



Uniwersytet Łódzki

Wydział Ekonomiczno-Socjologiczny

mgr Ewelina Niedzielska

Rozprawa doktorska

*Sentyment informacji niefinansowych jako czynnik wpływający na
reakcję rynku kapitałowego*

Rozprawa doktorska napisana pod kierunkiem promotor

dr hab. Joanny Fili

oraz promotora pomocniczego

dr. Pawła Sekuły

Łódź 2024

Spis treści	
Wstęp	4
Rozdział I – Informacja na rynku finansowym.....	10
1.1. Definicja informacji	12
1.2. Kluczowe aspekty informacji.....	15
1.3. Typy informacji na rynku finansowym	18
1.3.1. Mechanizm cenowy jako źródło informacji dla inwestorów	22
1.3.2. Informacja publiczna	27
1.3.3. Informacja prywatna	33
1.4. Asymetria informacji.....	37
1.5. Efekt momentum	41
Rozdział II – Odbiorca informacji na rynku finansowym	45
2.1. Wartość wewnętrzna instrumentu finansowego a jego cena	45
2.2. Inwestor w świetle neoklasycznej teorii finansów	48
2.2.1. Koncepcja oczekiwanej użyteczności	49
2.2.2. Stosunek inwestorów do ryzyka	51
2.2.3. Szacowanie prawdopodobieństwa	58
2.3. Inwestor w świetle behawioralnej teorii finansów	60
2.3.1. Teoria perspektywy.....	64
2.3.2. Inklinacje behawioralne inwestorów giełdowych	70
Rozdział III – Dziennikarz jako nadawca informacji na rynku finansowym	80
3.1. Kadrowanie jako immanentna cecha przekazywania informacji	82
3.2. Zewnętrzne czynniki selekcji informacji	89
3.3. Internalizowane czynniki selekcji informacji	94
3.3.1. Etyka.....	95
3.3.2. Wartości informacyjne.....	98
3.4. Wewnętrzne czynniki selekcji informacji	110
Rozdział IV – Sentyment tekstowy i jego zastosowanie w finansach	116
4.1. Big Data – ewolucja danych a proces badawczy	117
4.2. Eksploracja danych	124
4.3. Eksploracja danych a analiza tekstu.....	127
4.4. Język naturalny a eksploracja tekstu	131
4.5. Sentyment tekstowy	134

Rozdział V – Możliwość zastosowania eksploracji tekstu w finansach	146
5.1. Źródła danych tekstowych wykorzystywanych do badań w obszarze finansów .	149
5.2. Wykrywanie nadużyć	156
5.3. Analiza ryzyka.....	161
5.4. Zastosowanie analizy sentymentu, pochodzącego z publikacji prasowych, do badań na rynku finansowym	163
<i>Ameryka Północna</i>	166
<i>Ameryka Południowa</i>	168
<i>Azja</i>	169
<i>Europa</i>	171
Rozdział VI – Sentyment informacji niefinansowych jako zmienna objaśniająca reakcję ryнку	176
6.1. Metodyka i procedura badawcza.....	178
6.1.2. Opis procedury badawczej.....	179
6.1.3. Charakterystyka próby badawczej.....	199
6.1.4. Ograniczenia badawcze	203
6.2. Badania własne.....	204
6.2.1. Gromadzenie i przygotowanie danych	204
6.2.2. Wyniki pierwszego etapu badań	205
6.2.3. Wyniki drugiego etapu badań	208
6.2.4. Wyniki trzeciego etapu badań.....	214
6.3. Wnioski z badań	219
6.3.1. Podsumowanie uzyskanych wyników	219
6.3.2. Rekomendacje i dalsze kierunki badań	227
Zakończenie	230
Bibliografia	235
Spis tabel	257
Spis wykresów.....	257
Spis rysunków	258
Załączniki.....	259

Wstęp

Informacja i zapotrzebowanie na nią towarzyszy ludzkości od wieków. Dzięki informacji możliwe jest uczenie się, rozwój oraz adaptacja do nieustannie zmieniającego się otoczenia. Informacja zmniejsza poziom niewiedzy i motywuje do podejmowania określonych działań. Zapotrzebowanie na informację dotyczy każdej sfery życia: kulturowej, społecznej czy gospodarczej. Z perspektywy nauki o finansach informacja odgrywa rolę fundamentalną.

Analiza zachowań inwestorów, w reakcji na napływające na rynek informacje, stanowi przedmiot jednej z najpoważniejszych polemik, toczących się zarówno między reprezentantami różnych perspektyw badawczych, jak i praktykami rynkowymi. Jedną z najważniejszych koncepcji, mieszczących się w ramach tej polemiki, jest *Hipoteza Efektywności Informacyjnej*, zaprezentowana w latach 60. i 70. XX przez Eugena Fama. W jej ramach zakłada się, że nowe informacje, pojawiające się na rynku, zostają *w pełni i bardzo szybko* odzwierciedlone w cenach walorów. Z drugiej jednak strony, istnieją przesłanki, by sądzić, że sposób dyskontowania informacji przez uczestników rynku, charakteryzuje się asymetrią. Argumentów za takim podejściem dostarczają przedstawiciele paradygmatu behawioralnego, którzy wskazują, że inwestorzy obarczeni są różnego rodzaju uproszczeniami oraz błędami poznawczymi, co przekłada się na ich sposób reagowania na otrzymywane informacje. Pytanie o sposób dyskontowania informacji przez uczestników rynku finansowego, pozostaje więc wciąż kwestią nierozstrzygniętą, co stanowi przesłankę do podejmowania dalszych badań w tym obszarze.

Wskazywany problem wydaje się coraz bardziej istotny, z uwagi na to, że w ostatnich latach obserwowany jest znaczący wzrost liczby kreowanych, przetwarzanych oraz pobieranych informacji. Trend ten jest wynikiem powszechnego dostępu do Internetu, z którego obecnie w Polsce korzysta 88% populacji, przy czym dla 70% osób głównym powodem korzystania z sieci internetowej jest właśnie potrzeba poszukiwania informacji¹. Szacuje się, że tylko w ciągu ostatnich dwóch lat

¹ Zob. Kepios Pte. Ltd., *Digital 2023: Poland. The essential guide to the latest connected behaviours*, <https://indd.adobe.com/view/f37abde9-e4ac-4df0-8eef-3733deaa1cc2?allowFullscreen=true&wmode=opaque>. Dostęp: 1 kw. 2023.

wygenerowano 90% światowych zasobów danych, a w ciągu ostatnich 13 lat, liczba kreowanych danych wzrosła niemal 60-krotnie². W efekcie przedstawiciele różnych dyscyplin badawczych zaczęli podejmować próby eksploracji, powstających w ten sposób zbiorów, określanych mianem *Big Data*. Wydaje się, że rynki finansowe w naturalny sposób podążały za tym procesem. Widoczne jest to z jednej strony poprzez zmiany, jakie zachodzą w sposobie przekazywania danych, dotyczących obrotu instrumentami finansowymi, czy też zawierania transakcji. Jednocześnie zauważyć można, że niebagatelna część omawianych danych, prezentowana jest w postaci nieposiadającej struktury, czego egzemplifikacją są dane tekstowe. W ramach zagadnień, związanych z obszarem ekonomii i finansów, wskazać można przynajmniej dwie przyczyny tego zjawiska. Pierwszą jest wzrost znaczenia sprawozdawczości, którą określić można jako niefinansową. Drugą natomiast – wspomniany już – powszechny dostęp do Internetu, czego konsekwencją jest zwiększenie liczby źródeł informacji, związanych z analizą bieżącej sytuacji rynkowej. Dlatego też, z badawczego punktu widzenia, ważne wydaje się wypracowanie metod i narzędzi, umożliwiających eksplorację danych o charakterze tekstowym.

Powyższy problem mieści się w obszarze definiowanym jako przetwarzanie języka naturalnego, przez co rozumieć można wykorzystanie technologii komputerowej – w tym programowania – do analiz, poszukujących istotnych informacji jakościowych lub ilościowych w języku mówionym, lub pisany³. Ze względu na dynamiczny rozwój analizy tekstu w badaniach z obszaru finansów, w najnowszej literaturze anglojęzycznej coraz częściej – na określenie tego nurtu badawczego – używa się terminu „prognozowanie finansowe oparte na języku naturalnym” (ang. *Natural Language Based Financial Forecasting* (NLFF))⁴.

Efektom rozwijającego się zainteresowania tą tematyką jest rosnąca liczba publikacji, w ramach których podejmowane są próby wykorzystania zmiennych o charakterze tekstowym, do objaśniania reakcji inwestorów giełdowych.

² Zob. Duarte, F., *Amount of Data Created Daily (2023)*, <https://explodingtopics.com/blog/data-generated-per-day#how-much>, Dostęp: 1 kw. 2023.

³ Das. S. R., *Text and Context: Language Analytics in Finance. Foundations and Trends in Finance*, now Publishers Inc, Hanover 2014, s. 6.

⁴ Xing, F. Z., Cambria, E., Welsch, R. E., *Natural language based financial forecasting: a survey*, „Artificial Intelligence Review”, 2018. nr 50. s. 49–73.

Przeprowadzony przez autorkę przegląd literatury był podstawą do wysnucia wniosku, że rozwój badań z zakresu analizy tekstu na rynkach wschodzących można uznać za relatywnie słaby w porównaniu z rynkami rozwiniętymi. Powyższe obserwacje stanowiły przesłankę do stwierdzenia, że w omawianej tematyce istnieje luka badawcza. Stąd niniejsza dysertacja realizuje dwa cele: poznawczy i empiryczny.

Celem dysertacji w wymiarze poznawczym jest zidentyfikowanie oraz scharakteryzowanie kluczowych paradygmatów, opisujących sposoby dyskutowania informacji przez uczestników rynku finansowego.

Celem pracy w wymiarze empirycznym jest zbadanie i ocena sposobu dyskutowania informacji przez inwestorów giełdowych, z punktu widzenia dodatkowej perspektywy, jaką jest sentyment artykułów prasowych, na temat spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie.

Realizacji przyjętego celu pracy posłużyć ma weryfikacja hipotezy głównej oraz hipotez pomocniczych, które zostały sformułowane następująco:

Hipoteza główna (H): *Rodzaj sentymentu, zawarty w danych tekstowych, jest odzwierciedlony w reakcji rynku.*

Hipotezy pomocnicze:

H1: *Anglojęzyczne artykuły prasowe, dotyczące spółek notowanych w indeksie WIG20 w okresie od 01.01.2013 do 31.12.2022, charakteryzują się przewagą sentymentu pozytywnego względem sentymentu negatywnego.*

H2: *Występuje asymetria w sposobie reakcji rynku na pozytywny i negatywny sentyment, zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20.*

H3: *Sentyment zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20, stanowi determinantę reakcji inwestorów giełdowych.*

Prezentowana praca ma charakter teoretyczno–empiryczny i składa się z sześciu rozdziałów. W rozdziale pierwszym autorka scharakteryzowała rolę, jaką informacja odgrywa na rynkach finansowych. W tym celu przybliżone zostały kluczowe definicje informacji, jak również jej cechy. Omówione zostały także typy informacji na rynku

finansowym, przez co rozumieć należy mechanizm cenowy, informację publiczną, prywatną oraz poufną.

Istotą drugiego rozdziału było przybliżenie dwóch perspektyw, dotyczących odbiorców informacji na rynku finansowym. Pierwszą z nich jest, związane z hipotezą efektywności informacyjnej, założenie racjonalności uczestników rynku. Druga perspektywa ukazuje inwestora w świetle paradygmatu finansów behawioralnych. Autorka wskazała podstawowe różnice, występujące między oboma paradygmatami.

W rozdziale trzecim scharakteryzowano nadawców informacji, których – w ramach niniejszej dysertacji – rozumie się jako dziennikarzy, zajmujących się tematyką rynków finansowych. Autorka dysertacji zwróciła szczególną uwagę na dwa kluczowe aspekty, mogące mieć potencjalny wpływ, na sposób publikowania informacji przez dziennikarzy. Są nimi kadrowanie oraz czynniki selekcji informacji.

W rozdziale czwartym omówione zostały pojęcia, związane z pojawieniem się nowego typu danych, jakimi są *Big Data*. Scharakteryzowany został również proces ich eksploracji. W tym kontekście przedstawiono problematykę analizy tekstu oraz pracy z językiem naturalnym. W rozdziale tym zaprezentowane zostało także pojęcie sentymentu, które w niniejszej dysertacji, rozumiane będzie jako pozytywne lub negatywne nacechowanie publikacji dziennikarskich, na temat rynków finansowych. Autorka dysertacji wskazała na konieczność ujmowania tego zjawiska w ramach mechanizmu rynkowego i definiowania poprzez klarowne wskazywanie miejsca jego powstawania, w procesie od zdarzenia rynkowego do podjęcia przez inwestorów odpowiednich decyzji. Koncepcja ta stanowi autorski wkład w rozwój wiedzy w tym obszarze i wpisuje się problematykę ekonomii i finansów.

W rozdziale piątym dokonano przeglądu dotychczas prowadzonych badań, związanych z tematyką wykorzystania, zawartego w artykułach prasowych, sentymentu do objaśnienia reakcji inwestorów giełdowych. Ocena aktualnego dorobku naukowego, dotyczyła piśmiennictwa światowego, co wymagało pogłębionych studiów literaturowych w języku angielskim. Autorka przedstawiła możliwie szerokie zastosowanie tego typu analiz, zarówno pod względem geograficznym, jak

i tematycznym. Takie podejście pozwoliło pokazać różne obszary, w ramach których stosowana jest koncepcja analizy tekstu.

W rozdziale szóstym zaprezentowane zostały wyniki badań własnych, przeprowadzonych na podstawie próby badawczej składającej się z 35 podmiotów, notowanych w indeksie WIG20 w okresie od 01.01.2013 do 31.12.2022 roku. Badanie realizowane było z wykorzystaniem ponad 56 tysięcy anglojęzycznych artykułów prasowych i składało się z trzech etapów. W etapie pierwszym przeprowadzona została ekstrakcja, zawartego w anglojęzycznych artykułach prasowych, sentymentu. W tym kontekście zastosowane zostało podejście słownikowe, do którego wybrano – popularny w obszarze badań z zakresu finansów i ekonomii – słownik *Loughran-McDonald Master Dictionary*⁵. Celem etapu drugiego była weryfikacja drugiej hipotezy pomocniczej, poprzez zastosowanie metody analizy zdarzeń. Analizie podlegała reakcja inwestorów giełdowych na otrzymanie informacji, cechujących się skrajnym sentymentem pozytywnym oraz negatywnym, w oknie zdarzenia od dnia poprzedzającego pojawienie się artykułu prasowego, do pięciu dni po pojawieniu się artykułu. Siła owej reakcji mierzona była poprzez oszacowanie skumulowanych średnich ponadnormatywnych stóp zwrotu. W ramach ostatniego etapu weryfikacji podlegała trzecia hipoteza pomocnicza. W tym celu zbudowany został model logitowy, w którym zmiennymi objaśniającymi dzienne stopy zwrotu z akcji były wartości pozytywnego i negatywnego sentymentu, zawartego w publikacjach prasowych. Wybór tej metody podyktowany był jej powszechnością w dotychczas prowadzonych – w omawianej tematyce – badaniach na rynkach rozwiniętych. Empiryczna część pracy zwieńczona została wnioskami, w których autorka podsumowała uzyskane wyniki, na tle piśmiennictwa światowego oraz polemiki, toczącej się między przedstawicielami paradygmatu finansów neoklasycznych oraz nurtu behawioralnego. Jednocześnie wskazane zostały potencjalne kierunki badań w analizowanym obszarze.

Zgodnie z wiedzą autorki jak dotąd, nie przeprowadzono badań, pozwalających ocenić znaczenie sentymentu, zawartego w artykułach prasowych, dla objaśnienia dziennych stóp zwrotu spółek notowanych na Giełdzie Papierów Wartościowych

⁵ Zob. Loughran, T., McDonald, B.: *Loughran-McDonald Master Dictionary*. <https://sraf.nd.edu/loughranmcdonald-master-dictionary/>, Dostęp: 1 kw. 2023.

w Warszawie. Biorąc pod uwagę rosnącą liczbę analiz, dotyczących tego tematu, które realizowane są na rynkach zachodnich, stwierdzić można, że na krajowym gruncie istnieje luka badawcza, którą autorka starała się wypełnić.

Rozdział I – Informacja na rynku finansowym

Jednym z kluczowych problemów niniejszej pracy jest prawidłowe zdefiniowanie pojęcia informacji. Zagadnienie to stanowi teoretyczny fundament wielu dyscyplin, co przyczynia się do istnienia obfitej literatury związanej z tym tematem. Stwierdzić można, że przedstawiciele owych dyscyplin czerpią wiedzę od siebie nawzajem, systematycznie wzbogacając dotychczasowy dorobek naukowy. Jednocześnie jednak, wielość perspektyw powoduje trudności związane z uchwyceniem meritum przedmiotu. Z tego powodu – w pierwszej kolejności – nieodzowne staje scharakteryzowanie znaczenia oraz istoty informacji, zarówno w kontekście rynku finansowego w ogóle, jak i w kontekście analizowanego problemu badawczego.

Próba zdefiniowania zagadnienia i roli informacji w warunkach rynkowych jest tym trudniejsza, że w literaturze z obszaru finansów często stosuje się ją wymiennie z takimi określeniami jak „dane” czy „wiedza”. Taki skrót myślowy pozornie wydaje się uprawniony. Jednym z kluczowych źródeł wykorzystywanych w badaniach na temat rynku są wszak wartości wolumenów czy cen walorów, które w praktyce nie są niczym innym, jak zbiorem danych, na podstawie których pozyskać można wiedzę o dynamice giełd. Mimo to – zdaniem autorki – niejasność granic między powyższymi pojęciami utrudniać może określenie miejsca i istoty informacji na rynku finansowym. Dlatego też w tej części pracy zaproponowany zostanie sposób rozróżniania powyższych terminów i sprowadzenia pojęcia informacji do podstawowej definicji, która wykorzystywana będzie w dysertacji.

Z punktu widzenia nauki o finansach informacja odgrywa rolę niemal konstytutywną. Analiza jej wpływu na zachowanie inwestorów oraz dynamikę cen walorów i ruchów na rynku stanowi bowiem oś jednej z najbardziej zagorzałych dyskusji, jaka toczy się między przedstawicielami różnych nurtów badawczych oraz praktykami. Jak można sądzić, punktem wyjścia tej polemiki jest, zaproponowana przez Eugena Fama, *hipoteza efektywności informacyjnej* (ang. *efficient market hypothesis*), dotycząca sposobów kształtowania się cen aktywów wskutek otrzymywanej przez

inwestorów informacji⁶. *Rynek efektywny informacyjnie* może być potocznie definiowany jako taki, w którym nowa informacja, związana z daną spółką (lub jej otoczeniem), jest w pełni i bardzo szybko odzwierciedlana w cenach akcji. W ten sposób, w dowolnym momencie poprzez cenę waloru przedstawiane są wszystkie dostępne inwestorom komunikaty, a nowa informacja wywołuje zmianę ceny w sposób skokowy, do takiego poziomu, który odzwierciedlać będzie jej treść⁷. Uruchomienie takiego mechanizmu możliwe jest w sytuacji, w której każdy z uczestników gry rynkowej ma zapewniony szybki dostęp do tejże informacji⁸. Przedmiotem sporu pozostają jednak takie kwantyfikatory jak *szybko* oraz *w pełni*. Dlatego też konieczne wydaje się zarówno zaprezentowanie samej idei rynku efektywnego informacyjnie, jak i związanych z nią wątpliwości.

Dodatkowym argumentem przemawiającym za koniecznością kontynuowania rozważań na temat informacji w przestrzeni rynku finansowego jest fakt, że opierające się o nią teorie najczęściej łączą się również z problemem predykcji cen walorów i potencjalnych stóp zwrotu. Na podstawie prezentowanych koncepcji dojść można bowiem do wniosku, że wszelkie próby antycypacji dynamiki rynku mogą być uznane za bezzasadne. Taką perspektywę odnaleźć można zwłaszcza w idei efektywności informacyjnej, czy też w modelu błędzenia losowego. Z drugiej jednak strony dotychczasowe badania, w tym te przeprowadzone na polskim rynku kapitałowym, oraz przegląd literatury przedmiotu sugerują istnienie rynkowych anomalii takich jak efekt momentum lub dryf poinformacyjny, które stoją w sprzeczności z założeniami hipotezy efektywności informacyjnej. Celem autorki nie jest oczywiście udzielenie jednoznacznej odpowiedzi na pytanie o możliwość lub brak możliwości dokonywania predykcji cen instrumentów finansowych, a raczej przedstawienie argumentów, które przyświecają zwolennikom obu punktów widzenia.

⁶ Snarska, M., *Efektywność informacyjna giełdowych rynków akcji*, [w:] *Efektywność giełdowego rynku akcji w Polsce z perspektywy dwudziestolecia*, (red.) Czekaj J., PWE, Warszawa 2014, s. 39.

⁷ Doman, M., *Mikrostruktura giełd papierów wartościowych*, Wydawnictwo Uniwersytetu w Poznaniu, Poznań 2011, s. 23.

⁸ Żurek, M., *Inklinacje behawioralne na rynkach kapitałowych w świetle modeli SEM*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2016, s. 16.

1.1. Definicja informacji

W celu podjęcia próby określenia istoty informacji autorka proponuje wyjść od jednej z najbardziej ogólnych definicji, stworzonej przez Jacka Rochestera, który twierdzi, że „informacja to nic innego jak uporządkowany zbiór faktów i danych”⁹. Wynika z tego, że dowolny zbiór faktów i danych o rzeczywistości stanowiłby informację, o ile tylko jego struktura byłaby w jakiś sposób uporządkowana. Powyższą koncepcję zestawień można z definicją zaproponowaną przez profesor Barbarę Będowską-Sójkę, która – patrząc na informację przez pryzmat komunikatu – pisze:

Pojęcie informacji w kontekście rynku kapitałowego odnosi się przede wszystkim do wszelkich komunikatów werbalnych informujących o zdarzeniach, które mogą mieć wpływ na wycenę rynkową instrumentów finansowych. Informacjami są komunikaty o stosunkowo jednoznacznej treści, w postaci publikowanych raportów, wskaźników, wypowiedzi członków zarządu etc., przekazywane inwestorom, które odnoszą się bezpośrednio do zdarzeń zachodzących na rynkach kapitałowych, w spółkach i ich otoczeniu gospodarczym [...] ¹⁰.

Pomijając problematyczne określenie „komunikatu werbalnego”, wydaje się, że definicja ta trafnie odzwierciedla sedno tego, czym jest informacja na rynku finansowym. Należy jednak zwrócić uwagę, że pojawia się w niej jedno założenie, które w ramach badań nad zachowaniem inwestorów, sprawia wrażenie fundamentalnego. Mowa tu bowiem o takim komunikacie, który będzie miał wpływ na wycenę instrumentów finansowych. Za Jackiem Rochesterem przyjąć należałoby, że dobrze znane odbiorcy fakty o rzeczywistości, jeśli wpisują się w przedstawione przez niego kryterium, mają dla niego walor informacyjny. Barbara Będowska-Sójka zdaje się nie zgadzać z takim uproszczeniem.

Po pierwsze, patrząc na problem ogólnie, można przypuszczać, że nie każdy komunikat będzie miał wpływ na ceny walorów. Z perspektywy pojedynczego inwestora wątpliwe jest natomiast, że każdy komunikat będzie miał dla niego walor informacyjny. Trudno nie zadać więc pytania o to, które komunikaty można by utożsamiać z informacją, a które nie.

⁹ Burgin, M., *Theory Of Information: Fundamentality, Diversity And Unification*, World Scientific, Los Angeles 2009, s. 194. [za: Rochester, J. B., 1996].

¹⁰ Będowska-Sójka, B., *Wpływ informacji na ceny instrumentów finansowych. Analiza danych śróddziennych*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2014, s. 19.

Częściową odpowiedź na to pytanie uzyskać można, przesuwając środek ciężkości z poziomu samej informacji na poziom jej odbiorcy. Jeśli bowiem dwóch inwestorów, rozpatrujących możliwości ulokowania kapitału na tym samym rynku, otrzyma ten sam komunikat w dokładnie tym samym czasie, to czy komunikat ten będzie dla obu wyposażony w taki sam walor informacyjny? Dla przykładu przyjąć można, że jeden z nich możliwość taką analizuje po raz pierwszy, a drugi jest na tym rynku aktywny od wielu lat. Przywołać można tu stwierdzenie Maxa Boisota, który tłumaczy, że „informacją jest to, co obserwator wydobędzie z danych jako funkcja jego lub jej oczekiwań, lub przeszłej wiedzy”¹¹. Boisot sugeruje więc, że informacja jest w pewien sposób zależna od jej odbiorcy oraz że na zależność tę wpływ ma jego dotychczasowa wiedza. Mając to w pamięci, nie można jednak odpowiedzieć na pytanie, dla którego z przykładowych inwestorów komunikat będzie miał większą wartość informacyjną. Naturalnie założyć można, że doświadczony inwestor będzie w stanie szybciej i lepiej przetworzyć komunikat (np. na podstawie swojej dotychczasowej praktyki lub terminologicznej biegłości), a tym samym informatywność tego komunikatu będzie dla niego większa. Wątpliwości budzić może tu jednak stwierdzenie Deode Nauty:

informacja to nowość: to, co jest już wiadome, nie jest informacją. Coś może być informacją, do tego stopnia, w jakim jest nieznane, nieoczekiwane, zaskakujące lub nieprawdopodobne¹².

Takie definiowanie omawianego zagadnienia bliskie jest badaczom finansów z zakresu teorii sygnału. Warto jednak zaznaczyć, że w ramach tej literatury informacja rozumiana może być jako element składowy sygnału, jak również sygnał traktowany może być jako mniejsza część informacji. Przykładowo, Stefan Mynarski zakłada, że ilość informacji, jaką zawiera dany sygnał, zależy od poziomu niewiedzy, którą charakteryzuje się odbiorca, tj. im większa zajdzie w tym aspekcie zmiana w efekcie odebrania sygnału, tym większą informatywnością cechuje się sygnał¹³. Z kolei, wspomniana już, Barbara Będowska-Sokój twierdzi, że sygnał należy rozumieć jako tę

¹¹ Boisot, M. H., *Knowledge Assets: Securing Competitive Advantage in the Information Economy*, Oxford University Press, Oxford 1998, s. 49.

¹² Nauta, D., *The Meaning of Information*, Mouton & Co. The Hague, Paryż 1972, s. 19.

¹³ Mynarski, S., *Elementy teorii systemów i cybernetyki*, PWN, Warszawa 1979, s. 155.

część informacji, która jest nieoczekiwana¹⁴. Niezależnie jednak od tego, jak strukturyzujemy relację zachodzącą między sygnałem a informacją, z powyższych rozważań wysnuć można dwa istotne wnioski. Po pierwsze, informacja nie jest tożsama z komunikatem, który należałoby rozumieć raczej jako treść informacji. Po drugie, dla oceny waloru informacyjnego danego komunikatu, niezbędne jest włączenie w proces postaci odbiorcy.

Z uwagi na wymienione przesłanki, autorka proponuje rozumieć informację jako proces, którego immanentnym elementem jest nie tylko sam komunikat (w rozumieniu treści), ale również jego nadawca i odbiorca. Tak szerokie spojrzenie na informację pozwala głębiej przyjrzeć się związanym z nią problemom. Po pierwsze, należy zastanowić się kim – zwłaszcza w kontekście rynku kapitałowego – jest nadawca komunikatu. W najprostszym ujęciu można oczywiście stwierdzić, że takowym nadawcą jest emitent, który kieruje komunikat do swoich akcjonariuszy. Należy jednak zauważyć, że w codziennej praktyce rynkowej, autorami komunikatów są na przykład analitycy biur maklerskich czy dziennikarze. Co więcej, z regularnie prowadzonych przez Stowarzyszenie Inwestorów Indywidualnych badań wynika, że częstym źródłem informacji o wydarzeniach rynkowych i spółkach są na przykład ogólnodostępne portale internetowe¹⁵. Katalog nadawców zdaje się więc ulegać systematycznemu rozszerzaniu, a założenie, że jedynymi nadawcami komunikatów są spółki lub nawet szerzej – zorganizowane podmioty rynkowe – wydaje się coraz trudniejsze do podtrzymania. Należy raczej stwierdzić, że spółki emitujące papiery wartościowe są obecnie autorami jedynie części komunikatów, jakie otrzymują inwestorzy.

Kolejno, sama istota komunikatu wydaje się wymagać zdefiniowania. Na tej płaszczyźnie bowiem rozgrywają się przynajmniej dwie poważne polemiki, które autorka będzie starała się przybliżyć w dalszej części pracy. Po pierwsze, sam fakt otrzymywania przez inwestorów komunikatu nie jest oczywisty, z czym wiąże się, szeroko dyskutowane w ramach literatury przedmiotu, zagadnienie asymetrii informacji. Drugi punkt sporny

¹⁴ Będowska-Sójka, B., *Wpływ informacji...*, op.cit., s. 19.

¹⁵ Zob. Żuławski, M., *Oto najpopularniejsze źródła informacji dla inwestorów [OBI 2022]*, <https://www.sii.org.pl/15945/aktualnosci/badania-i-rankingi/oto-najpopularniejsze-zrodla-informacji-dla-inwestorow-obi-2022.html> [dostęp: 16.02.2023].

opiera się o samą formę komunikatów. Dla analityków, technicznych komunikatem, któremu nadadzą walor informacyjny, będzie na przykład zmiana poziomu cen instrumentów finansowych. Z kolei dla analityków fundamentalnych, komunikat może przybierać formę raportu okresowego spółki, jak również artykułu o stanie sektora, którym są zainteresowani.

Trzecie zagadnienie wiąże się z odbiorcą komunikatu. Należy w tym miejscu przyznać, że wcześniejsza odpowiedź na pytanie o to, dla którego z przykładowych inwestorów komunikat będzie charakteryzował się wyższym stopniem informatywności, nie jest pełna. Może zdarzyć się bowiem tak, że dla pierwszego z nich nadana treść będzie po prostu niezrozumiała. Informacją, jak twierdzi Carl Friedrich von Weizsäcker, może być tylko to, co może zostać zrozumiane¹⁶. Takie założenie rozszerza dyskusję o próbę odpowiedzi na pytanie: jak przebiega proces rozumienia inwestora, a nawet szerzej – jak przebiega jego proces uczenia się. Jeżeli „rozumienie” pojmować będziemy w najszerszym znaczeniu – jako proces kognitywny¹⁷, to trudno poprzestać jest na potocznej definicji informacji, utożsamiającej ją z komunikatem. Dzieje się tak, chociażby ze względu na to, że procesy poznawcze obarczone są czynnikami emocjonalnymi, uprzedzeniami oraz uproszczeniami. Innymi słowy, istnieje wiele przesłanek by sądzić, że ten sam komunikat będzie niósł dla inwestorów zupełnie różne wartości informacyjne. Zresztą w ramach tej problematyki od lat w środowisku akademickim toczy się dyskusja. Zwolennicy hipotezy efektywności rynku *implicite* zakładają racjonalność zachowania inwestorów giełdowych. Z kolei badacze z zakresu finansów behawioralnych doszukują się w dynamice rynku przejawów „irracjonalności” jego uczestników.

1.2.Kluczowe aspekty informacji

W ogólnym ujęciu, pojęcie informacji rozpatrywać można na trzech płaszczyznach: ilościowej, semantycznej oraz pragmatycznej¹⁸. Na potrzeby dalszych rozważań, autorka niniejszej dysertacji zarysuje każdy z tych aspektów. Początki naukowej refleksji nad

¹⁶ von Weizsäcker, C. F. F., *The Structure of Physics*, Springer, Dordrecht 2006, s. 229.

¹⁷ Por. Burgin, M., *Theory of Information: Fundamentality, Diversity and Unification*, World Scientific, Los Angeles 2009, s. 189.

¹⁸ Zob. Woleński, J., *Informacja i semantyka*, „Filozofia Nauki”, 1997, 5/1, 59–64.

informacją związane są z teorią ilościową, którą łączyć należy – z jednej strony z zagadnieniami natury czysto technicznej, a z drugiej – z modelami probabilistyczno-statystycznymi oraz entropią informacyjną. Za autora tego nurtu badawczego powszechnie uznaje się Claude’a Shannona, który w 1948 roku opublikował pracę pt. *A mathematical theory of communication*. Należy zaznaczyć, że osią problemu Shanona nie była informacja w rozumieniu przekazywanej treści, a informacja pojmowana jako jej ilość, do wyznaczenia której wykorzystywane było prawdopodobieństwo (w tym przypadku oparte na częstotliwości pojawiania się znaków). Shannon w swojej pracy nie podał precyzyjnej definicji informacji, ale umiejscowił ją w kontekście niepewności. Tym samym, informację można wiązać ze stopniem nieokreśloności, która będzie się zmniejszać w efekcie przekazywania wiadomości¹⁹. Z kolei miarę tejże niepewności określił mianem entropii²⁰. Teoria ilościowa w ramach nauk społecznych może wydawać się jednak najmniej użyteczna²¹. Próby jej przełożenia na te dziedziny określa się wręcz jako niemożliwe, z uwagi na to, że w rzeczywistości społecznej występować mogą sytuacje, które nie mieszczą się w modelu probabilistycznym (choćby ze względu na nieznaną rozkładu prawdopodobieństwa)²². Mimo to, kanadyjski badacz Jing Chen w swojej pracy z 2004 roku stwierdził, że reguły matematyczne, które prezentował Shannon, dotyczą nie tylko systemów komunikacyjnych, ale odnoszą się również do wszystkich organizmów żywych, w tym do człowieka. Tym samym istnieje możliwość zastosowania teorii ilościowej do obserwacji rynków finansowych. Na tej podstawie Chen wysnuł trzy istotne wnioski. Pierwszy związany jest z kosztem pozyskania informacji. Im większa jest wartość informacji, tym jej pozyskanie jest bardziej kosztowne. Po drugie, wartość informacji jest odwrotnie proporcjonalna do liczby osób, które są w stanie ją zrozumieć, przy czym mowa tu zarówno o rozumieniu mechanizmów giełdowych, jak i technologii produkcji lub potencjału popytowego produktów. Po trzecie, ilość informacji, jaką otrzyma inwestor, może być zdefiniowana jako ilość informacji wygenerowanej

¹⁹ Wędrawska, E., *Ilościowe i jakościowe koncepcje informacji*, „Ekonomiczne Problemy Usług” 2010, 57, 49–56, s. 50.

²⁰ Shannon, Cl. E., *A Mathematical Theory of Communication*, „The Bell System Technical Journal”, 1948, 27/3, 379–423, s. 20.

²¹ Zob. Burgin, M., *Theory of Information...*, op.cit., s. 6.

²² Wędrawska, E., *Ilościowe i jakościowe...*, op.cit. s. 51.

i pomniejszonej o jej dwuznaczność. Ową dwuznaczność Chen utożsamia z poziomem asymetrii informacji, która wyznacza związek między źródłem informacji, a jej odbiorcą. Innymi słowy to, ile informacji może uzyskać dany inwestor zależne jest od jego wiedzy. Dlatego też proces rozumienia informacji powinien być umieszczany w szerokim kontekście procesu uczenia się²³.

O ile w pierwszym wrażeniu teoria ilościowa odnosi się raczej do kwestii technicznych (dotyczących procesu transmisji), o tyle aspekt semantyczny związany jest komunikowaniem treści²⁴. Trudno mówić tu jednak o wspólnej teorii. Należałoby raczej rozpatrywać semantykę jako zbiór koncepcji, którego wspólnym mianownikiem jest koncentracja na znaczeniowej stronie wiadomości²⁵. Ujmując rzecz inaczej, semantyczny aspekt informacji nawiązuje do treści wypowiedzi²⁶. Jednym z założeń tych koncepcji jest posiadanie przez odbiorcę komunikatu określonego tezaurya pojęciowego, to jest zbioru znaczeń, opisów i charakterystyk słów, najczęściej w obrębie tego samego języka²⁷. W tym rozumieniu informację zawierają te wiadomości, które wywołują zmianę w teaurusie adresata²⁸. Informacja pojmowana jest więc „jako system relacyjny, w którym zmiana tezaurya T następuje pod wpływem wiadomości ze zbioru W odebranej przez odbiorcę ze względu na cel, jaki chce on osiągnąć dzięki otrzymanej wiadomości”²⁹. W ten sposób perspektywa semantyczna nawiązuje do teorii sytuacji³⁰, w ramach której zakładane jest istnienie pewnego otoczenia informacyjnego, gdzie transmisja wiadomości możliwa jest dzięki temu, że między różnymi sytuacjami istnieją zależności³¹. W efekcie aspekt semantyczny uznać można za kluczowy z punktu widzenia funkcji poznawczych, co pozwala rozumieć informację jako czynnik wiedzotwórczy³². Autorzy zwracają więc w tym kontekście uwagę zarówno na treść

²³ Chen, J., *Information Theory and Market Behavior*, „ICFAI Journal of Behavioral Finance”, 2005, 2/4, 25–45, s. 27.

²⁴ Burgin, M., *Theory of Information...*, op.cit., s. 301.

²⁵ Zob. Wędrowska, E., *Ilościowe i jakościowe...*, op.cit., s. 52.

²⁶ Będowska-Sójka, B., *Wpływ informacji...*, op.cit., s. 17 [za: Szaniawski 2987, s. 249].

²⁷ Kowalczyk, E., *O istocie informacji*, Wydawnictwo Komunikacji i Łączności, Warszawa 1986, s. 91.

²⁸ Wędrowska, E., *Ilościowe i jakościowe...*, op.cit., s. 54.

²⁹ Ibidem, s. 52.

³⁰ Zob. Dretske, F., *Knowledge and the Flow of Information*, MIT Press, Cambridge 1981.

³¹ Szczepański, R., *Semantyczne Teorie Informacji*, „Studia Metodologiczne”, 2015, 34, s. 62.

³² Machlarz, A., *O ogólnej definicji pojęcia informacji semantycznej*, „Filozofia Nauki”, 2011, 19/1, 71–93, s. 72.

komunikatu, jak i na zaplecze wiedzy, jakim dysponuje odbiorca. To z kolei prowadzi do konstatacji, że informacja w ujęciu semantycznym może być nacechowana pewną dawką subiektywizmu³³. Takie założenie ponownie prowadzić może do wniosku, że identyczny sygnał przekazany dwóm różnym odbiorcom, może być zinterpretowany przynajmniej na dwa różne sposoby³⁴. Ten tok myślenia wydaje się wpisywać w dyskusję na temat sposobu interpretowania komunikatów przez inwestorów giełdowych.

Brakującym elementem w powyższej deskrypcji jest trzeci – pragmatyczny – aspekt informacji, który odnosi się do jej wpływu na działanie odbiorcy komunikatu³⁵. Istotna jest oczywiście nie tylko treść wiadomości, ale również konsekwencje jej wystąpienia dla przebiegu procesu decyzyjnego. Podstawowym problemem rozpatrywanym w ramach prezentowanej rozprawy, jest wszak pytanie o to, jaki wpływ na dynamikę rynków finansowych ma informacja oraz czy semantyczny subiektywizm odbioru jej treści może wpływać na jej aspekt pragmatyczny. Tak postawione pytania stanowić będą podstawę dla dalszych rozważań.

1.3. Typy informacji na rynku finansowym

Jednym z częściej dyskutowanych zagadnień w ramach nauk ekonomicznych jest problem efektywności. Mówi się o nim zarówno w kontekście makroekonomicznym, zarządczym, jak i w kontekście rynków finansowych. W pierwszym ujęciu definiuje się ją poprzez pojęcie efektywności Pareta, z którą to „mamy do czynienia, gdy nie istnieje możliwość reorganizacji produkcji lub dystrybucji w taki sposób, by komuś działało się lepiej, bez pogorszenia sytuacji innej osoby”³⁶. W tym sensie, efektywność opisuje stan, który określić można jako optymalny, to znaczy taki, w którym poziom dobrobytu społecznego w danych warunkach jest maksymalny³⁷. W ujęciu zarządczym rozumiana jest ona jako działanie związane z gospodarowaniem zasobami, które pozwala – przy ich minimalnym wykorzystaniu – na maksymalne realizowanie postawionych celów³⁸.

³³ Wędrowska, E., *Ilościowe i jakościowe...*, op.cit., s. 53.

³⁴ Będowska-Sójka, B., *Wpływ informacji...*, op.cit., s. 17 [za: Szaniawski 2987, s. 249].

³⁵ Burgin, M., *Theory of Information...*, op.cit., s. 307.

³⁶ Samuelson, P. A., Nordhaus, W. D., *Ekonomia*, Dom Wydawniczy Rebis, Poznań 2017, s. 161.

³⁷ Podstawka, M., (red.), *Finanse. Instytucje, instrumenty, podmioty, rynki, regulacje*, PWN, Warszawa 2017, s. 20.

³⁸ Ciołek, M., *Efektywność informacyjna polskiego rynku akcji*, CeDeWu, Warszawa 2015, s. 15 [za: Frąckiewicz].

Dla uwarunkowań związanych z rynkiem finansowym efektywność rozpatrywana jest na trzech komplementarnych płaszczyznach: alokacyjnej, transakcyjnej oraz informacyjnej. Jak twierdzi Oskar Starzeński te trzy formy nie mogą występować w systemie gospodarczym niezależnie. Korzystna alokacja kapitału nie może wszak odbywać się ani bez sprawnie funkcjonującego systemu przepływu informacji, ani bez odpowiednio niskich kosztów pośrednictwa transakcji rynkowych³⁹.

Pierwsza płaszczyzna związana jest z przepływem kapitału między podmiotami dysponującymi jego nadwyżkami a podmiotami, które go poszukują w celu sfinansowania swoich planów inwestycyjnych⁴⁰. Wiesław Dębski dokonuje w tym kontekście rozróżnienia na proces alokacji oraz relokacji kapitału. Alokacja kapitału przebiega na rynku pierwotnym. Podmiot, chcący pozyskać kapitał, emituje w tym celu papiery wartościowe. Pozyskane dzięki temu środki tworzą potencjał sfinansowania celów rozwojowych (np. zwiększenia zdolności produkcyjnych). Tym samym nie mogą być one traktowane jako bezpośrednie powiększenie zysku ani przeznaczone na działalność konsumpcyjną. Jednostki posiadające nadwyżkowe środki poszukują natomiast takich możliwości inwestycyjnych, które zapewnią im satysfakcjonującą stopę zwrotu. W efekcie, będą starały się przeznaczać pieniądze na przedsięwzięcia, które są w stanie takową stopę zwrotu zapewnić. Mechanizm ten w literaturze przedmiotu definiowany jest jako alokacyjna funkcja rynku finansowego. Na rynku wtórnym dokonywana jest natomiast relokacja wcześniej zainwestowanych środków. Stanowi to konsekwencję ciągłej wyceny, jaka ma miejsce na giełdach oraz – w szerokim ujęciu – zmieniających się warunków gospodarczych. W tym procesie niektóre podmioty podejmują decyzje o „wyjściu” z danej inwestycji, a inne – w oczekiwaniu na wyższy zwrot z zaangażowanych środków – o „wejściu”. W ten sposób kapitał transferowany jest do przedsięwzięć charakteryzujących się najwyższą rentownością⁴¹.

O efektywności transakcyjnej można mówić wówczas, gdy konkurencyjność pośredników rynkowych oraz infrastruktura techniczna, wpływają na obniżanie kosztów

³⁹ Starzeński, O., *Analiza rynków finansowych*, Wydawnictwo C.H. Beck, Warszawa 2011, s. 41.

⁴⁰ Podstawka, M., *Finanse. Instytucje...*, op.cit., s. 567.

⁴¹ Dębski, W., *Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki*, Wydawnictwo PWN, Warszawa 2014, s. 110–111.

wymiany handlowej oraz skrócenie czasu niezbędnego do jej przeprowadzenia między inwestorami⁴². W tak sformułowanej definicji w istocie poruszone są dwa zagadnienia. Po pierwsze, chodzi tutaj o poziom kosztów, a po drugie, o tempo realizacji umów. Im wyższy poziom efektywności transakcyjnej danego rynku, tym zarówno koszty, jak i czas realizacji transakcji powinny przyjmować niższe wartości. Warunków do wystąpienia takiego stanu rzeczy powszechnie upatruje się w zagadnieniu konkurencyjności. Zakłada się tutaj, że rywalizujący ze sobą agenci wymuszają na sobie nawzajem obniżanie kosztów transakcji⁴³. O pełnej tego typu efektywności można by więc mówić wówczas, gdy między pośrednikami zachodziłaby również pełna konkurencyjność cenowa⁴⁴.

Trzecia płaszczyzna efektywności rozpatrywana jest między innymi na tle zdolności inwestorów do pozyskiwania i przetwarzania informacji oraz podejmowania na tej podstawie decyzji inwestycyjnych. W tym kontekście kluczową – zwłaszcza dla rozważań o roli informacji na rynku finansowym – jest zaproponowana przez Eugena Famę w latach 60. i 70. XX wieku hipoteza efektywności informacyjnej rynku (ang. *efficient market hypothesis, EMH*). Warto zaznaczyć, że problematyką tą zajmowano się już wcześniej. Wystarczy wspomnieć o dysertacji Luisa Bacheliera z 1900 roku. Trudno jednak dyskutować ze stwierdzeniem, że to hipotezy postawione przez Famę stały się przedmiotem doniosłej dyskusji wśród osób zainteresowanych funkcjonowaniem rynków. W związku z powyższym, punktem wyjścia dla rozważań autorki będzie definicja rynku efektywnego informacyjnie zaproponowana przez tego właśnie badacza. W artykule z 1970 roku *Efficient Capital Markets: A Review of Theory and Empirical Work* stwierdza on, że: „rynek efektywny to taki, w którym ceny zawsze w pełni odzwierciedlają dostępne informacje”⁴⁵. Oczywiście w ten sposób zdefiniowane pojęcie może budzić wiele zastrzeżeń. Jak bowiem rozumieć stwierdzenie „zawsze w pełni”? Jak szybko ceny powinny odzwierciedlać informację oraz co uznać za

⁴² Por. Żurek, M., *Inklinacje behawioralne na rynkach kapitałowych w świetle modeli SEM*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2016, s. 16.

⁴³ Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna a Wolters Kluwer business, Warszawa 2012, s. 14.

⁴⁴ Ciołek, M., *Efektywność informacyjna...*, op.cit., s. 15–16.

⁴⁵ Fama, E. F., *Efficient Capital Markets: A Review of Theory and Empirical Work*, „The Journal of Finance”, 1970, 25/2, 28–30, s. 384.

informację, którą można określić jako „dostępną”? Sam Fama zresztą zaznacza, że tak skonstruowana definicja jest zbyt ogólna i nie pozostawia przestrzeni dla empirycznej weryfikacji⁴⁶. Dlatego też proponowana hipoteza wymaga dalszego objaśnienia. Jego podstawą będzie popularny podział na trzy formy efektywności, który zapożyczony został przez Famę od Harry’ego Robertsa⁴⁷.

Pierwszą z nich jest efektywność w formie słabej. W jej ramach informację zdefiniować można jako sekwencję historycznych cen instrumentów⁴⁸. Kluczowym problemem w tym przypadku jest odpowiedź na pytanie, czy dynamika cen danego instrumentu niesie ze sobą wartość informacyjną, umożliwiającą dokonywanie na jej podstawie predykcji cen w kolejnych okresach. W tym kontekście – w dalszej części rozdziału – przytoczona zostanie teoria błędzenia losowego, której zwolennicy opisują rynek jako taki, w którym następujące po sobie zmiany cen są od siebie niezależne⁴⁹. Na takim rynku dokonywanie przez inwestorów prognoz byłoby niemożliwe. W opozycji do tego twierdzenia stoją Joseph Grossman i Sanford Stiglitz. Ich zdaniem cenę rozumieć można jako nośnik informacji. Dzieje się tak w wyniku różnorodności wiedzy inwestorów, między którymi cena transferuje wiedzę od inwestorów poinformowanych do niepoinformowanych⁵⁰.

Druga forma efektywności, to jest efektywność półsilna, charakteryzuje rynek, w którym cena instrumentów stanowi odzwierciedlenie wszystkich informacji o charakterze publicznym⁵¹. Zdaniem autorki dysertacji, w obliczu rosnącej liczby źródeł informacji, jej ilości oraz różnorodności, istotną kwestią staje się przedyskutowanie definicji pojęcia „informacja publiczna”. Obecnie zauważyć można w tej sprawie dwa przeciwstawne podejścia. Zdaniem części badaczy informacja publiczna to ta, którą udostępnia do publicznej wiadomości emitent danego waloru⁵². Inni autorzy wskazują jednak na szeroki wachlarz informacji, do których dostęp ma inwestor, wliczając

⁴⁶ Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 384.

⁴⁷ Ciołek, M., *Efektywność informacyjna...*, op.cit., s. 35 [za: H. Robert 1967].

⁴⁸ Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 383.

⁴⁹ Fama, E. F., *Random Walks in Stock Market Prices*, „Financial Analysts Journal”, 1965, 21/5, s. 56.

⁵⁰ Zob. Stiglitz, J. E., Grossman, S. J., *On the Impossibility of Informationally Efficient Markets*, „On the Impossibility of Informationally Efficient Markets”, 70.3 (1980), s. 393–408.

⁵¹ Snarska, M., *Efektywność informacyjna...*, op.cit., s. 40.

⁵² Zob. Snarska, M., *Efektywność informacyjna...*, op.cit., s. 40.

w poczet informacji publicznych również takie, których autorami są podmioty niemające bezpośredniego związku z emitentem⁵³. Zagadnienie to rozwinięte zostanie w kolejnej części rozdziału.

Ostatnia – silna forma efektywności – związana jest z problemem dostępu do unikalnych informacji na temat aktywności podmiotów gospodarczych przez pojedynczych inwestorów lub ich wąskie grupy⁵⁴. Autorka porusza w tym kontekście dwie kwestie. Pierwsza związana jest pytaniem o to, czy informacje prywatne oraz poufne traktować możemy synonimicznie. Literatura przedmiotu wskazuje bowiem, że współczesne funkcjonowanie rynków finansowych wymaga rozróżnienia tych pojęć⁵⁵. Druga kwestia to szereg regulacji prawnych, możliwych do odnalezienia zarówno na gruncie prawa krajowego⁵⁶, jak i unijnego⁵⁷, w których znajduje się odniesienie do wskazywanego problemu.

1.3.1. Mechanizm cenowy jako źródło informacji dla inwestorów

Zanim Eugene Fama opisał hipotezę efektywności informacyjnej, w artykule z 1965 roku *The Behavior of Stock-Market Prices* podjął się wersyfikacji hipotezy błędzenia losowego. Zaznaczenie tego faktu jest o tyle istotne, że zdaniem Famy to właśnie ten model stanowił najlepsze odwzorowanie dynamiki cen na rynku akcji⁵⁸. Zgodnie z nim, przyszłe ceny akcji nie są przewidywalne w większym stopniu niż seria przypadkowo skumulowanych numerów⁵⁹. Najprostsza definicję tego zjawiska podał jeden z jego najbardziej gorliwych zwolenników, czyli Burton G. Malkiel mówiąc: „błądzenie losowe

⁵³ Zob. Mikołajek-Gocejna, M., Urbaś T., *Racjonalni inwestorzy zamiast racjonalnych oczekiwań w hipotezie rynku efektywnego*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, 2017, 481, s. 38. Zob. także Ciołek, M., *Efektywność informacyjna...*, op.cit., s. 36.

⁵⁴ Zob. Statman, M., *Behavioral Finance: Finance with Normal People*, „Borsa Istanbul Review”, 14.2 (2014), s. 72.

⁵⁵ Zob. Buzala P., *Silna a pól silna forma efektywności informacyjnej rynku kapitałowego – doświadczenia z GPW w Warszawie na podstawie zaleceń analityków*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego Finanse, Rynki Finansowe, Ubezpieczenia”, 2015, 862/75, s. 99–101.

⁵⁶ Zob. Ustawa z dnia 29 lipca 2005 r. o obrocie instrumentami finansowymi (Dz.U. 2005 nr 183 poz. 1538).

⁵⁷ Zob. Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 596/2014 z dnia 16 kwietnia 2014 r. w sprawie nadużyć na rynku (rozporządzenie w sprawie nadużyć na rynku) oraz uchylające dyrektywę 2003/6/WE Parlamentu Europejskiego i Rady i dyrektywy Komisji 2003/124/WE, 2003/125/WE i 2004/72/WE.

⁵⁸ Ciołek, M., *Efektywność informacyjna...*, op.cit., s. 30.

⁵⁹ Fama, E. F., *Random Walks in Stock Market Prices*, „Financial Analysts Journal”, 1965, 21/5 55–59, s. 34.

jest procesem, którego kolejne etapy i kierunki rozwoju pozostają bez związku z dotychczasowymi”⁶⁰. W statystycznym rozumieniu, jak twierdzi sam Fama, oznacza to, że następujące po sobie zmiany cen są niezależnymi zmiennymi losowymi o identycznym rozkładzie prawdopodobieństwa⁶¹. Tym samym, „rynek błędzenia losowego to taki, w którym kolejne zmiany, występujących na nim cen walorów, są niezależne”⁶². Owa niezależność fałszywie rozumiana bywa jako odzwierciedlenie mechanizmu, w którym ceny kształtują się w oderwaniu od rzeczywistych wydarzeń gospodarczych czy politycznych. W ramach takiego rozumowania, ceny akcji byłyby bowiem jedynie akumulacją „sygnałów” przypadkowo generowanego hałasu, tj. czynników psychologicznych lub innych – właściwych uczestnikom gry rynkowej. To jednak zdaniem Famy – nawet jak na dość bezkompromisową teorię – jest podejściem nazbyt rygorystycznym⁶³. Błędzenie losowe oznacza jedynie, że rozkład prawdopodobieństwa dla zmian cen w danym okresie „t” jest niezależny od sekwencji zmian cen w okresie wcześniejszym. Jednocześnie, chociaż Fama zdaje sobie sprawę, że znalezienie takiego szeregu zmian cen, który charakteryzowałby się idealną niezależnością, nie jest możliwe, a założenie to nie stanowi kompletnego odzwierciedlenia rzeczywistości, to twierdzi, że można uznawać je za prawdziwe tak długo, jak tylko zależność serii zmian cen nie przekroczy pewnego „akceptowalnego minimum”⁶⁴.

W praktyce powyższe założenie oznacza, że predykcja wahań cen na rynku akcji – z uwagi na to, że ich determinanty mają charakter losowy – nie jest możliwa⁶⁵. W związku z tym, wiedza o sekwencji zmian danego okresu nie stanowi żadnej wartości dodanej w szacowaniu rozkładu prawdopodobieństwa dla zmian w okresach

⁶⁰ Malkiel, B. G., *Błądząc po Wall Street. Dlaczego nie można wygrać z rynkiem*, WIG-Press, Warszawa 2003, s. 4.

⁶¹ Fama, E. F., *Random Walks...*, op.cit., s. 34.

⁶² Ibidem, s. 56.

⁶³ Ibidem, s. 36.

⁶⁴ Ibidem, s. 35.

⁶⁵ Schumaker, R., Chen, H., *Textual Analysis of Stock Market Prediction Using Financial News Articles*, „AMCIS 2006 Proceedings”, 2006, 185, s. 3.

następnych⁶⁶. Z tego względu krótkoterminowa dynamika jest nieprzewidywalna⁶⁷, a więc uzyskanie ponadprzeciętnych zysków z inwestycji również nie jest możliwe⁶⁸.

W istocie Fama formułował swoją myśl na podstawie rozprawy Luisa Bacheliera z 1900 roku, *Théorie de la Spéculation*, który zauważył, że do zmiany ceny akcji dojść może tylko w sytuacji, w której na rynek napłynie jakaś nowa informacja, a z uwagi na to, że pojawiające się informacje jako takie mają charakter losowy, dynamika cen na rynku również powinna kształtować się w taki sposób⁶⁹. Model błędzenia losowego zakłada bowiem, że wszystkie publiczne informacje są już zawarte w cenie instrumentu, a w konsekwencji aktywne handlowanie nie może przynieść ponadprzeciętnych zysków⁷⁰. Zdaniem Famy założenie o niezależności cen może być ważne tak długo, jak wiedza o poprzednich wynikach, nie może prowadzić do wzrostu oczekiwanych zysków⁷¹.

Powyższy opis nieodzownie związany jest ze słabą formą efektywności informacyjnej rynku, to jest taką, w której za zbiór komunikatów uznać można historyczne ceny instrumentów⁷². Fundamentem tej koncepcji jest założenie mówiące o tym, że aktualne ceny stanowią odzwierciedlenie informacji o historycznych notowaniach. W myśl hipotezy efektywnego rynku, nowe informacje, które mogą mieć wpływ na dynamikę cen instrumentów, pojawiają się w sposób losowy. Z uwagi na powyższe, próby dokonywania na tej podstawie prognoz, nie mogą być skuteczne⁷³. Z punktu widzenia hipotezy efektywności informacyjnej wiedza o dynamice cen instrumentów finansowych nie powinna być więc traktowana jako wartościowa informacja. Nie jest bowiem możliwe zbudowanie na jej podstawie strategii inwestycyjnej. Charakterystykę tę potwierdzają badania, które na polskim rynku dla lat 1995–2005 prowadził Jan Hájek. Przez wykorzystanie szeregu testów (błędzenia

⁶⁶ Fama, E. F., *Random Walks...*, op.cit., s. 35.

⁶⁷ Malkiel, B. G., *Błądząc po Wall Street...*, op.cit., s. 4.

⁶⁸ Schumaker, R., Chen H., *Textual Analysis...*, op.cit., s. 3.

⁶⁹ Gurgul, H., *Analiza zdarzeń...*, op.cit., s. 14.

⁷⁰ Chun, S. H., Kim, S. H., *Data Mining for Financial Prediction and Trading: Application to Single and Multiple Markets*, „Expert Systems with Applications”, 26.2 (2004), s. 131–39, s. 132.

⁷¹ Schumaker, R., Chen H., *Textual Analysis...*, op.cit., s. 56.

⁷² Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 383.

⁷³ Snarska, M., *Efektywność informacyjna giełdowych rynków akcji*, [w:] *Efektywność giełdowego rynku akcji w Polsce z perspektywy dwudziestolecia*, (red.) Czekał J., PWE, Warszawa 2014, s. 35–40.

losowego, autokorelacji, pierwiastka jednostkowego, wariancji czy też nieparametrycznych testów rang) wykazał, że Giełda Papierów Wartościowych w Warszawie jest efektywna w formie słabej⁷⁴. Wniosek ten potwierdzają badania przeprowadzone przez Dorotę Żebrowską-Suchodolską, która poddała analizie dynamikę indeksu WIG20 oraz subindeksów spółek, wchodzących w jego skład dla czterech okresów (główny okres, obejmujący lata 2000–2006, podzielony został na trzy szeregi odpowiadające okresowi spadków, trendowi bocznemu oraz okresowi wzrostów). Poza nielicznymi wyjątkami (np. okres wzrostów dla indeksu WIG20) wyniki potwierdziły hipotezę związaną ze słabą formą efektywności⁷⁵. Pozytywną weryfikację hipotezy słabej przeprowadził również Adam Waszkowski, przy czym należy zaznaczyć, że wyniki te dotyczyły jedynie sektora spożywczego⁷⁶.

Z drugiej jednak strony, cena częstokroć opisywana jest po pierwsze jako podstawowe źródło informacji, a po drugie – jako mechanizm pozwalający na jej przepływ⁷⁷. Rzecz jasna, nie chodzi tutaj o cenę jako taką, ale o jej zmienność, którą rozumieć należy jako miarę oddającą siłę wahań cen instrumentów⁷⁸. Rysują się więc w tym kontekście dwa podejścia, w ramach których cena stanowi, lub też nie stanowi, źródło użytecznej informacji dla inwestorów rynkowych. Na przykład Barry Gordon i Libby Rittenberg analizując funkcjonowanie Giełdy Papierów Wartościowych w latach 1993–1994 doszli do wniosku, że teoria rynków efektywnych w ogóle nie jest adekwatna dla opisu rynku polskiego. Zasugerowali oni, że w wielu przypadkach ważniejszą rolę odgrywają czynniki psychologiczne i społeczne (np. pewność siebie inwestorów, postawy, a nawet moda)⁷⁹.

Zauważyć można jednak, że wiele dotychczas przeprowadzonych badań nie pozwala na jednoznaczne opisanie polskiej giełdy jako rynku w pełni efektywnego lub

⁷⁴ Zob. Hajek, J., *Weak-form Efficiency Test in the Central European Capital Markets*, „Politica Ekonomiczna”, 6, s. 773–791.

⁷⁵ Zob. Żebrowska-Suchodolska, D., *Weryfikacja słabej formy efektywności za pomocą testów statystycznych na GPW w Warszawie*, „Optimum Studia Ekonomiczne”, 2009, 1(41).

⁷⁶ Zob. Waszkowski, A., *Hipoteza efektywności rynku-weryfikacja dla indeksu WIG-Spożywczy*, „Problemy Rolnictwa Światowego”, 2011, 11(26)/4.

⁷⁷ Por. Będowska-Sójka, B., *Wpływ informacji...*, op.cit., s. 25.

⁷⁸ Doman, M., *Mikrostruktura Giełd Papierów Wartościowych*, Wydawnictwo Uniwersytetu w Poznaniu, Poznań 2011, s. 35.

⁷⁹ Zob. Gordon, B., Rittenberg, L., *The Warsaw Stock Exchange: A Test of Market Efficiency*, „Comparative Economic Studies”, 37/2.

nieefektywnego w sensie słabym. Przykładowo Krzysztof Kompe oraz Aleksandra Matuszewska-Janicka wykazali, że słaba forma efektywności dotyczy indeksów WIG oraz WIG20, a także niektórych spółek (Orlen, TPSA, KGHM). Natomiast indeks średnich spółek (wówczas MIDWIG) okazał się nieefektywny. Na tej podstawie sądzić można, że uzyskanie ponadprzeciętnych stóp zwrotu byłoby więc możliwe, ale w ograniczonym do konkretnego indeksu stopniu⁸⁰. Związek między wielkością spółek a efektywnością zauważyły również Dorota Witkowska oraz Dorota Żebrowska-Suchodolska. W przeprowadzonych przez nie badaniach najwyższą efektywnością wykazywały się spółki indeksowane w WIG20. Ponadto, autorki, dzieląc okres badawczy (1994–2006) na dziewiętnaście podokresów, zwróciły uwagę, że o ile – wraz z rozwojem polskiej giełdy – odnotować można zwiększającą się efektywność rynku, to w przypadku mniejszych spółek brak efektywności zdarza się relatywnie często⁸¹. Być może obserwacja związana z diachronicznym ujęciem efektywności uzasadnić może wyniki uzyskane przez Gordona i Rottenberga, z uwagi na to, że ich badanie obejmowało bardzo wczesny okres funkcjonowania Giełdy Papierów Wartościowych w Warszawie.

Ta dyskusja wydaje się jednak jedynie pochodną dwóch innych problemów. Pierwszy z nich dotyczy klasyfikacji inwestorów działających na rynku. Jeżeli bowiem założymy, że inwestorzy – zarówno pod względem posiadania, jak i reakcji na informację – stanowią grupę homogeniczną, to stanowisko Famy wydaje się zasadne. Jeżeli jednak uznać, że inwestorzy nie mają jednolitego zestawu informacji, a ich zdolności interpretacyjne i reakcje mogą się od siebie różnić, wówczas jego koncepcja narażona jest na krytykę. Badacze Joseph Grossman i Sanford Stiglitz, dyskutując z założeniami Famy, wykazali, że podstawą dla utrzymania koncepcji konkurencyjnego rynku kapitałowego jest między innymi różnorodność w wiedzy inwestorów, tj. różne możliwości absorbowania przez nich informacji⁸².

⁸⁰ Zob. Kompa, K., Matuszewska-Janicka, A., *The Warsaw Stock Exchange: A Test of Market Efficiency*, „International Advances in Economic Research”, 2009, 15, s. 59–70.

⁸¹ Zob. Witkowska, D., Żebrowska-Suchodolska, D., *Badanie słabej formy efektywności informacyjnej GPW*, „Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania”, 2008, 9, s. 155–165.

⁸² Zob. Stiglitz, J. E., Grossman, S. J., *On the Impossibility of Informationally Efficient Markets*, „On the Impossibility of Informationally Efficient Markets”, 70.3 (1980), s. 393–408.

W ich propozycji, ceny odgrywają rolę przekąznika wiedzy od inwestorów poinformowanych do niepoinformowanych, a poziom informatywności systemu cen zależny jest od liczebności tych pierwszych⁸³. Drugi problem związany jest z odpowiedzią na pytanie o determinanty zmian cen. W modelach przypisujących cenie funkcję nośnika informacji najczęściej mowa jest bowiem o takich informacjach, które wcześniej musiały zostać pozyskane przez część inwestorów przy odpowiednim nakładzie pracy lub kapitału. Zakłada się więc, że napływ informacji publicznej i prywatnej determinuje zachowanie cen i wolumenów. Inwestorzy niepoinformowani, odkrywają informację na rynku kapitałowym obserwując procesy dynamiki cen, wolumenów i innych wielkości związanych z analizowanymi instrumentami⁸⁴. W tym rozumieniu cenę można więc traktować jako informację, ale informację o charakterze wtórnym wobec trudniejszych do zaobserwowania informacji fundamentalnych. W tej sytuacji konieczne jest omówienie dwóch kolejnych typów informacji, na których bazuje hipoteza efektywności informacyjnej: publicznej i prywatnej.

1.3.2. Informacja publiczna

O rynku efektywnym informacyjnie w formie półsilnej możemy mówić wówczas, gdy cena instrumentów stanowi odzwierciedlenie wszystkich informacji o charakterze publicznym⁸⁵. Innymi słowy, istotą zagadnienia jest odpowiedź na pytanie czy ceny aktywów dostosowują się do tego typu informacji⁸⁶. O ile w przypadku hipotezy słabej typ informacji, o którym mowa, wydaje się jasny, o tyle w przypadku formy półsilnej sytuacja staje się mniej oczywista.

Po pierwsze, należy zaznaczyć, że nie zakłada się tutaj istnienia odrębnego typu informacji niż ten, o którym mowa we wcześniejszej formie efektywności. Należy raczej stwierdzić, że w ramach półsilnej efektywności, zbiór branych pod uwagę informacji został poszerzony o informacje niecenowe⁸⁷. Ta forma efektywności idzie więc o krok dalej poprzez założenie o dyskontowaniu w cenach, zarówno informacji historycznych,

⁸³ Stiglitz Joseph E., Grossman Sanford J., 'On the Impossibility of Informationally Efficient Markets', *On the Impossibility of Informationally Efficient Markets*, 70.3 (1980), s. 393–408, s. 394.

⁸⁴ Będowska-Sójka, B., *Wpływ informacji...*, op.cit., s. 24.

⁸⁵ Snarska, M., *Efektywność informacyjna...*, op.cit., s. 40.

⁸⁶ Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 383.

⁸⁷ Czekaj, J., Woś, M., Żarnowski, J., *Efektywność giełdowego rynku akcji w Polsce. Z perspektywy dziesięciolecia*, Wydawnictwo Naukowe PWN, Warszawa 2001, s. 40–41.

jak i bieżących⁸⁸. Co do poszerzenia tego katalogu, Fama w artykule z 1970 roku wypowiedział się jednak dość lakonicznie, podając jedynie dwa przykłady: ogłoszeń o rocznych zarobkach spółek i podziale akcji⁸⁹. Mimo to, przytaczane w literaturze przedmiotu egzemplifikacje sugerują, że chodzi tu o informacje przekazywane inwestorom przez emitentów w formie raportów, strategii, wyników lub planów⁹⁰. Jednocześnie spotkać można się ze stwierdzeniem, że w skład takiej grupy wchodzi również informacje „pochodzące z innych źródeł, które mogą mieć wpływ na cenę akcji”⁹¹. Przykładowo, w zestawieniu informacji, które określić można jako ogólnie dostępne, Maciej Ciołek wymienia: notowania historycznych cen oraz wolumenów, komunikaty prasowe, raporty bieżące, dane makroekonomiczne, jak również wiadomości, które są powszechnie publikowane⁹².

Z metodologicznego punktu widzenia, badania prowadzone nad pólśilną efektywnością informacyjną sklasyfikować można na te wykorzystujące analizę portfelową oraz analizę zdarzeń. W pierwszym przypadku autorzy konstruuja portfele, zawierające akcje spółek, na podstawie jednego lub kilku wskaźników fundamentalnych. Celem badań jest weryfikacja możliwości osiągnięcia statystycznie istotnych, ponadnormalnych stóp zwrotu, dla których punktem odniesienia są stopy zwrotu z indeksów szerokiego rynku lub innych, np. spółek najwyższej kapitalizacji lub sektorowych.

Tego typu badanie na polskim rynku przeprowadzili między innymi Marcin Czupryna oraz Małgorzata Snarska. Ich próba badawcza obejmowała ponad szesnaście tysięcy obserwacji. Analizie poddane zostały wskaźniki: kapitalizacji, ceny do zysku (P/E) oraz ceny do wartości księgowej (P/BV). Przeprowadzona estymacja, wykorzystująca kapitalizację spółek jako zmienną objaśniającą ponadnormalną stopę zwrotu, pozwoliła stwierdzić, że czynnik ten jest istotną statystycznie determinantą, kształtującą stopy zwrotu portfela referencyjnego. Dla całego okresu wpływ ten został

⁸⁸ Schumaker, R. P., Chen, H., *A Quantitative Stock Prediction System Based on Financial News*, „Information Processing and Management”, 45.5 (2009), s. 571–583 (s. 572).

⁸⁹ Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 383.

⁹⁰ Snarska, M., *Efektywność informacyjna...*, op.cit., s. 40.

⁹¹ Mikołajek-Gocejna, M., Urbaś T., *Racjonalni inwestorzy zamiast racjonalnych oczekiwań w hipotezie rynku efektywnego*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, 2017, 481, s. 38.

⁹² Ciołek, M., *Efektywność informacyjna...*, op.cit., s. 36.

oszacowany jako nieprzekraczający 1%⁹³. Wpływ kapitalizacji analizowany był również przez Damiana Bulskiego i Mariana Górskiego. Autorzy – na podstawie badań własnych – doszli jednak do wniosku, że w długim i średnim okresie nie jest możliwe osiągnięcie stóp zwrotu wyższych niż rynkowe⁹⁴.

Marcin Czupryna i Małgorzata Snarska wykazali również, że na kształtowanie się stóp zwrotu z portfela referencyjnego wpływ ma wskaźnik P/E, ale w tym przypadku czynnik determinacji dla całego badanego okresu nie przekracza 0,1%⁹⁵. W swoich badaniach ten wskaźnik wykorzystał również Paweł Sekuła. Otrzymane przez niego wyniki pozwoliły wysnuć wniosek o możliwości osiągnięcia stóp zwrotu wyższych niż średni zwrot z indeksu WIG. Paweł Sekuła podkreślił jednocześnie, że nadwyżkowe stopy zwrotu charakteryzowały się niestabilnością (w niektórych przypadkach portfele osiągały wyniki wręcz niższe od indeksu) i w związku z tym nie znalazł podstaw do odrzucenia hipotezy pól silnej efektywności⁹⁶. Do podobnego wniosku doszli Damian Bulski i Marian Górski, u których wyniki osiągane przez portfele skonstruowane na podstawie wskaźnika P/E różniły się w zależności od okresu (od -51,4% do 48,8%). Autorzy postulowali więc aby dalszemu testowaniu podlegały portfele budowane w oparciu o różne wartości wskaźnika ceny do zysku⁹⁷.

Damian Bulski i Marian Górski doszli do analogicznej konkluzji, wykorzystując do budowy portfeli wskaźnik P/BV. Ich zdaniem wyniki, sugerujące nieefektywność rynku w sensie pól silnym, widoczne są zwłaszcza w drugiej połowie pierwszej dekady wieku (okres badania obejmował lata 2005–2011)⁹⁸. Marcin Czupryna i Małgorzata Snarska natomiast opisują ten wskaźnik, wprowadzając jako mający istotny wpływ na kształtowanie się stopy zwrotu z portfela, ale osiągający najniższy współczynnik determinacji spośród analizowanych czynników fundamentalnych (0,05%)⁹⁹.

⁹³ Czupryna, M., Snarska, M., *Efektywność pól silna*, [w:] *Efektywność giełdowego rynku akcji w Polsce z perspektywy dwudziestolecia*, (red.) Czekaj J., PWE, Warszawa 2014, s. 107–117.

⁹⁴ Bulski D., Górski M., *Wyniki testów hipotezy pól silnej efektywności informacyjnej Giełdy Papierów Wartościowych w Warszawie w latach 2005–2011*, „Problemy Zarządzania”, 2012, 10/4, s. 141–168.

⁹⁵ Czupryna, M., Snarska, M., *Efektywność pól silna...*, op.cit., s. 107–117.

⁹⁶ Sekuła, P., *Empiryczny test strategii fundamentalnej*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, 2012, 271, s. 280–289.

⁹⁷ Bulski, D., Górski, M., *Wyniki testów...*, op.cit., s. 141–168.

⁹⁸ Ibidem, s. 141–168.

⁹⁹ Czupryna, M., Snarska, M., *Efektywność pól silna...*, op.cit., s. 107–117.

W badaniach prowadzonych przez Sekułę proponowane są dodatkowo trzy wskaźniki oparte na zyskach spółek, tj. dynamika zysku netto, stopa wzrostu EBIT oraz wskaźnik rentowności kapitału własnego (ROE). W pierwszym przypadku autor zaobserwował zależność wysokości ponadnormatywnych stóp zwrotu od stopnia koncentracji portfela (portfele rozbudowywane były o każde pięć kolejnych spółek w zależności od poziomu ROE; najmniejszy portfel składał się z pięciu spółek, największy z dwudziestu). Jednocześnie stopy zwrotu trzech z czterech portfeli wykazały istotność statystyczną. Wyjątek stanowił portfel piętnastoelementowy. Podobny wynik dotyczył dynamiki zysku netto, tzn. istotność statystyczną wykazały stopy zwrotów z portfeli o tej samej liczebności, chociaż każdy pozwolił osiągnąć wyniki wyższe od indeksu WIG. Mimo że w przypadku portfeli konstruowanych na podstawie zysku operacyjnego również osiągnięto stopy zwrotu przewyższające średnie roczne zwroty z indeksu, to wyniki nie okazały się statystycznie istotne, a co za tym idzie, nie można było na tej podstawie odrzucić hipotezy zgodnej z teorią efektywności informacyjnej¹⁰⁰.

Trzeba jednak zauważyć, że rozbieżne wnioski z przytaczanych badań mogą mieć swoje uzasadnienie w czynnikach niezwiązanych *sensu stricto* z teorią i praktyką rynków finansowych. Przykładowo dla Marcin Czapryna i Małgorzata Snarska oraz Pawła Sekuły benchmarkiem był indeks WIG, podczas gdy Damian Bulski i Marian Górski jako taki potraktowali średnią nieważoną stopę zwrotu z akcji badanych spółek. Sposób budowania portfeli również nie był jednorodny. Marcin Czapryna i Małgorzata Snarska konstruowali portfele kwintylowe, a Damian Bulski i Marian Górski decylowe. Różne były również okresy badawcze, a co za tym idzie liczba obserwacji. Niemniej wydaje się, że na podstawie dotychczasowych analiz portfelowych nie jest możliwe jednoznaczne stwierdzenie czy pólśilna hipoteza efektywności informacyjnej opisuje rzeczywistość polskiego rynku finansowego, czy też nie.

Trudności w zdefiniowaniu informacji publicznej – choć w pełni uzasadnione – prowadzą przynajmniej do dwóch konkluzji. Po pierwsze, należy zaznaczyć, że mowa tu nie tylko o informacjach przekazywanych przez emitenta w sposób bezpośredni, ale również o takich, które pozyskać może inwestor, a na które emitent nie ma

¹⁰⁰ Sekuła, P., *Empiryczny test...*, op.cit., s. 280–289.

bezpośredniego wpływu. Z uwagi na niebywały przyrost zarówno źródeł informacji, jak i jej ilości, lista oraz dostępność do tego typu danych wydaje się wciąż poszerzać. Po drugie, różnorodność informacji publicznej prowokuje do zadania pytania o sposób przyjęcia lub odrzucenia hipotezy o pólslnej efektywności. Wynika to chociażby z możliwości gradacji jej dostępności, a tym samym z różnicowania nakładu pracy, jaki jest niezbędny do jej zdobycia. Problem ten obrazują badania nad reakcją inwestorów na rekomendacje biur maklerskich. Wykorzystując analizę zdarzeń, autorzy często poddają obserwacji dwie daty zdarzenia. Pierwszą jest udostępnienie informacji komercyjnym klientom danego biura, czyli moment wydania rekomendacji. Drugą datą natomiast jest jej upublicznienie, czyli udzielenie dostępu szerokiemu gronu odbiorców. Przegląd literatury pozwala stwierdzić, że dotychczas polski rynek przebadany został pod względem wszystkich typów rekomendacji w skali trzystopniowej: kupuj, trzymaj, sprzedaj, oraz pięciostopniowej: kupuj, akumuluj, trzymaj, redukuje, sprzedaj.

Badania przeprowadzone przez Pawła Mielcarza i innych wskazują, że rekomendacja „kupuj” prowadzi do statystycznie istotnych ponadnormatywnych stóp zwrotu z akcji, a zjawisko to widoczne jest w dniu przekazania informacji klientom biur maklerskich, kiedy to średnia nadwyżkowa stopa zwrotu osiąga 0,79%. Ponadto autorzy zauważyli, że jeśli za dzień zdarzenia przyjęta zostanie data upublicznienia rekomendacji, to hipotezę o braku nadwyżkowych stóp zwrotu można odrzucić na podstawie wyników z kilku dni poprzedzających zdarzenie. Dodatkowo ujemne nadwyżkowe stopy zwrotu odnotowane zostały w drugim i trzecim dniu po udostępnieniu rekomendacji opinii publicznej. Może to sugerować, że niektórzy komercyjni klienci instytucji rekomendujących, stosują strategię inwestycyjną polegającą na kupowaniu akcji z rekomendacją „kupuj” tuż przed podaniem jej do publicznej wiadomości i sprzedawaniu ich natychmiast po podwyżce wywołanej przez inwestorów, którzy nie mają dostępu do informacji dostarczanych klientom komercyjnym¹⁰¹. Wnioski te wydają się potwierdzone przez rezultaty osiągnięte przez Pawła Buzalę. Autor w badaniach przeprowadzonych w 2012 oraz 2015 roku wykazał, że konsekwencją zarówno wydania,

¹⁰¹ Zob. Mielcarz, P., Podgórski, B., Wermczuk, P., *Positive recommendations and abnormal returns on the Warsaw Stock Exchange in 2005–2006*, [w:] *The Problems of Company Value Management*, (red.) Urbańczyk, E., Publishing House Print Group Daniel Krzanowski, Lewiston 2007, afiliacja: WSPiZ.

jak i upublicznienia rekomendacji „kupuj” są ponadnormatywne stopy zwrotu (przy czym nie w każdej próbie charakteryzowały się one statystyczną istotnością)¹⁰².

Przytoczone badania stoją jednak w sprzeczności z wnioskami prezentowanymi na przykład przez Henryka Gurgula, który w ramach swoich badań wprowadził hipotezę, że uczestnicy rynku postrzegają rekomendację „kupuj” jako pozytywny sygnał, a tym samym po jej ogłoszeniu cena akcji powinna rosnąć, ale ostatecznie wykazał, że rekomendacja ta nie ma statystycznie istotnego wpływu na cenę akcji¹⁰³. Do tej samej konkluzji w 2018 roku doszedł Kamil Polak, który – w wynikach swoich badań – nie znalazł podstaw do przyjęcia hipotezy o statystycznie istotnym wpływie tej rekomendacji na ceny walorów¹⁰⁴.

W przypadku rekomendacji „sprzedaj” wyniki osiągnięte przez Henryka Gurgula i Kamila Polaka, również pozwoliły na wyciągnięcie tożsamyh wniosków. W przypadku badań Henryka Gurgula ta rekomendacja była właściwie jedyną, dla której można było zaobserwować ponadnormatywne stopy zwrotu, co widoczne było na dzień przed jej upublicznieniem¹⁰⁵. Podobnie w badaniach przeprowadzonych przez Polaka wynik okazał się ujemny i istotnie statystyczny na dzień przed przekazaniem informacji do ogólnego dostępu. Logikę tę potwierdzają również wyniki analiz Pawła Mielcarza i Błażej Podgórskiego. Nie dostrzegli oni statystycznie istotnej przeceny walorów w dniu upublicznienia rekomendacji, ale zauważyli ją na 5 oraz 2 dni przed zdarzeniem. Dodatkowo analogiczny rezultat widoczny był w dniu wydania rekomendacji klientom biur, a nawet do siedmiu dni wcześniej. Autorzy przypuszczają, że z jednej strony może być to oczywiście wynik wycieku wcześniej nieudostępnianych informacji, ale równie

¹⁰² Buzala, P., *Reakcja inwestorów na rekomendacje giełdowe. Implikacje dla efektywności rynku akcji*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego Finanse, Rynki Finansowe, Ubezpieczenia”, 2012, 689/50, s. 11–20; oraz Buzala, P., *Silna a pól silna forma efektywności informacyjnej rynku kapitałowego – doświadczenia z GPW w Warszawie na podstawie zaleceń analityków*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego Finanse, Rynki Finansowe, Ubezpieczenia”, 2015, 862/75, s. 97–109.

¹⁰³ Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna a Wolters Kluwer business, Warszawa 2012.

¹⁰⁴ Zob. Polak, K., *Wpływ opublikowanej rekomendacji na reakcje inwestorów*, „Studia i Materiały Wydziału Zarządzania UW”, 2018, 27/2, s. 127–135.

¹⁰⁵ Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna a Wolters Kluwer business, Warszawa 2012.

dobrze może być związane z wcześniejszymi spadkami cen na rynku¹⁰⁶. Na podstawie badań przeprowadzonych przez Pawła Buzalę zauważyć można podobne prawidłowości. Wprawdzie w tym przypadku w momencie wydania rekomendacji możliwe było osiągnięcie ponadnormatywnej i statystycznie istotnej stopy zwrotu, ale była ona relatywnie niższa od wyników osiąganych między pierwszym a piątym dniem przed zdarzeniem. Analogiczna sytuacja miała miejsce w kontekście upublicznienia rekomendacji, gdzie w dniu zdarzenia osiągnięto ujemną, ponadnormatywną i statystycznie istotną stopę zwrotu na poziomie 0,49%, ale w relacji do dni wcześniejszych była ona niższa, a dziesięć dni przed zdarzeniem osiągnęła poziom - 2,62%¹⁰⁷.

Warto podkreślić, że część z przytoczonych badań obrazuje zmiany w poziomie ponadnormatywnych stóp zwrotu po momencie definiowanym jako zdarzenie. Przykładowo w badaniach Pawła Buzalę od dnia upublicznienia rekomendacji „sprzedaj” do piątego dnia po zdarzeniu stopy zwrotu rosły zgodnie z kierunkiem rekomendacji (od -0,71% do -1,25%), wykazując się statystyczną istotnością. Podobny mechanizm widoczny był przy rekomendacji „kupuj”, przy czym trwał on dłużej, bo aż do piętnastego dnia po zdarzeniu¹⁰⁸. Obserwację tę można umieścić w kontekście zjawiska dryfu poinformacyjnego i asymetrii informacji. Zagadnienie to zostanie szerzej omówione w dalszej części rozdziału.

1.3.3. Informacja prywatna

Meir Statman sugeruje, że kluczem do rozróżnienia informacji publicznej od prywatnej jest jej unikalność oraz możliwość wglądu w wewnętrzną sytuację spółki, a inwestorzy dysponujący informacjami, które charakteryzują się wysokim poziomem obu tych czynników, mają szansę na uzyskanie lepszych wyników na rynku, od inwestorów dysponujących jedynie powszechnie dostępną, tj. publiczną informacją¹⁰⁹. Wydaje się, że ten tok myślenia bliski jest logice proponowanej przez Eugena Fama. Opisując

¹⁰⁶ Zob. Mielcarz, P., Podgórski, B., *Wpływ negatywnych i neutralnych na osiągnięcie ponadprzeciętnych stóp zwrotu na GWP w latach 2005–2006*, „Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania”, 2008, 7, s. 533–543.

¹⁰⁷ Buzala, P., *Reakcja inwestorów...*, op.it., s. 11–20.

¹⁰⁸ Ibidem, s. 11–20.

¹⁰⁹ Statman, M., *Behavioral Finance: Finance with Normal People*, „Borsa Istanbul Review”, 14.2 (2014), s. 65–73 (s. 72).

bowiem ostatnią – silną formę efektywności – definiuje ją poprzez odpowiedź na pytanie o to, czy jakaś grupa inwestorów ma monopolistyczny (a więc unikalny) dostęp do informacji, które związane są z procesem formowania się cen¹¹⁰.

Do tej pory zwroty: informacja prywatna i poufna stosowane były wymiennie. W ramach literatury przedmiotu wydaje się być to częstą praktyką, z uwagi na to, że granica między oboma pojęciami nie jest wyraźna. Zgodnie z Ustawą o obrocie instrumentami finansowymi informacja poufna to:

[...] określona w sposób precyzyjny – informacja dotycząca, bezpośrednio lub pośrednio, jednego lub kilku emitentów instrumentów finansowych, jednego lub kilku instrumentów finansowych albo nabywania lub zbywania takich instrumentów, która nie została przekazana do publicznej wiadomości, a która po takim przekazaniu mogłaby w istotny sposób wpłynąć na cenę tych instrumentów finansowych lub na cenę powiązanych z nimi pochodnych instrumentów finansowych¹¹¹.

Jako potencjalnych posiadaczy informacji poufnej wymienia się tutaj trzy grupy podmiotów. Pierwszą z nich są osoby, które pełnią w spółce emitenta funkcję, są w niej zatrudnione lub połączone ze spółką stosunkiem prawnym (członkowie zarządu, rady nadzorczej, prokurenci, pełnomocnicy, pracownicy, biegli rewidenci itp.). Druga grupa to osoby, które weszły w posiadanie informacji poufnej w wyniku popełnienia przestępstwa. Trzecia natomiast to osoby, które przy dołożeniu należytej staranności mogły dowiedzieć się, że informacja, którą posiadają, ma charakter poufny. Zgodnie z artykułem 156 Ustawy wszystkie te osoby objęte są zakazem wykorzystania takiej informacji, przez co rozumie się

nabywanie lub zbywanie, na rachunek własny lub osoby trzeciej, instrumentów finansowych, w oparciu o informację poufną będącą w posiadaniu tej osoby, albo dokonywanie, na rachunek własny lub osoby trzeciej, innej czynności prawnej powodującej lub mogącej powodować rozporządzenie takimi instrumentami finansowymi¹¹².

¹¹⁰ Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 383.

¹¹¹ Art. 154 ustawy z dnia 29 lipca 2005 r. o obrocie instrumentami finansowymi (Dz.U. 2005 nr 183 poz. 1538).

¹¹² Art. 156 ustawy z dnia 29 lipca 2005 r. o obrocie instrumentami finansowymi (Dz.U. 2005 nr 183 poz. 1538).

Informacja poufna rozpatrywana jest więc przez pryzmat wiążących się z nią nadużyć (ang. *Market abuse*). Zgodnie z Rozporządzeniem Parlamentu Europejskiego i Rady (UE) nr 596/2014:

nadużycie na rynku to pojęcie obejmujące bezprawne zachowanie na rynkach finansowych, a [...] należy je rozumieć jako wykorzystywanie informacji poufnych, bezprawne ujawnianie informacji poufnych oraz manipulacje na rynku¹¹³.

Taka forma działalności określana jest mianem *insider tradingu* i obarczona jest zasadniczymi, negatywnymi konsekwencjami, zarówno z punktu widzenia pojedynczych inwestorów, jak i całego rynku. Po pierwsze, możliwość uprawiania tego typu handlu wynika z faktu, że osoby wtajemniczone w działanie spółki (ang. *company insiders*) posiadają tylko im dostępną wiedzę i potencjalnie mogą wykorzystać ją do osiągnięcia na rynku lepszych wyników, niż pozostali inwestorzy¹¹⁴. Mowa jest więc o „nieuczciwym uzyskaniu korzyści [...] ze szkodą dla osób trzecich niemających dostępu do takich informacji”¹¹⁵. Po drugie, Michael Fishman i Kathleen Hagerty, którzy na początku lat 90. badali zjawisko *insider tradingu*, wskazują, że następstwem takich sytuacji może być dostrzeganie przez inwestorów nieuczciwości i strata przez nich zaufania do rynku kapitałowego¹¹⁶. Również w ramach Rozporządzeniem Parlamentu Europejskiego i Rady (UE) potencjał utraty zaufania przez inwestorów rynkowych, opisywany jest jako jedno z zagrożeń omawianego handlu¹¹⁷. Po trzecie, należy zauważyć, że oba zjawiska będą miały przełożenie na funkcjonowanie rynku jako takiego. Fishman i Hagerty w tym kontekście wspominają o możliwości utraty płynności rynku¹¹⁸. Natomiast w przytaczanym rozporządzeniu zwraca się uwagę, że taka aktywność ogranicza

¹¹³ Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 596/2014 z dnia 16 kwietnia 2014 r. w sprawie nadużyć na rynku (rozporządzenie w sprawie nadużyć na rynku) oraz uchylające dyrektywę 2003/6/WE Parlamentu Europejskiego i Rady i dyrektywy Komisji 2003/124/WE, 2003/125/WE i 2004/72/WE.

¹¹⁴ McKinley, K., *Stock Market Efficiency and Insider Trading*, „Issues in Political Economy”, 1999, s. 1.

¹¹⁵ Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 596/2014 z dnia 16 kwietnia 2014 r. w sprawie nadużyć na rynku (rozporządzenie w sprawie nadużyć na rynku) oraz uchylające dyrektywę 2003/6/WE Parlamentu Europejskiego i Rady i dyrektywy Komisji 2003/124/WE, 2003/125/WE i 2004/72/WE, pkt 23.

¹¹⁶ Fishman, M. J., Hagerty, K. M., *Insider Trading and the Efficiency of Stock Prices*, „The RAND Journal of Economics Journal of Economics”, 23.1 (1992), s. 106–122 (s. 106).

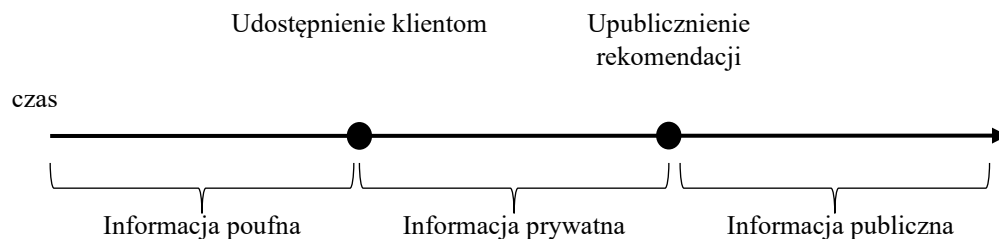
¹¹⁷ Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 596/2014 z dnia 16 kwietnia 2014 r. w sprawie nadużyć na rynku (rozporządzenie w sprawie nadużyć na rynku) oraz uchylające dyrektywę 2003/6/WE Parlamentu Europejskiego i Rady i dyrektywy Komisji 2003/124/WE, 2003/125/WE i 2004/72/WE, pkt 7 oraz 23.

¹¹⁸ Fishman, M. J., Hagerty, K. M., *Insider Trading...*, op.cit., s. 106.

przejrzystość rynku, która jest warunkiem koniecznym dla aktywności podmiotów gospodarczych lub też, że stanowi ona wręcz zagrożenie dla integralności rynków finansowych¹¹⁹.

Na tym etapie wydawałoby się więc jasne, że trzecia forma efektywności rynkowej związana jest z informacją poufną. Fama w artykule z 1991 roku *Efficient Capital Markets: II* pisze jednak: „zamiast testów silnej formy, sprawdzających czy poszczególni inwestorzy posiadają informację, która nie jest zawarta w cenach rynkowych, sugeruję używać bardziej opisowego tytułu, jakim są testy dla prywatnych informacji”¹²⁰. Można więc przypuszczać, że autor analizuje raczej problem informacji prywatnej niż poufnej. Oczywiście przyjąć można, że taka nieścisłość wynika z tłumaczenia. Mimo to, interesujące objaśnienie tego problemu zaproponował Piotr Buzala, który wskazuje, że silna forma efektywności rynku dotyczy tych wszystkich informacji, które nie są znane publicznie. Wskazuje on więc, że istnieje grupa informacji niepublicznych, przy czym grupa ta nie musi być jednolita. Za Buzalą opisać można to na przykładzie – wspomnianych wcześniej – rekomendacji biur maklerskich, co zaprezentowane jest na rysunku 1¹²¹.

Rysunek 1. Typy informacji na rynku finansowym



Źródło: Buzala, P., *Silna a pól silna forma efektywności informacyjnej rynku kapitałowego – doświadczenia z GPW w Warszawie na podstawie zaleceń analityków*.

¹¹⁹ Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 596/2014 z dnia 16 kwietnia 2014 r. w sprawie nadużyć na rynku (rozporządzenie w sprawie nadużyć na rynku) oraz uchylające dyrektywę 2003/6/WE Parlamentu Europejskiego i Rady i dyrektywy Komisji 2003/124/WE, 2003/125/WE i 2004/72/WE, pkt 23.

¹²⁰ Fama, E. F., *Efficient Capital Markets: II*, „The Journal of Finance”, 46.5 (1991), s. 1575–1617 (s. 1577).

¹²¹ Buzala P., *Silna a pól silna forma efektywności informacyjnej rynku kapitałowego – doświadczenia z GPW w Warszawie na podstawie zaleceń analityków*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego Finanse, Rynki Finansowe, Ubezpieczenia”, 2015, 862/75, s. 99–101.

W ten sposób przyjąć można, że w grupie informacji niepublicznych znajdują się zarówno informacje prywatne, jak i poufne, a kryterium definicyjnym jest nie tyle sam fakt dostępu do informacji, ile czas jej udostępnienia poszczególnym typom odbiorców. Można w tym miejscu zastanowić się, czy konieczność poniesienia kosztów związanych z dostępem do informacji powinna stanowić fundament rozróżniający ją na publiczną i prywatną. Koszty mogą bowiem wynikać nie z samego faktu otrzymywania informacji, ale na przykład skrócenia czasu wymaganego do jej pozyskania. Dzieje się tak zarówno w przypadku rekomendacji przekazywanych przez biura maklerskie, jak i w przypadku wielu płatnych baz danych, które agregują komunikaty ze źródeł pierwotnych.

Z powyższych rozważań rysuje się obraz informacji, związanej z rynkiem finansowym jako pojęcia dalekiego od koherencji. Składają się na to przynajmniej dwa czynniki. Z jednej strony, mamy do czynienia z różnymi typami informacji, spośród których wymienić można informację: cenową, publiczną, poufną i prywatną. Z drugiej natomiast, taki stan rzeczy podaje w wątpliwość homogeniczność inwestorów jako odbiorców informacji. Różnorodność w podejmowaniu przez nich decyzji inwestycyjnych może być bowiem wynikiem asymetrii związanej z jej dostępnością.

1.4. Asymetria informacji

Jeden z czołowych przedstawicieli szkoły austriackiej, Friedrich Hayek, w tekście *The Use of Knowledge in Society* z 1945 roku stwierdził, że szczególny charakter problemu racjonalnego porządku gospodarczego zdeterminowany jest przez fakt, że wiedza o okolicznościach, z których człowiek musi zrobić użytek, nigdy nie istnieje w skoncentrowanej lub zintegrowanej formie, ale występuje wyłącznie pod postacią rozproszonych, niekompletnych i często sprzecznych elementów. Przez to też, społecznym problemem ekonomii nie jest kwestia alokacji danych zasobów. Jest to raczej problem wykorzystania wiedzy, której nikt nie otrzymuje w całości¹²². Mimo tego spostrzeżenia, w czasach, kiedy Fama budował hipotezę efektywnego rynku, oczywistym wydawało się założenie o istnieniu informacji doskonałej. Jak twierdzi Joseph Stiglitz, przekonanie to było na tyle powszechne, że w ramach dowodów na podstawowe twierdzenia ekonomii badacze –w swoich zestawach założeń – nie artykułowali go

¹²² Zob. Hayek, F. A., *The Use of Knowledge in Society*, 1945, s. 519–520

*explicite*¹²³. Fama w artykule z 1970 postąpił jednak inaczej i wskazał trzy założenia rynku efektywnego, spośród których jedno określało dostęp do informacji jako powszechny i darmowy¹²⁴. Obie te cechy do dziś stanowią przedmiot dyskusji. Choć obecnie pojęcie asymetrii informacyjnej jest powszechnie znane, to jednak nie przestają zastanawiać konsekwencje, jakie niesie ono za sobą dla funkcjonowania rynku finansowego. W związku z tym, problem ten wymaga przynajmniej esencjonalnego omówienia.

Sytuacja związana z założeniem o istnieniu informacji doskonałej, jak może się zdawać, zmieniła się w roku 1971 roku, kiedy Walter Bagehot opublikował artykuł *The only game in town*¹²⁵. Autor wskazywał w nim, że problem *spreadu* mógłby istnieć, wyłącznie w oparciu o fakt, że niektórzy uczestnicy rynku są w posiadaniu prywatnych informacji oraz że rozmiar jego rozpiętości zależeć może od asymetrii informacyjnej¹²⁶. Chociaż Bagehot nie zdefiniował tego pojęcia, to w ramach niniejszej dysertacji wykorzystana zostanie prosta definicja Stiglitz, według którego asymetria informacji to sytuacja, w której „różne osoby dysponują różnymi wiadomościami”¹²⁷. Kluczowym czynnikiem, który wpływa na ten fakt, są koszty związane z pozyskaniem informacji, przy czym przez taki koszt należy rozumieć nie tylko aspekt finansowy, ale również czas, który inwestor poświęcić musi na dotarcie do danej informacji¹²⁸.

Zdaniem Sanforda Grossmana i Josepha Stiglitz koszty pozyskania informacji są punktem, w którym teoria Famy może się załamywać. Wykazali oni, że jeżeli hipoteza efektywności informacyjnej rynku byłaby prawdziwa, a informacje absorbowane przez ten rynek darmowe, to ten – jako twór za natury konkurencyjny – nie miałby racji bytu. Kiedy bowiem założenia Famy byłyby zgodne z rzeczywistością, każdy poinformowany inwestor czułby, że może przestać ponosić koszty pozyskania informacji i zachowywać się tak jak inwestor, których ich dotąd nie ponosił. W ten sposób wszyscy przyjmowałiby cenę za daną i nikt nie dysponowałby dodatkowymi informacjami, co uniemożliwiłoby

¹²³ Stiglitz, J. E., *Informacja i zmiana paradygmatu w ekonomii*, „Gospodarka Narodowa”, 3/2004, s. 89.

¹²⁴ Fama, E. F., *Efficient Capital Markets...*, op.cit., s. 387.

¹²⁵ Zob. Bagehot, W., *The Only Game in Town*, „Financial Analysts Journal”, 27:2, 1971, s. 12–14; Walter Bagehot był pseudonimem, pod którym artykuł opublikował Jack Treynor.

¹²⁶ Storkenmaier, A., *Financial Markets and Public Information*, KIT Scientific Publishing, 2011, s. 8–9.

¹²⁷ Stiglitz, J. E., *Informacja i zmiana...*, op.cit., s. 94

¹²⁸ Por. Henryk, G., *Analiza zdarzeń...*, op.cit., s. 16.

osiągnięcie przez rynek równowagi.

Fama wprawdzie jasno zaznaczył, że wymieniane przez niego warunki istnienia efektywnego rynku nie mają charakteru obligatoryjnego, są jedynie wystarczające do tego, żeby taki rynek zaistniał, a ich niewystąpienie nie przekreśla efektywności rynkowej, ale zdaniem Grossmana i Stiglitz'a właśnie to sprowadza całą koncepcję do nonsensu. Jeżeli informacja nie byłaby darmowa, ceny nie mogłyby jej w pełni odzwierciedlać. W ten sposób rynek nie mógłby być efektywny. Darmowość informacji nie jest więc warunkiem wystarczającym, ale niezbędnym do tego, aby rynek mógł być efektywny w sensie informacyjnym. Z drugiej strony, jeśli informacja byłaby bezpłatna, przez co wszyscy mieliby do niej dostęp, inwestorzy pozbawieni byłiby bodźców różnicujących przekonania. W efekcie handel na takim rynku mógłby po prostu zaniknąć¹²⁹.

Problem kosztów pozyskania informacji, a tym samym jej powszechności, rodzi poważne konsekwencje zarówno dla praktyki funkcjonowania na rynku, jak i dla jego dynamiki jako całości. Po pierwsze, jeżeli przyjmujemy za prawdę, że otrzymanie informacji może wiązać się z poniesieniem kosztów, wówczas oczywiste staje się, że jedynie część inwestorów będzie skłonna lub będzie w stanie takie koszty ponieść. W efekcie pozostali inwestorzy mogą reagować na wydarzenia rynkowe ze zwłoką. Będzie się tak działo przynajmniej z dwóch powodów. Pierwszy to taki, że inwestorzy będą korzystać z informacyjnej funkcji mechanizmu cenowego. Drugi to fakt, że nawet ta sama informacja będzie docierać do niektórych inwestorów z opóźnieniem z uwagi na to, że korzystają oni z bezpłatnych serwisów informacyjnych¹³⁰. Trzeba jednak zaznaczyć, że Fama pisał o „wystarczającej” liczbie inwestorów, która miałaby powszechny dostęp do informacji. Nie można jednak oszacować, jak liczna grupa inwestorów mogłaby zostać za takową uznana. Jeśli miałby on na myśli inwestorów instytucjonalnych, którzy z racji na dysponowanie odpowiednimi zasobami, mają lepszy dostęp do informacji, to *implicite* zakładałby, że inwestorzy indywidualni nie mogą wpływać na zmniejszenie poziomu efektywności rynkowej.

¹²⁹ Stiglitz, J. E., Grossman S. J., *On the Impossibility...*, op.cit., s. 404; Należy zaznaczyć, że Grossman i Stiglitz wprawdzie zakładają, że różnorodność przekonań inwestorów ma charakter endogeniczny, ale ich zdaniem wynika z faktu ponoszenia wydatków na informację.

¹³⁰ Ciołek, M., *Efektywność informacyjna polskiego rynku akcji*, CeDeWu, Warszawa 2015, s. 33.

Konsekwencje asymetrii informacyjnej – w kontekście rynku kapitałowego jako całości – rozpatrywane są z perspektywy jego płynności, a w szczególności głębokości. Szczegółowo problem ten w swoim modelu zaprezentował Jiang Wang, którego zdaniem wolumen obrotu zawsze skorelowany będzie pozytywnie z bezwzględną zmianą ceny, a korelacja ta przyjmuje wyższe wartości wraz ze wzrostem asymetrii. Wynika to z faktu, że inwestorzy, którzy nie posiadają żadnych unikalnych informacji, nie mogą być pewni motywacji konkurentów, którzy takimi informacjami dysponują. Ponoszą więc oni ryzyko konkurowania z inwestorami, będącymi w posiadaniu nieznanych im informacji. W miarę wzrastania poziomu asymetrii pomiędzy dwoma klasami inwestorów wolumen obrotów powinien spadać¹³¹.

Prawidłowość ta wydaje się związana z opisanym przez Geogra Akerlofa zjawiskiem negatywnej selekcji (ang. *adverse selection*). W warunkach niepewności strona popytowa nie może mieć gwarancji co do realnej wartości przedmiotu zakupu. Akerlof wyjaśniał ten problem między innymi na przykładzie rynku używanych samochodów. Sprzedawcy z oczywistych względów dysponują pełniejszą informacją o historii pojazdów. Strona popytowa natomiast działa w warunkach niepewności. Ponieważ kupujący nie mogą być pewni stanu technicznego auta, będą starali się obniżyć cenę, do poziomu który uwzględni prawdopodobieństwo zakupu pojazdu dobrego lub złego (prawdopodobieństwo to wynosiłoby 50%). W ten sposób sprzedawcy samochodów, będących w dobrym stanie, musieliby dokonywać transakcji poniżej oszacowanej przez nich wartości, w konsekwencji czego zaczęłyby wycofywać się z rynku, na którym pozostaną tylko pojazdy o niskiej jakości¹³². Podobnie zdaje się być w przypadku modelu Wanga. Wzrost asymetrii informacji powoduje jednocześnie pogłębienie problemu negatywnej selekcji. Inwestorzy, którzy nie dysponują prywatną informacją, będą oczekiwać niższej ceny nabycia waloru, aby zrekompensować ryzyko handlu w warunkach asymetrii, co będzie miało przełożenie na poziom wolumenu obrotu. Dlatego – zdaniem Wanga – wolumen ten będzie skorelowany z poziomem

¹³¹ Wang, J., *A Model of Competitive Stock Trading Volume*, „Journal of Political Economy”, 102.1 (1994), s. 127–68, s. 129.

¹³² Zob. Akerlof, G. A., *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, [w:] *An economic theorist's book of tales*, (red.) Akerlof, G. A., Cambridge University Press, 1984.

asymetrii informacji¹³³. Można zwrócić uwagę, że logika wyводу Grossmana i Stiglitz'a prowadzi ich jednak do innych wniosków. Wychodząc od problemu kosztów pozyskania informacji, sądzą oni, że jeżeli takie koszty będą bardzo niskie (lub bardzo wysokie) średnia oraz wariancja wolumenu obrotu będą wynosiły zero. Podobnie będzie działo się w przypadku, w którym inwestorzy otrzymywać będą informacje o wysokiej precyzji. Wprawdzie, jak twierdzą, rynek będzie mógł osiągać stan równowagi, a cena rynkowa ujawni większość informacji, jakie mają inwestorzy poinformowani, ale jednocześnie w kontekście płynności – z uwagi na homogeniczność inwestorów – będzie to rynek płytki¹³⁴.

Przy obecnym poziomie zrozumienia mechanizmów funkcjonowania rynku finansowego asymetrię informacji określić można jako jeden z poważniejszych problemów badawczych. Jej występowanie powodowane jest przede wszystkim kosztami pozyskania informacji, przez które rozumieć należy zarówno kapitał, jak i czas, poświęcony na dotarcie do niej. W rzeczywistości więc, jak można przypuszczać, formułowane przez inwestorów oczekiwania co do przyszłych cen, ufundowane są na danych niepełnych. W konkretnym momencie inwestor posiada jedynie część wiedzy, która może ulec zmianie, ilekroć na rynek napłynie nowa informacja. Ma to niebywale znaczenie w kontekście szeroko dyskutowanej w finansach racjonalności inwestorów rynkowych. Jak twierdzi Małgorzata Snarska, ową racjonalność zakładać można bowiem jedynie w obrębie pewnego podzbioru informacji, jakim dysponuje dany inwestor¹³⁵.

1.5. Efekt momentum

W opozycji do logiki, leżącej u podstaw hipotezy efektywności informacyjnej oraz modelu błędzenia losowego, stoi tak zwany *efekt momentum*, powszechnie określany mianem rynkowej anomalii. Efekt ten zdefiniować można jako „średniookresową tendencję cen akcji do kontynuacji wzrostów lub spadków”¹³⁶. Jego przejawem jest obserwacja, że akcje spółek, których ceny notowały w ostatnim czasie wzrosty, kontynuują ten trend w przyszłości. Z kolei te akcje, których ceny spadały, odnotowują

¹³³ Wang, J., *A Model of Competitive...*, op.cit., s. 129.

¹³⁴ Stiglitz, J. E., Grossman, S. J., *On the Impossibility...*, op.cit., s. 403–404.

¹³⁵ Snarska, M., *Efektywność informacyjna...*, op.cit., s. 41.

¹³⁶ Ibidem, s. 47.

spadki również w kolejnych okresach¹³⁷. Na podstawie tego spostrzeżenia zbudowano dwa typy strategii inwestycyjnych: momentum oraz kontrariańską. Rekomendacje wynikające z przyjęcia założeń inwestowania kontrariańskiego zasadniczo uznać można za sprzeczne z tymi wpływającymi ze strategii momentum. Zgodnie z tą koncepcją, należy bowiem pozyskiwać akcje spółek, które dotąd traciły na wartości, a sprzedawać te, które odnotowywały w ostatnim czasie wzrost. W istocie jednak różnica związana jest z brany pod uwagę horyzontem inwestycyjnym oraz okresem przyjętym do szacowania historycznych stóp zwrotu. W przypadku strategii kontrariańskiej mowa jest albo o krótkich okresach (tygodnia¹³⁸, miesiąca¹³⁹) albo bardzo długich (kilkuletnich¹⁴⁰)¹⁴¹. Wynika to z faktu, że średnioterminowy charakter efektu momentum związany jest z okresem od kilku do kilkunastu miesięcy. W przypadku krótszych lub dłuższych okresów widoczna jest tendencja do odwracania się kierunku zmian cen walorów¹⁴².

Ten drugi przypadek opisany został przez prekursorów poruszanego zagadnienia: Wenera DeBonda i Richarda Thaler, którzy w artykule z 1985 roku scharakteryzowali efekt zwycięzców i przegranych. Ich analiza uwzględniała portfel składający się z „przegranych akcji”, to jest takich spółek, których akcje w ostatnich 36 miesiącach znajdowały się w dolnym decylu stóp zwrotu oraz z „akcji wygranych” (znajdujących się w tym samym okresie w górnym decylu). Wyniki ich badania wskazywały na to, że w ciągu kolejnych trzech lat od momentu skonstruowania portfela, akcje przegrane przekroczyły wyniki akcji wygranych o ponad 24%¹⁴³. Tego typu badania zaproponowali również Narasimhan Jegadeesh i Sheridan Titman, którzy wykazali, że budowanie portfeli w oparciu o strategię momentum pozwoliło wygenerować znaczące

¹³⁷ Wójtowicz, T., *Efekt momentum na GPW w Warszawie w latach 2003–2010*, „Ekonomia Menedżerska”, 2011, 9, s. 63–74, s. 64.

¹³⁸ Zob. Jegadeesh, N., Titman, S., *Overreaction, Delayed Reaction, and Contrarian Profits*, „Review of Financial Studies”, 1995, 8/4, s. 973–993; Lehmann, B., *Fads, Martingales, and Market Efficiency*, „Quarterly Journal of Economics”, 105/1, s. 1–28.

¹³⁹ Zob. Jegadeesh, N., *Evidence of Predictable Behavior of Security Returns*, „Journal of Finance”, 1990, 45/ 3, s. 881–898; Jegadeesh, N., Titman, S., *Overreaction, Delayed Reaction, and Contrarian Profits*, „Review of Financial Studies”, 1995, 8/ 4, s. 973–993.

¹⁴⁰ Zob. De Bondt, W. F. M., Thaler, F., *Does the Stock Market Overreact?*, „The Journal of Finance”, 1985, 40/3, s. 793.

¹⁴¹ Szyszka, A., *Zjawisko kontynuacji stóp zwrotu na Gieldzie Papierów Wartościowych w Warszawie*, Bank i Kredyt, 2006, 8, s. 37–49, s. 38.

¹⁴² Wójtowicz, T., *Efekt momentum...*, op.cit., s. 64.

¹⁴³ Zob. De Bondt, W. F. M., Thaler, R., *Does the Stock Market Overreact?*, „The Journal of Finance”, 1985, 40/3.

ponadprzeciętne zwroty w okresach od 3 do 12 miesięcy. Przykładowo, selekcja spółek na podstawie ich zwrotu z ostatnich 6 miesięcy i przetrzymywanie ich w portfelu przez analogiczny okres, pozwoliło osiągnąć średnioroczny zwrot na poziomie ponad 12%¹⁴⁴.

Na rynku polskim badania nad efektem momentum przeprowadzili między innymi Adam Szyszka i Tomasz Wójtowicz. Zgodnie z wynikami Szyszki, strategię momentum można uznać za skuteczną w kontekście Giełdy Papierów Wartościowych w Warszawie, ale jej efektywność maleje wraz z okresem utrzymywania akcji w portfelach. Najlepsze rezultaty osiągnięte zostały, kiedy zajmowanie długich lub krótkich pozycji zdeterminowane było 9-miesięcznymi historycznymi stopami zwrotu, przy ich utrzymywaniu przez kolejne 3 miesiące¹⁴⁵. Z kolei wyniki badań Wójtowicza wskazywać mogą na relatywnie krótki okres efektywności omawianej strategii, ponieważ zjawisko momentum znajdowało każdorazowo swoje potwierdzenie jedynie w przypadku 1-miesięcznego okresu formacji. Jednocześnie, jeśli podstawą klasyfikacji były wyniki z ostatniego miesiąca, wszystkie wyniki okazały się dodatnimi zyskami. Dla okresów 3- i 6-miesięcznych wyniki nie były tak jednoznaczne. Im dłuższy był okres formacji portfela, tym mniej stóp zwrotu można było uznać za istotne¹⁴⁶. Powyższa obserwacja, jak i kolejne badania tego autora, sugerować mogą skrócenie czasu trwania efektu momentum¹⁴⁷.

Występowanie efektu momentum niewątpliwie stoi w sprzeczności z istotą modelu błędzenia losowego, a tym samym podaje w wątpliwość twierdzenie Famy o natychmiastowym dyskontowaniu przez rynek pojawiających się na nim informacji. W związku z powyższym istotne staje się scharakteryzowanie czynników, mogących wpływać na to, że na przykład stopy zwrotu w pewnym okresie mają tendencję do autokorelacji. Na tym polu rozważa się dwie – choć wydaje się, że wzajemnie uzupełniające się – grupy okoliczności, które wyjaśniałyby powody opisywanego zjawiska. Pierwsza grupa związana jest z rynkiem jako całością i jego reakcją na

¹⁴⁴ Jegadeesh, N., Titman, S., *Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency*, „The Journal of Finance”, 1993, 48/1 (1993), 65–91, s. 89.

¹⁴⁵ Szyszka, A., *Zjawisko kontynuacji stóp zwrotu...*, op.cit., s. 42. W przytaczanym badaniu brane były pod uwagę kwartalne okresy szacowania stóp zwrotu.

¹⁴⁶ Zob. Wójtowicz, T., *Zjawisko momentum a efekty kalendarzowe*, „AGH Managerial Economics”, 2012, 11, s. 115–25. Co ciekawe, wyniki tych dwóch ostatnich grup uzależnione były w dużej mierze od dnia tygodnia, w którym dokonywano klasyfikacji spółek do portfeli.

¹⁴⁷ Zob. Wójtowicz Tomasz, *Efekt momentum...*, op.cit., s. 63–74.

informację. Druga natomiast uwzględnia czynniki natury psychologicznej, co przenosi środek ciężkości rozważań na proces decyzyjny inwestorów jako jednostek.

W kontekście całego rynku mówi się o zbyt wolnej reakcji na napływające informacje lub o jego nadreaktywności. W pierwszym przypadku, jak twierdzi Adam Szyszka, jeśli sednem problemu byłaby „systematyczna nieprawidłowość w przetwarzaniu informacji”, musielibyśmy przyjąć, że w długim okresie możliwość inwestowania kontrariańskiego jest wynikiem relatywnie słabej reakcji graczy na wiadomości, które napłyną na rynek w przyszłości, a które będą stały w sprzeczności do pierwotnych poglądów inwestorów na temat akcji danej spółki¹⁴⁸. O ile takie spostrzeżenie może rodzić wątpliwości związane z obieraniem konkretnej strategii budowania portfela, to wpisuje się ono w – rozpatrywane w ramach finansów behawioralnych – zagadnienie konserwatyzmu poznawczego inwestorów¹⁴⁹. Z drugiej strony, powodem występowania efektu momentum może być również odwrotne zjawisko, tj. nadreaktywność rynku. W takiej sytuacji, w średnim okresie (kilku lub kilkunastu miesięcy) ceny tak naprawdę przechodziłby proces oddalania się od wewnętrznej wartości waloru (co powodowałaby trend w stopach zwrotu), a w kolejnych okresach rynek dokonywałby korekty. Trzecim wyjaśnieniem może być to, że rynek nie doszacowuje lub przeszacowuje informacje w zależności od ich typu¹⁵⁰. Propozycje te zostaną omówione w dalszej części dysertacji.

¹⁴⁸ Szyszka, A., *Zjawisko kontynuacji stóp zwrotu...*, op.cit., s. 39.

¹⁴⁹ Zob. Zielonka, P., *Gielda i psychologia. Behawioralne aspekty inwestowania na rynku papierów wartościowych*, CeDeWu, Warszawa 2017, s. 67–70.

¹⁵⁰ Por. Szyszka, A., *Zjawisko kontynuacji stóp zwrotu...*, op.cit., s. 39.

Rozdział II – Odbiorca informacji na rynku finansowym

Mechanizmy, opisane w rozdziale pierwszym, prowadzą do pytania o to, w jaki sposób inwestorzy dokonują wyceny aktywów, a tym samym, na podstawie jakich przesłanek podejmują decyzje o lokowaniu swoich środków. Z odpowiedzią na to pytanie ściślej związany jest – opisywany w poprzednim rozdziale – typ efektywności, to jest efektywność alokacyjna. Na potrzeby dalszych rozważań przyjęta zostanie w tym kontekście perspektywa inwestora. Jak zostało już zarysowane, działanie w zgodzie z efektywnością alokacyjną oznacza inwestowanie kapitału w takie przedsiębiorstwa lub projekty, które mogą zaoferować najlepsze warunki, przez co rozumieć należy stopę zwrotu oraz ryzyko, jakim jest ona obciążona¹⁵¹. Zagadnienia te scharakteryzowane zostaną w świetle dwóch paradygmatów. Pierwszy, który opiera się na założeniu o racjonalności zachowań inwestorów, nazwać można teorią neoklasyczną¹⁵². Drugi, który w powszechnym rozumieniu kontrastuje z powyższym założeniem, to paradygmat finansów behawioralnych. Oba te podejścia obrazują odmienne sposoby rozumienia i przetwarzania informacji przez inwestorów giełdowych, co znajduje swoje odzwierciedlenie w wycenianiu wartości instrumentów finansowych.

2.1. Wartość wewnętrzna instrumentu finansowego a jego cena

Warto zauważyć, że podstawowe twierdzenie hipotezy efektywności informacyjnej podlegało ewolucji. Można sądzić, że środek ciężkości początkowych definicji oscylował głównie wokół tematu wartości wewnętrznej (ang. *intrinsic value*)¹⁵³. Problem takiego definiowania efektywności informacyjnej polega jednak na tym, że badacze musieliby mieć wiedzę na temat prawdziwej wartości poszczególnych walorów¹⁵⁴. Sam Fama twierdził jednak, że – w obliczu niepewności – owa wartość pozostaje dla obserwatorów nieznana. Mimo to zakładał, że pewnym przybliżeniem szacunków wartości wewnętrznej poszczególnych walorów jest ich cena, co nie oznacza, że stanowi ona ich idealne

¹⁵¹ Ciołek, M., *Efektywność informacyjna polskiego rynku akcji*, CeDeWu, Warszawa 2015, s. 15.

¹⁵² Określenia tego używa między innymi Profesor Jerzy Gajdka.

¹⁵³ Zob. Buzala, P., *Silna a pólslna forma efektywności informacyjnej rynku kapitałowego – doświadczenia z GPW w Warszawie na podstawie zaleceń analityków*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia”, 2015, nr 862/75, s. 100.

¹⁵⁴ Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna Ekonomiczna, Kraków 2006, s. 18.

odzwierciedlenie. Ponieważ między osobami podejmującymi decyzje inwestycyjne zawsze mogą istnieć niezgodności dotyczące wyceny instrumentów, rzeczywista cena rynkowa i wartość wewnętrzna mogą się od siebie różnić. Fama proces ten określił mianem rynkowego szumu¹⁵⁵. Niemniej, z uwagi na to, że na rynku ceny aktywów oscylują losowo wokół ich wartości fundamentalnych, to – w ujęciu średnim – cenę definiować można jako najlepsze oszacowanie wartości wewnętrznej¹⁵⁶.

Koncepcję dotyczącą losowych ruchów cen, przypadkowości ich rozkładu i braku możliwości ich wcześniejszego oszacowania – w tym samym czasie co Fama – zaproponował również Paul Samuelson w artykułach *Proof that Properly Anticipated Prices Fluctuate Randomly* (1965) oraz *Proof That Properly Discounted Present Values of Assets Vibrate Randomly* (1973). Samuelson zauważył, że jeśli prawidłowo oszacowana przyszła wartość danego instrumentu byłaby znana, to jego cena przesunęłaby się w tym kierunku natychmiast¹⁵⁷. Co więcej, podobnie jak Fama, zakładał, że ceny instrumentów reagują na nowe informacje w sposób, który również określić można jako natychmiastowy¹⁵⁸. Mimo to, analizując sposób kształtowania się cen na rynku, zdefiniował przy tym pojęcie „ceny doskonale efektywnej”, przez co rozumiał mechanizm, w którym ceny odzwierciedlają absolutnie wszystkie informacje. Spełniając ten warunek, instrumenty znajdują się w stanie „ciągłej równowagi”, przy czym stan ten nie jest w czasie stały. Okres między momentem pojawienia się nowej informacji, która zmieni wartość wewnętrzną waloru, a dostosowaniem się do niej ceny, jest okresem nierównowagi. Długość tego okresu jest natomiast przez Samuelsona traktowana jako miara stopnia efektywności cen¹⁵⁹.

Rozpatrywanie problemu efektywności poprzez relację, jaka zachodzi między wewnętrzną wartością instrumentów a ich ceną rynkową, kontynuowana była również w późniejszym okresie. Na przykład Fischer Black w swoim artykule z 1986 roku *Noise*,

¹⁵⁵ Fama, E. F., *The Behaviour of Stock-Market Prices*, „The Journal of Business”, 1965, nr 38/1, s. 36.

¹⁵⁶ Delcey, T., *Samuelson vs Fama on the efficient market hypothesis: The point of view of expertise*, „Economa – History/Methodology/Philosophy”, 2019, nr 9/1 s. 41.

¹⁵⁷ Schuster, T., *News Events and Price Movements. Price Effects of Economic and Non-Economic Publications in the News Media*, „SSRN Electronic Journal”, 2003, s. 6–7.

¹⁵⁸ Erdas, M. L., *Validity of Weak-Form Market Efficiency in Central and Eastern European Countries (CEECs): Evidence from Linear and Nonlinear Unit Root Tests*, „Review of Economic Perspectives”, 2019, nr 19/4, s. 400. [za: Kofarbai i Zubairu, 2016: 261].

¹⁵⁹ Francis, J. C., *Inwestycje. Analiza i zarządzanie*, „WIG-Press”, Warszawa 2000, s. 236.

stwierdził, że rynek efektywny zdefiniować można jako taki, który mieści się we współczynniku dwukrotności, co oznacza, że cena instrumentu powinna być większa niż połowa lub mniejsza niż dwukrotność jego wartości. Współczynnik ten Black oszacował oczywiście w sposób arbitralny, ale stwierdził przy tym, że

intuicyjnie wydaje mi się to jednak uzasadnione, biorąc pod uwagę źródła niepewności co do wartości i siły, które mają skłonność do powrotu ceny do wartości [wewnętrznej]. Według tej definicji prawie wszystkie rynki są sprawne prawie przez cały czas. „Prawie wszystkie” oznacza co najmniej 90%¹⁶⁰.

Należy jednak zaznaczyć, że propozycje Famy i Samuelsona nie są ze sobą tożsame. Fama opierał się bowiem na interpretacji analiz empirycznych, podczas gdy Samuelson wykorzystywał logikę dedukcyjną, koncentrując się przede wszystkim na formalnym aspekcie swoich wniosków. Ponadto, chociaż obaj poruszali problem fluktuacji cen, to scharakteryzowali je poprzez odmienne procesy stochastyczne (w przypadku Famy było to błędzenie losowe, a Samuelsona martyngał). Różne były również instrumenty, które stanowiły przedmiot analizy: Fama zajmował się akcjami, a Samuelson instrumentami typu *futures*. Autorzy ci mieli też odmienny stosunek do normatywnej implikacji swoich modeli. Podczas gdy Fama skupiał się głównie na wiedzy interesującej praktyków rynkowych, Samuelson koncentrował się na wiedzy potencjalnie interesującej rządowych decydentów. Pomimo znacznej różnicy, zarówno między modelami, jak i ich autorami, obie propozycje pokazały jednak, że jeśli inwestorzy mają prawidłowe oczekiwania, może to skutkować przypadkowymi wahaniami cen¹⁶¹.

W celu zgłębienia kwestii związanych z motywacjami inwestorów giełdowych dotyczącymi alokacji kapitału, autorka proponuje zwrócić uwagę właśnie na zagadnienie „prawidłowych oczekiwań”. Fama zauważa bowiem, że większość dostępnych prac opiera się na założeniu, że warunki równowagi rynkowej można określić w kategoriach oczekiwanej stopy zwrotu i w ogólnym ujęciu opisać je za pomocą poniższego modelu:

$$E(\tilde{p}_{j,t+1}|\Phi_t) = [1 + E(\tilde{r}_{j,t+1}|\Phi_t)]p_{jt}, \text{ gdzie:}$$

p_{jt} – cena instrumentu „j” w czasie „t”; $p_{j,t+1}$ – cena w czasie „t+1” łącznie z reinwestycją jakiegokolwiek śródkresowego zwrotu na walorze; $r_{j,t+1}$ – jednookresowy zwrot $(p_{j,t+1} - p_{j,t})/p_{j,t}$; Φ – ogólny symbol dla dowolnego zestawu informacji, który jak się zakłada, jest w pełni odzwierciedlony w cenie „t”.

¹⁶⁰ Black, F., *Noise*, „Journal of Finance”, 1986, nr 41/3, s. 533.

¹⁶¹ Delcey, T., *Samuelson vs Fama...*, op.cit., s. 43–45.

Na podstawie powyższego wzoru stwierdzić można, że wartość równowagi oczekiwanego zwrotu skonstruowana na podstawie danego zestawu informacji, byłaby zdeterminowana konkretną teorią oczekiwanych zwrotów, z jakiej akurat korzysta inwestor. Warunkowy oczekiwany zwrot przedstawiony w powyższym modelu ma sugerować jednak, że niezależnie od modelu oczekiwanych zwrotów, jaki zostanie zastosowany, informacja zawarta w danym zestawie zostanie w pełni zdyskontowana do określenia równowagi oczekiwanych zwrotów. W tym właśnie sensie – jak mówi Fama – informacja jest w pełni odzwierciedlona w procesie formowania się ceny¹⁶². Według Macieja Ciołka definicja ta oznacza, że ceny stanowią odzwierciedlenie wszystkich dostępnych inwestorom informacji, o ile równowaga rynku w sferze oczekiwanych stóp zwrotu, osiągnięta została niezależnie od przyjętego do kalkulacji modelu, oraz przy założeniu, że informacje zostały w pełni wykorzystane¹⁶³.

2.2. Inwestor w świetle neoklasycznej teorii finansów

Jednym z najistotniejszych zagadnień w kontekście zarówno reakcji inwestorów na informację, jak i efektywności alokacyjnej, jest racjonalność uczestników rynku. Jeżeli przyjąć hipotezę rynku efektywnego Famy za taką, która odzwierciedla rzeczywistość, to *implicite* przyjąć należy również twierdzenie o racjonalnym zachowaniu inwestorów¹⁶⁴. Merton Miller oraz Franco Modigliani, którzy analizowali relację, jaka zachodzi między polityką dywidendową spółek a cenami akcji, stwierdzili, że zachowanie racjonalne oznacza, że inwestorzy zawsze wolą bogactwo większe od mniejszego, a to, czy przyrost majątku przybiera formę płatności gotówkowych, czy wzrostu rynkowej wartości posiadanych akcji, pozostaje dla nich obojętne¹⁶⁵. W ogólnym ujęciu – w ramach rynku finansowego – osoba racjonalna może zostać zdefiniowana jako taka, która jest w stanie dokonać poprawnej interpretacji otrzymanej informacji, jak również porównania różnych alternatyw według własnej funkcji użyteczności, na podstawie czego może podjąć decyzję o wyborze optymalnego

¹⁶² Fama, E. F., *Efficient Capital Markets: A Review of Theory and Empirical Work*, „The Journal of Finance”, 1970, nr 25/2, s. 384.

¹⁶³ Zob. Ciołek, M., *Efektywność informacyjna...*, op.cit., s. 30–31.

¹⁶⁴ Zob. Gurgul, H., *Analiza zdarzeń...*, op.cit., s. 15.

¹⁶⁵ Miller, M. H., Modigliani, F., *Dividend Policy, Growth, and the Valuation of Shares*, „The Journal of Business”, 1961, nr 34/4, s. 412.

rozwiązania¹⁶⁶. Założenie racjonalności zachowań inwestorów można uznać za powszechny zabieg nauk ekonomicznych¹⁶⁷. Należy jednak zwrócić uwagę na wynikające z niego implikacje. Częstość przyjmuje się bowiem, że racjonalni uczestnicy rynku stanowią dominującą grupę i dzięki temu udaje im się szybko i skutecznie wyeliminować wszelkie przejawy irracjonalności ze strony innych uczestników rynku. Zdaniem Famy, domniemywać można, że w rezultacie takiego mechanizmu rynek powinien zachowywać się tak, jakby wszyscy jego uczestnicy działali racjonalnie¹⁶⁸. Z uwagi więc na wagę tego zagadnienia, koncepcja racjonalnego inwestora wymaga głębszej charakterystyki. Autorka proponuje spojrzeć na nią przez pryzmat trzech zagadnień, to jest: oczekiwanej użyteczności, jaką inwestor przypisuje podejmowanym przez siebie decyzjom, preferowanemu przez niego stosunkowi do ryzyka oraz sposobowi, w jaki szacuje on prawdopodobieństwo zdarzeń.

2.2.1. Koncepcja oczekiwanej użyteczności

Mianem racjonalnego określić można takiego inwestora, który dąży do maksymalizacji swojej oczekiwanej użyteczności. Koncepcję tę zdefiniować można jako ważoną sumę prawdopodobieństw wystąpienia danych zdarzeń oraz ich użyteczności i opisać wzorem zaproponowanym przez Daniela Bernoulliego¹⁶⁹:

$$EU = \sum_i p_i U(x_i), \text{ gdzie:}$$

EU – oczekiwana użyteczność; p_i – prawdopodobieństwo wystąpienia i -tego wyniku (zdarzenia); $U(x_i)$ – użyteczność i -tego wyniku (zdarzenia).

Powszechnie za twórców samej koncepcji uznaje się jednak Johna von Neumanna i Oskara Morgensterna, którzy uzupełnili powyższą definicję o postulaty związane z aksjomatami racjonalności. Zaproponowany przez nich model podejmowania decyzji w warunkach ryzyka wydaje się ogólnie aprobowany przez przedstawicieli paradygmatu neoklasycznego¹⁷⁰. Równocześnie stanowi on przedmiot wielu zarzutów ze strony

¹⁶⁶ Szyszka, A., *Behavioural Finance and Capital Markets: How Psychology Influences Investors and Corporations*, Palgrave Macmillan, Nowy Jork 2013, s. 9.

¹⁶⁷ Zob. Gurgul, H., *Analiza zdarzeń...*, op.cit., s. 14.

¹⁶⁸ Zob. Fama, E. F., *The Behaviour of Stock-Market...*, op.cit., s. 39.

¹⁶⁹ Zielonka, P., *Gięda i psychologia. Behawioralne aspekty inwestowania na rynku papierów wartościowych*, CeDeWu, Warszawa 2017, s. 33.

¹⁷⁰ Gajdka, J., *Behawioralne finanse przedsiębiorstw. Podstawowe podejścia i koncepcje*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2013, s. 22.

środowisk podających w wątpliwość założenie o racjonalności inwestorów rynkowych. Dlatego też wymagają one przynajmniej krótkiego objaśnienia.

Pierwszym ze wspomnianych aksjomatów, jaki opisany został przez Morgensterna i von Neumanna, jest „aksjomat porównywalności” (nazywany również aksjomatem kompletności lub zwrotności). Zgodnie z nim założyć można, że uczestnicy rynku zdolni są do porównania alternatywnych rozwiązań i wyników. W związku z tym – niejako *implicite* – przyjąć można, że dysponują oni kompetencjami, schematami i narzędziami, umożliwiającymi im dokonywanie takowych porównań. Przegląd literatury z zakresu finansów behawioralnych pozwala stwierdzić, że owo założenie stanowi jeden z poważniejszych punktów spornych między przedstawicielami obu omawianych nurtów, co opisane zostanie w dalszej części pracy. Drugi aksjomat określany jest mianem „aksjomatu przechodniości” (nazywanego również aksjomatem tranzytywności). W tym przypadku zakłada się, że wybory, przed którymi stoi inwestor, stanowią dla niego pewną logiczną i konsekwentną całość. W sytuacji bowiem, w której ma on do wyboru trzy alternatywy: A, B oraz C, a preferuje opcję A nad B oraz B nad C, to – mając ostatecznie do wyboru opcję A lub C – preferować powinien opcję A nad C¹⁷¹. W tym sensie, preferencje racjonalnych osób mają charakter przechodni. Trzeci aksjomat to „zasada niezależności”, która mówi o tym, że jeśli inwestor uznaje dwie opcje A i B za pozbawione ryzyka i równie atrakcyjne, to powinien pozostać obojętny wobec wyboru. Aksjomat ten określany bywa również mianem zasady eliminacji. Jego istota polega na stwierdzeniu, że wybór między różnymi opcjami zależeć powinien od różnicy w ich wynikach. Oznacza to, że jeżeli jakieś alternatywy zaowocują tym samym wynikiem, to nie powinny mieć wpływu na preferencje inwestora¹⁷². Kolejnym aksjomatem jest „ekwiwalent stałości” lub inaczej – zasada niezmienności. Założenie to związane jest ze sposobem prezentacji opcji wyboru. Innymi słowy, atrakcyjność wyborów, przed którymi stoi inwestor, nie powinna być determinowana ich strukturą, o ile jego wynikom odpowiadają takie same prawdopodobieństwa. Piotr Zielonka podaje w tym kontekście następujący przykład:

¹⁷¹ Zielonka, P., *Gięda i psychologia...*, op.cit., s. 33.

¹⁷² Gajdka, J., *Behawioralne finanse przedsiębiorstw...*, op.cit., s. 24.

dwupoziomowa loteria z 50% szansą wygranej na każdym poziomie, zakończona w przypadku sukcesu studolarową wypłatą, powinna dla decydenta oznaczać to samo, co jednopoziomowa loteria z 25% szansą wygranej i wypłatą identyczną jak w przypadku pierwszej loterii¹⁷³.

Różny sposób prezentacji problemu powinien więc pozostawać bez wpływu na preferencje decyzyjne. Konsekwentnie, również zmiana układu przedstawienia alternatyw nie powinna rzutować na ostateczny wybór¹⁷⁴. Za podstawowy aksjomat uznaje się również „warunek ciągłości”, który mówi o tym, że jeśli trzy dowolne wyniki uporządkujemy w skali od najmniej do najbardziej preferowanego, to wówczas stworzyć można taką loterię, w której za każdym razem skrajne wyniki równoważne będą wynikowi środkowemu¹⁷⁵. W kontekście aksjomatów racjonalności Jerzy Gajdka wymienia także „zasadę dominacji”, zgodnie z którą przyjąć można, że jeśli w przynajmniej jednym aspekcie opcja A jest lepsza od opcji B, a w pozostałych aspektach opcje te są przynajmniej jednakowo dobre, to inwestor powinien preferować opcję A nad opcją B¹⁷⁶.

2.2.2. Stosunek inwestorów do ryzyka

Powyższa charakterystyka, chociaż przybliża sposób, w jaki inwestor – określany mianem racjonalnego – powinien podejmować decyzje, nie uwzględnia jednego z największych problemów, przed którym ów inwestor stoi, to jest problemu ryzyka. Należy zaznaczyć, że zagadnieniu temu poświęconych jest wiele prac badawczych i modeli z zakresu finansów, tworzonych zarówno z perspektywy pojedynczego inwestora, jak i sposobów wyceny instrumentów czy kształtowania się równowagi rynkowej. Celem tej części pracy nie będzie jednak szczegółowy opis koncepcji i punktów widzenia związanych z ryzykiem inwestycyjnym, a zarysowanie obrazu inwestora, który kształtowany jest przez założenia, stosowane w ramach neoklasycznego paradygmatu finansów.

¹⁷³ Zielonka, P., *Gięda i psychologia...*, op.cit., s. 34.

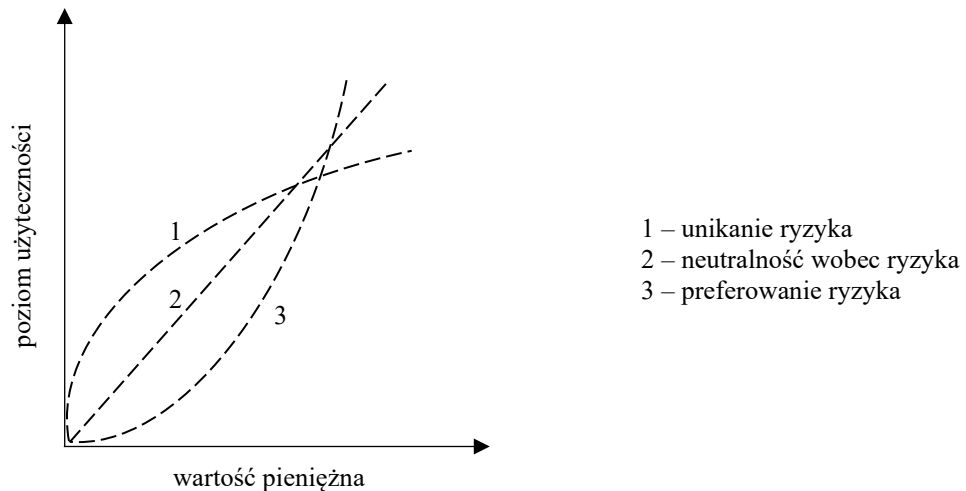
¹⁷⁴ Gajdka, J., *Behawioralne finanse przedsiębiorstw...*, op.cit., s. 24.

¹⁷⁵ Zielonka, P., *Gięda i psychologia...*, op.cit., s. 34.

¹⁷⁶ Gajdka, J., *Behawioralne finanse przedsiębiorstw...*, op.cit., s. 24.

Za punkt wyjścia do rozważań nad tym tematem przyjąć można powszechnie znaną koncepcję funkcji użyteczności. W jej ramach zakłada się, że inwestor postrzega pieniądź poprzez jego wartość użytkową (co określić można również jako użyteczność bogactwa), przez co każda dodatkowa jednostka pieniądza, rozumiana będzie jako wzrost użyteczności. Z punktu widzenia pojedynczego inwestora funkcja ta kształtować będzie się zgodnie z zasadą malejącej krańcowej użyteczności. Sformułowanie to definiować można jako mechanizm, w którym każdy kolejny przyrost pieniądza, wprawdzie zwiększać będzie użyteczność w ujęciu całkowitym, ale stopień tego zwiększania będzie coraz mniejszy¹⁷⁷. Racjonalny inwestor powinien zachowywać się w myśl zasady maksymalizowania użyteczności. Poprzestanie jednak na takiej charakterystyce, budowałoby monolityczny obraz uczestników rynku, co jednocześnie stać mogłoby w sprzeczności do zdefiniowanego wcześniej problemu wahania cen instrumentów wokół wartości wewnętrznej. Sytuacja jednak kształtować będzie się inaczej, jeśli do powyższych rozważań włączony zostanie problem ryzyka. Dzięki temu inwestorów sklasyfikować można jako tych, którzy cechują się awersją, neutralnością, bądź skłonnością do ryzyka.

Rysunek 2. *Funkcja użyteczności, a stosunek inwestora do ryzyka*



Źródło: Dębski, W., *Rynek finansowy i jego mechanizmy, Podstawy teorii i praktyki*, PWN, Warszawa 2014.

¹⁷⁷ Dębski, W., *Rynek finansowy i jego mechanizmy, Podstawy teorii i praktyki*, PWN, Warszawa 2014, s. 512.

Pierwszy przypadek – awersja do ryzyka – jest generalnym założeniem ekonomii w kontekście podejmowania decyzji w warunkach niepewności¹⁷⁸. Oczywiście niechęć do podejmowania ryzyka nie jest równoznaczna z brakiem motywacji do uczestniczenia w procesach rynkowych w ogóle. Jednakże w przypadku takiego inwestora, ocena prawdopodobieństwa zysku musi być na tyle duża, aby pokonać naturalną niechęć do ryzyka. Im wyższy jest poziom jego awersji, tym wyżej musi takową szansę ocenić, by podjąć decyzję o alokacji środków. Ponadto inwestor charakteryzujący się takim podejściem, aby zmniejszyć poziom ryzyka, skłonny będzie poświęcić część oczekiwanej stopy zwrotu¹⁷⁹. Oznacza to, że może wyrazić chęć wybrania pewnej wypłaty i zrezygnowania z udziału w loterii (cechującej się na przykład dużą wariancją), której oczekiwana wartość byłaby równa wysokości wypłaty gwarantowanej¹⁸⁰. W takiej sytuacji funkcja użyteczności przyjmuje postać linii wklęsłej¹⁸¹. Jest to związane ze wspomnianą malejącą krańcową użytecznością pieniądza, co oznacza, że wraz ze wzrostem jego wartości wzrost użyteczności maleje¹⁸².

Z kolei inwestor skłonny do podejmowania ryzyka powinien wybrać udział w loterii, nawet jeśli pewna wypłata jest jedną z opcji do wyboru¹⁸³. Oznacza to, że może on angażować kapitał w przedsięwzięcie, którego prawdopodobieństwo poniesienia straty – z matematycznego punktu widzenia – ocenił wyżej niż prawdopodobieństwo zysku, a im jego preferencja ryzyka jest silniejsza, tym wyższe musi być ryzyko, aby podjąć decyzję o odstąpieniu od inwestycji. Z drugiej jednak strony, inwestorzy o takiej skłonności, oczekiwać będą w zamian odpowiednio wyższej stopy zwrotu¹⁸⁴. Obojętność inwestora wobec ryzyka uwidoczni się natomiast w sytuacji braku preferencji wobec powyższego wyboru¹⁸⁵. W wyniku tego będzie się on kierował głównie potrzebą

¹⁷⁸ Tversky, A., Kahneman, D., *Advances in Prospect-Theory – Cumulative Representation of Uncertainty*, „Journal of Risk and Uncertainty”, 1992, nr 5/4, s. 298.

¹⁷⁹ Dębski, W., *Rynek finansowy...*, op.cit., s. 512–514.

¹⁸⁰ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 34.

¹⁸¹ Dobrowolski, K., *Teoria rynków efektywnych i model racjonalnego inwestora – od warunków ryzyka do warunków konfliktu*, „Współczesna Gospodarka”, 2015, nr 6/2, s. 4.

¹⁸² Gajdka, J., *Behawioralne finanse przedsiębiorstw...*, op.cit., s. 23 [za: Jajuga, 2006: 195–196]

¹⁸³ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 34.

¹⁸⁴ Dębski, W., *Rynek finansowy...*, op.cit., s. 512–514.

¹⁸⁵ Por. Zielonka, P., *Gielda i psychologia...*, op.cit., s. 34.

maksymalizacji oczekiwanej wartości¹⁸⁶. Ponieważ w tym przypadku funkcja użyteczności przyjmuje kształt liniowy, to przyrost użyteczności – związany ze zmianą poziomu bogactwa o jedną jednostkę – oceniany jest przez niego tak samo, niezależnie od tego, czy dotyczy on początku, czy dalszych punktów, znajdujących się na krzywej¹⁸⁷.

Inwestorzy charakteryzują się awersją do ryzyka. W związku z tym istotnym pytaniem staje się to, w jaki sposób owo ryzyko minimalizować. Odpowiedź na to pytanie umiejscowić można w kontekście teorii zaproponowanej w 1952 roku przez Harrego Markowitza, to jest teorii portfelowej. Markowitz zauważył, że inwestor, podejmując decyzję, styka się z dwoma rodzajami parametrów: oczekiwaną stopą zwrotu oraz ryzykiem, a celem jego aktywności jest podjęcie takiej decyzji, która przy dowolnie wybranym poziomie ryzyka, będzie charakteryzowała się maksymalną stopą zwrotu. Jednocześnie wskazał, że zredukowanie poziomu ryzyka możliwe jest poprzez dywersyfikację aktywów. Dzięki temu mechanizmowi możliwe jest zmniejszenie zmienności stopy zwrotu względem założonej wartości oczekiwanej¹⁸⁸. Markowitz stwierdził bowiem, że: „jeżeli dwa oryginalne portfele mają taką samą wariancję, wówczas zwykle wariancja wynikowego (złożonego) portfela będzie mniejsza niż wariancja któregośkolwiek portfela pierwotnego¹⁸⁹”. Zaznaczył on również, że dla zmniejszenia wariancji nie wystarczy zwiększanie liczby aktywów w portfelu. Kluczowe jest bowiem unikanie alokowania kapitału w instrumenty charakteryzujące się wysoką kowariancją¹⁹⁰. Z koncepcji zaproponowanej przez Markowitza wynikają dwie istotne implikacje.

Po pierwsze, aby zminimalizować ryzyko, inwestorzy powinni łączyć aktywa w portfele. Po drugie, ryzyko aktywów powinno być mierzone jako udział w ryzyku portfela, a nie jednostkowo¹⁹¹. Zdefiniował on tym samym pojęcie portfela efektywnego, które rozumieć można jako kombinację walorów, która charakteryzować będzie się:

¹⁸⁶ Dobrowolski, K., *Teoria rynków efektywnych i model racjonalnego inwestora – od warunków ryzyka do warunków konfliktu*, „Współczesna Gospodarka”, 2015, nr 6/2, s. 4.

¹⁸⁷ Dębski, W., *Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki*, PWN, Warszawa 2014, s. 514.

¹⁸⁸ Francis, J. C., *Inwestycje. Analiza i zarządzanie*, WIG-Press, Warszawa 2000, s. 251.

¹⁸⁹ Markowitz, H. M., *Portfolio Selection*, „The Journal of Finance”, 1952, nr 7/1, s. 89–90.

¹⁹⁰ Ibidem, s. 89.

¹⁹¹ Brigham, E. F., Gapenski, L. C., *Zarządzanie Finansami*, Polskie Wydawnictwo Ekonomiczne, Warszawa, 2000, T.1, s. 34.

największą oczekiwaną stopą zwrotu wśród portfeli o tym samym poziomie ryzyka lub najniższym poziomem ryzyka w grupie portfeli o tej samej oczekiwanej stopie zwrotu¹⁹².

Chociaż teoria Markowitza określiła sposób mierzenia ryzyka inwestycyjnego, to nie opisała precyzyjnie relacji, jaka zachodzi między ryzykiem a wymaganymi stopami zwrotu¹⁹³. Refleksję nad tym problemem kontynuowało niezależnie czterech badaczy: Jack L. Treynor, John V. Lintner, Jan Mossin oraz William F. Sharpe. Efektem ich pracy było powstanie klasycznego modelu wyceny aktywów kapitałowych CAPM (ang. *Capital Asset Pricing Model*). Podobnie jak hipoteza efektywności informacyjnej, tak i ten model stał się przedmiotem dyskusji, spotykając się z krytyką głównie ze strony teoretyków rynku. Należy jednak przyznać, że jednocześnie przez praktyków rynkowych – jest on jednym z najchętniej stosowanych¹⁹⁴.

William Sharpe, wyprowadzając warunki równowagi rynku kapitałowego, odwołał się do dwóch założeń. Po pierwsze, uznał istnienie jednolitej, wspólnej dla wszystkich stopy wolnej od ryzyka oraz że wszyscy inwestorzy mogą dokonywać wymiany handlowej środków na tych samych warunkach. Po drugie, Sharp przyjął założenie o jednorodności oczekiwań inwestorów. Miał przy tym na myśli, że zgadzają się oni co do perspektyw różnych inwestycji, to jest ich: oczekiwanych wartości, odchyłeń standardowych oraz współczynników korelacji. Stwierdził również, że przy tak sformułowanych warunkach, każdy inwestor postrzegać będzie swoje alternatywy w ten sam sposób. Chociaż Sharpe zwrócił uwagę na restrykcyjność swojego stanowiska, to stwierdził jednocześnie, że „właściwym sprawdzianem teorii nie jest realizm jej założeń, ale akceptowalność jej implikacji¹⁹⁵”.

Wykorzystywany dziś powszechnie model CAPM, nie został zaprezentowany przez żadnego z czterech badaczy, uznawanych za jego autorów. W rzeczywistości przedstawiony został on przez Famę w artykule z 1968 roku *Risk, Return and*

¹⁹² Francis, J. C., *Inwestycje...*, op.cit., s. 251.

¹⁹³ Brigham, E. F., Gapenski, L. C., *Zarządzanie finansami*, op.cit., s. 34.

¹⁹⁴ Jajuga, K., *Równowaga rynku kapitałowego – pół wieku historii rodziny CAPM*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia”, 2013, nr 768/63, s. 182.

¹⁹⁵ Sharpe, W. F., *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, „The Journal of Finance”, 1964, nr 19/3, s. 433–434.

*Equilibrium: Some Clarifying Comments*¹⁹⁶. Zapisać można go za Krzysztofem Jajugą w następujący sposób¹⁹⁷:

$$E(R) = R_f + \beta(R_M - R_f), \text{ gdzie:}$$

$E(R)$ - oczekiwana stopa zwrotu; R_f - stopa wolna od ryzyka; R_M - stopa zwrotu z portfela rynkowego; β - współczynnik beta.

Dzięki modelowi CAPM – analizie problemu ryzyka z perspektywy paradygmatu neoklasycznego – zwrócić można uwagę na kilka istotnych kwestii. Po pierwsze w jego ramach zakłada się, że na rynku występują papiery wartościowe pozbawione ryzyka, a inwestorzy, znając związane z nimi oczekiwane stopy zwrotu, dzielą portfele między aktywa tego typu, a pozostałe aktywa obciążone ryzykiem¹⁹⁸. Po drugie, inwestor podejmujący ryzyko oczekiwać będzie – wynikającej z tego faktu – premii. Po trzecie, dodanie do portfela kolejnych aktywów, może wpłynąć na ogólny poziom jego ryzyka. Najczęściej mówi się w tym kontekście o ryzyku systematycznym (lub też niedywersyfikowalnym) oraz specyficznym (lub też dywersyfikowalnym). Źródeł pierwszego typu ryzyka poszukiwać można w szerokich zmianach gospodarczych, społecznych czy politycznych. Wynikająca z nich zmienność stóp zwrotu może więc do pewnego stopnia dotyczyć wszystkich lub przynajmniej większości notowanych instrumentów. Drugi typ ryzyka związany jest natomiast z konkretnym emitentem. Jako przykłady podać można tutaj: strajki, nietrafione inwestycje, czy też błędne decyzje zarządcze. W tym przypadku zmienność walorów nie będzie miała więc charakteru, który określić można by jako systemowy¹⁹⁹. Należy zauważyć, że przypisanie ryzyka do którejś z klas, zależne jest również od tego, jaka relacja zachodzi między danym aktywem a portfelem rynkowym. Jeżeli zmienność ceny konkretnego waloru jest niezależna od rynku, to dodanie go do portfela nie powiększy jego ryzyka w sposób znaczący. W związku z tym ryzyko to określić można jako specyficzne dla emitenta, a więc możliwe do zdywersyfikowania. Natomiast, jeśli sytuacja jest odwrotna (to znaczy dynamika ceny waloru jest analogiczna do dynamiki rynku), wówczas dodatnie takiego

¹⁹⁶ Fama, E. F., *Risk, Return and Equilibrium: Some Clarifying Comments*, „The Journal of Finance”, 1968, nr 23/1, s. 29–40.

¹⁹⁷ Zob. Jajuga, K., *Równowaga rynku kapitałowego – pół wieku historii rodziny CAPM*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia”, 2013, nr 768/63.

¹⁹⁸ Damodaran, A., *Finanse korporacyjne. Teoria i praktyka*, Wydawnictwo Helion, Gliwice 2017, s. 280.

¹⁹⁹ Francis, J.C., *Inwestycje...*, op.cit., s. 291–292.

waloru zwiększy poziom ryzyka całego portfela, a więc ryzyko ma charakter rynkowy. W przytaczanym modelu ryzyko to mierzone jest wskaźnikiem beta²⁰⁰.

Jednym z zarzutów, jaki stawiany jest modelowi CAPM jest fakt, że w istocie, analizowane w jego ramach ryzyko, oparte jest na funkcji jednego czynnika, tj. zmienności dochodów z walorów względem dochodów rynkowych (jest to więc tak zwany model jednoczynnikowy). Istnieją jednak przesłanki, by sądzić, że relacja między ryzykiem a dochodem ma charakter bardziej złożony – to znaczy, związana jest z funkcją większej liczby czynników²⁰¹. Dodatkowo wielu autorów wskazuje na wspomnianą zbyt dużą restrykcyjność założeń modelu²⁰². Na kanwie tych spostrzeżeń badacze zaczęli proponować rozwiązania alternatywne wobec modelu CAPM. I tak w 1979 roku Steve A. Ross opisał arbitrażowy model wyceny – APM (ang. *The Arbitrage Theory of Capital Asset Pricing – APT*). Niewątpliwą zaletą tej koncepcji jest fakt uwzględnienia wpływu więcej niż jednej zmiennej na poziom zwrotu z poszczególnych walorów. W ramach APT uwzględnić można bowiem dowolną liczbę czynników ryzyka, a co za tym idzie – dochód rozumieć można jako funkcję kilku faktorów. Wskazywaną w literaturze przedmiotu wadą tej propozycji jest jednak to, że autor nie zdefiniował, które czynniki powinny zostać wzięte pod uwagę podczas analiz, ani jaka powinna być ich ostateczna liczba. Innymi słowy – od strony ekonomicznej – owe czynniki nie zostały zidentyfikowane²⁰³.

Za odpowiedź z kolei na ten problem uznać można modele zaproponowane przez Eugena Fama i Kenneth'a Frencha. W 1993 roku, budując tak zwany model trójczynnikowy, wskazali, że oczekiwana stopa zwrotu zależna jest od rynkowej stopy zwrotu oraz zwrotów z portfeli, których zadaniem jest naśladowanie dodatkowych czynników ryzyka. Pierwszy z tych czynników jest różnicą między zwrotami z portfela małych akcji i portfela dużych akcji (SMB). Drugi to różnica między zwrotami z portfela o wysokiej wartości księgowej do wartości rynkowej a portfelami o niskiej wartości tego

²⁰⁰ Damodaran, A., *Finanse korporacyjne...*, op.cit., s. 281–284.

²⁰¹ Brigham, E. F., Gapenski, L. C., *Zarządzanie finansami*, op.cit., s. 133–134.

²⁰² Zob. Dębski, W., *Rynek finansowy...*, op.cit., s. 534.

²⁰³ Brigham, E. F., Gapenski, L. C., *Zarządzanie finansami*, op.cit., s. 134–136.

wskaźnika (HML)²⁰⁴. Kilka lat później, ci sami autorzy zbudowali model, w którym do powyższej listy dodali jeszcze dwa czynniki, to jest: różnicę między stopami zwrotu ze zdywersyfikowanego portfela akcji o wysokiej i niskiej rentowności operacyjnej (RMW) oraz różnicę między stopami zwrotu ze zdywersyfikowanych portfeli akcji firm o niskich i wysokich inwestycjach (CMA)²⁰⁵.

2.2.3. Szacowanie prawdopodobieństwa

Ważnym elementem decyzji związanych z maksymalizacją oczekiwanej użyteczności jest prawidłowe oszacowanie prawdopodobieństwa wystąpienia przyszłych zdarzeń. Reprezentanci nurtu neoklasycznego zakładają, że racjonalny inwestor posługuje się w tym zakresie regułą Bayes'a. Dodatkowo przyjmując, że uczestnicy rynku dokonują wnioskowania na podstawie rachunku prawdopodobieństwa, twierdzą również, że „wraz z napływem nowych informacji ludzie potrafią automatycznie aktualizować prawdopodobieństwo zajścia danych zdarzeń²⁰⁶”. Schemat modelu Bayes'a zapisać można w następujący sposób²⁰⁷:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ oraz } P(B|A) = \frac{P(B \cap A)}{P(A)} \Rightarrow P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$P(A|B)$ – prawdopodobieństwo warunkowe zajścia zdarzenia A pod warunkiem zajścia zdarzenia B;
 $P(B|A)$ – prawdopodobieństwo warunkowe zajścia zdarzenia B pod warunkiem zajścia zdarzenia A;
 $P(A \cap B)$ – prawdopodobieństwo zajścia zdarzenia łącznego A oraz B; $P(B)$ – prawdopodobieństwo zajścia zdarzenia B; $P(A)$ – prawdopodobieństwo zajścia zdarzenia A.

Problem szacowania prawdopodobieństwa jest jednym z punktów spornych między przedstawicielami omawianych paradygmatów. Badacze związani z nurtem behawioralnym zgłaszają bowiem w tym kontekście fundamentalne zastrzeżenia. Po pierwsze wskazują, że sposób szacowania przez graczy rynkowych prawdopodobieństwa wystąpienia danego zdarzenia nie odpowiada regule Bayes'a, a człowiek ma tendencję do przeszacowywania lub niedoszacowywania możliwości wystąpienia niektórych zdarzeń, w zależności od indywidualnej oceny sytuacji.

²⁰⁴ Fama, E. F., French, K. R., *Common risk factors in the returns of stocks and bonds*, „Journal of Financial Economics”, 1993, nr 33/1, s. 3–56.

²⁰⁵ Zob. Fama, E. F., French, K. R., *A five-factor asset pricing model*, „Journal of Financial Economics”, 2015, nr 116, s. 1–22.

²⁰⁶ Czerwonka, M., Gorlewski, B., *Finanse behawioralne. Zachowania inwestorów i rynku*, Oficyna Wydawnicza Szkoły Głównej Handlowej, Warszawa 2012, s. 12.

²⁰⁷ Borowski, K., *Finanse behawioralne. Modele*, Wydawnictwo Difin, Warszawa 2014, s. 27.

Po drugie, podają w wątpliwość zdolności inwestorów do wykorzystywania opisanego modelu w praktyce rynkowej. Kluczową barierą są tutaj różnego typu heurystyki i uproszczenia poznawcze, które utrudniają podejmowanie decyzji zgodnie z regułami matematycznymi. Dotyczy to zwłaszcza sytuacji, w których czas na ich podjęcie jest ograniczony. Wyniki tych spostrzeżeń wyrażone zostały przede wszystkim poprzez alternatywną propozycję funkcji prawdopodobieństwa, która omówiona zostanie w dalszej części rozdziału.

Z powyższej charakterystyki wysnuć można wniosek, że racjonalnych inwestorów rynkowych charakteryzują trzy kluczowe cechy. Po pierwsze, zawsze dokonują oni wyborów w taki sposób, aby zmaksymalizować oczekiwaną użyteczność. Po drugie, najczęściej wykazują oni awersję do ryzyka, co oznacza, że podejmują je tylko wówczas, kiedy mogą osiągnąć dzięki temu większe korzyści. Po trzecie, w odpowiedzi na pojawiające się informacje, modyfikują oni swoje przekonania, związane z prawdopodobieństwem wystąpienia zdarzeń, zgodnie z regułą Bayes'a. W świetle finansów neoklasycznych, inwestor charakteryzuje się więc spójnym zestawem preferencji, a jego kryterium decyzyjne opiera się na umiejętności oszacowania – dla wszystkich rozpatrywanych alternatyw – poziomu użyteczności, której funkcja cechuje się następującymi właściwościami: (1) użyteczność rośnie wraz z dochodem, tzn. marginalna użyteczność dochodu jest zawsze dodatnia; (2) krańcowa użyteczność dochodu jest natomiast malejąca.

Można stwierdzić, że na podstawie tych założeń od lat toczy się dyskusja między zwolennikami finansów neoklasycznych a przedstawicielami nurtu behawioralnego. Ci drudzy wskazują na przykład, że model Bayes'a nie jest adekwatny dla opisywania rzeczywistych wyborów w warunkach niepewności²⁰⁸. Jednocześnie podkreślają oni, że zasada maksymalizacji oczekiwanej użyteczności – w przypadku niektórych sytuacji decyzyjnych – ma charakter kontrfaktyczny, ponieważ proces decyzyjny człowieka obarczony jest inklinacjami psychologicznymi, które – w klasycznym rozumieniu –

²⁰⁸ De Bondt, W. F. M., Thaler R., *Does the Stock Market Overreact?*, „The Journal of Finance”, 1985, nr 40/3, s. 793.

skutkują różnego typu anomaliami²⁰⁹. Ponadto, twórcy aksjomatów racjonalności wydają się w pewnym miejscu przeceniać ludzkie zdolności. Przykładowo, przyjęta w ramach hipotezy racjonalności zasada niezmienności, niejako zakłada, że osoba podejmująca decyzje powinna umieć „przeniknąć” formę prezentacji problemu, żeby pozostać wobec niej niewzruszona²¹⁰. Wreszcie, badania prowadzone w ramach finansów behawioralnych, sugerują, że relacja między ryzykiem, jakim obarczona jest decyzja, a samą decyzją, wygląda zgoła inaczej, niż zakładano w modelu racjonalnym. Okazuje się bowiem, że ludzie często preferują małe prawdopodobieństwo dużej wygranej nad oczekiwaną wartość całej perspektywy. Co więcej, poszukiwanie ryzyka można określić jako powszechne. Dzieje się tak na przykład w sytuacji, gdy ludzie dokonują wyboru między pewną stratą a dużym prawdopodobieństwem większej straty²¹¹. Z uwagi na powyższe, mechanizm decyzyjny proponowany w ramach finansów behawioralnych, wymaga bliższego wyjaśnienia.

2.3. Inwestor w świetle behawioralnej teorii finansów

Badacze zajmujący się analizą rynków z punktu widzenia efektywności informacyjnej wskazują, że wiele trudności nastrocza stwierdzenie o natychmiastowym reagowaniu rynku na pojawiającą się informację. Fama twierdzi bowiem, że na rynku efektywnym konkurencja pomiędzy inteligentnymi uczestnikami prowadzić będzie do sytuacji, w której w dowolnym momencie rzeczywiste ceny poszczególnych aktywów stanowić będą odzwierciedlenie „efektu informacji” opartego zarówno na wydarzeniach przeszłych, jak i antycypowanych. Innymi słowy, na rynku efektywnym w dowolnym momencie aktualna cena waloru będzie dobrym estymatorem wartości wewnętrznej²¹². Jeśli więc na rynku znajdują się inwestorzy, których Fama określa mianem „bystrych”, to w średnim ujęciu pełny efekt pojawienia się nowej informacji będzie widoczny w cenie waloru niemal natychmiast²¹³.

²⁰⁹ Żurek, M., *Inklinacje behawioralne na rynkach kapitałowych w świetle modeli SEM*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2016, s. 27.

²¹⁰ Gajdka, J., *Behawioralne finanse przedsiębiorstw...*, op.cit., s. 24.

²¹¹ Tversky, A., Kahneman, D., *Advances in Prospect-Theory...*, op.cit., s. 298.

²¹² Fama, E. F., *Random Walks in Stock Market Prices*, „Financial Analysts Journal”, 1965, nr 21/5, s. 56.

²¹³ Fama, E. F., *The Behaviour of Stock-Market...*, op.cit., s. 39.

W tym miejscu należy zaznaczyć, że potoczne rozumienie hipotezy efektywności rynku doprowadziło do jej wielu uproszczeń. Fama stwierdził przecież, że w rzeczywistości dochodzi do sytuacji, w której nowe informacje nie są jednoznaczne, tj. nie są precyzyjne i wiążą się z niepewnością. W związku z tym „natychmiastowe dostosowywania się cen” ma dwie implikacje. Po pierwsze, co zostało już wspomniane, nowa cena waloru może być początkowo przeszacowana lub niedoszacowana względem wartości wewnętrznej. Po drugie jednak, opóźnienie w dostosowywaniu się cen do kolejnych wartości wewnętrznych samo w sobie będzie niezależną zmienną losową, która może poprzedzać pojawiającą się informację (np. kiedy rynek przewiduje pojawienie się jakiejś informacji), a czasem następować może dopiero po pojawieniu się informacji. Tak czy inaczej, w efekcie tego procesu, następujące po sobie zmiany cen poszczególnych instrumentów powinny być niezależnymi zmiennymi o charakterze losowym²¹⁴.

Niemniej, nawet tak objaśniony problem natychmiastowości dostosowywania się zmian cen do informacji, nie jest wolny od niejasności. Wiele badań wskazuje na to, że inwestorzy dyskontują informacje w różnym tempie. Z jednej strony zależne bywa to od typu informacji, która jest przedmiotem analizy. Z drugiej natomiast, wpływ na tę sytuację mają czynniki, które wobec samej informacji mają charakter egzogeniczny, jak np. płynność rynku, jego transakcyjna efektywność, średni poziom obrotów akcjami poszczególnych spółek, pozycja emitenta, czy liczba transakcji, która zawierana jest podczas jednej sesji giełdowej²¹⁵. Powyższe przykłady mogą podać w wątpliwość założenie modelu błędzenia losowego, mówiącego o tym, że skoro informacja pojawia się na rynku w sposób losowy, to sama dynamika cen również będzie miała charakter losowy.

Przedstawiciele nurtu finansów behawioralnych stoją na stanowisku, że ludzki mózg często przetwarza informacje za pomocą skrótów i różnego rodzaju filtrów o charakterze poznawczym i emocjonalnym. Ich zdaniem takim samym mechanizmom podlegają inwestorzy rynkowi, którzy – pozornie zachowując się w sposób irracjonalny – naruszają koncepcje decyzyjności proponowane w ramach paradygmatu

²¹⁴ Ibidem, s. 39.

²¹⁵ Gurgul, H., *Analiza zdarzeń...*, op.cit., s. 17.

neoklasycznego²¹⁶. Jednym z punktów spornych, toczącej się w tym kontekście dyskusji, jest to, czy zachowania takie stanowią jedynie odstępstwo od normy, czy też mają charakter powszechny. Z jednej strony można za Thomasem Kuhnem stwierdzić, iż „anomalie ujawniają się tylko na gruncie paradygmatów. Im ściślejszy i więcej obejmujący jest paradygmat, tym bardziej czułym staje się on wskaźnikiem anomalii”²¹⁷. Sam fakt występowania tego typu zjawisk, nie musi przecież naruszać istnienia całego paradygmatu. Z drugiej jednak strony, istnieje bogaty dorobek badawczy, wskazujący na to, że nieprawidłowości w obszarach podejmowania decyzji inwestycyjnych, czy też szeroko pojętych decyzji finansowych, mają charakter powszechny. Co więcej – wpływają przez to na efektywność rynków oraz wyniki działalności przedsiębiorstw²¹⁸.

Warto jednak zaznaczyć, że oba paradygmaty wydają się różnić w sposób fundamentalny głównie z perspektywy teleologicznej. W literaturze z zakresu finansów behawioralnych często wskazuje się, że na przykład teoria Johna von Neumanna i Oskara Morgensterna ma charakter podejścia normatywnego. Koncepcje behawioralne z kolei, stawiają sobie za cel opisanie rzeczywistych sposobów postępowania inwestorów oraz wynikających z nich konsekwencji dla funkcjonowania rynków. Co za tym idzie, wpisane są w schemat deskryptywny²¹⁹.

Należy jednocześnie zauważyć, że współcześnie analizowane zagadnienia – w ramach paradygmatu behawioralnego – poprzedzone są wcześniejszymi teoriami, sugerującymi istnienie wad w neoklasycznym podejściu do rynków finansowych. Przykładowo w latach 50. XX wieku Herbert Simon postulował, że koncepcja *homo economicus* (przez co rozumiał człowieka racjonalnego) musi zostać zrewidowana²²⁰. Jego zdaniem, ze względu na psychologiczne ograniczenia, faktyczne ludzkie dążenie do racjonalności może być w najlepszym wypadku prostą aproksymacją tego, co powszechnie rozumie się przez to pojęcie²²¹. Rzeczywiste sytuacje decyzyjne wiążą

²¹⁶ Baker, H. K., Nofsinger, J. R., *Behavioral finance: Investors, corporations, and markets*, John Wiley & Sons, New Jersey, 2010, s. 3.

²¹⁷ Kuhn, T. S., *Struktura rewolucji naukowych*, Fundacja Aletheia, Warszawa 2001, s...

²¹⁸ Baker, H. K.; Nofsinger, J. R., *Behavioral Finance...*, op.cit. s. 3.

²¹⁹ Szyszka, A., *Behavioural Finance and Capital Markets: How Psychology Influences Investors and Corporations*, Palgrave Macmillan, Nowy Jork 2013, s. 46.

²²⁰ Simon, H., *A Behavioral Model of Rational Choice*, „The Quarterly Journal of Economics”, 1955, nr 69/1, s. 99–118.

²²¹ Ibidem, s. 99–118.

się bowiem z tak dużą ilością – koniecznych do zdobycia oraz przetworzenia – informacji, że wskazanie tak zwanej optymalnej alternatywy, jest właściwie niemożliwe²²². Podobnie, w 1978 roku Harvey Leibenstein zaproponował teorię selektywnej racjonalności (ang. *selective rationality*), w ramach której zwrócił uwagę na to, że racjonalność podlega stopniowaniu, a poszczególne jednostki mają w pewnym sensie wybór, jak „racjonalnie” mają się zachowywać w różnych kontekstach. Zakres, w jakim będą one ową racjonalność przejawiać, może być zależny od ich osobowości²²³. W 2002 roku Nagrodę Banku Szwecji im. Alfreda Nobla w dziedzinie ekonomii „za zintegrowanie spostrzeżeń z badań psychologicznych z naukami ekonomicznymi, zwłaszcza dotyczącymi ludzkich osądów i podejmowania decyzji w warunkach niepewności”²²⁴ otrzymał Daniel Kahneman – jeden z czołowych przedstawicieli nurtu behawioralnego. Zaproponowana przez niego oraz Amosa Tverskiego teoria perspektywy została jednak stworzona w latach 70. Piętnaście lat później wyróżnienie to – za wkład w ekonomię behawioralną – otrzymał Richard Thaler, ale jedna z jego kluczowych prac *Judgment and Decision Making under Uncertainty: What Economists Can Learn from Psychology* opublikowana została w roku 1980²²⁵. Często spotykane w literaturze określanie finansów behawioralnych mianem „nowych” nie jest więc w pełni zgodne z rzeczywistością. Można wręcz stwierdzić, że – w pewnym okresie – oba paradygmaty rozwijały się symultanicznie, dochodząc jednak do rozbieżnych wniosków.

Hersh Shefrin wskazuje na dwie kluczowe różnice między podejściem behawioralnym a neoklasycznym, które plasują się w założeniach związanych z racjonalnością oraz teorią oczekiwanej użyteczności. W pierwszym przypadku istotą problemu wydaje się to, co Shefrin nazywa „sentymentem”. Zwolennicy finansów behawioralnych traktują nastroje jako główną determinantę cen rynkowych, będących wynikiem systematycznych błędów popełnianych przez inwestorów. Z kolei w ramach szkoły neoklasycznej zakłada się, że sentyment jest kwestią nieistotną, a inwestorzy,

²²² Żurek, M., *Inklinacje behawioralne...*, op.cit., s. 25.

²²³ Leibenstein, H., *General X-efficiency Theory and Economic Development*, Oxford University Press, Oxford 1978, s. 21.

²²⁴ Nobelprize.org: The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2002, <https://www.nobelprize.org/prizes/economic-sciences/2002/summary>. Dostęp: 7 mar. 2023.

²²⁵ Zob. Thaler, R., *Judgment and Decision Making under Uncertainty: What Economists Can Learn from Psychology*, „Risk Analysis in Agriculture Research and Educational Developments”, January 16–18, 1980, Tucson, Arizona.

co do zasady, wolni są od uprzedzeń związanych z przetwarzaniem informacji. Co za tym idzie, podczas gdy badacze nurtu behawioralnego przypisują zjawiska związane z dynamiką cen czynnikom psychologicznym, neoklasyści tłumaczą je poprzez problem fundamentalnego ryzyka lub – zmiennej w czasie – awersji do ryzyka. Drugą cechą odróżniającą podejście behawioralne od tradycyjnego, jest założenie oczekiwanej użyteczności. Neoklasyczni teoretycy wyceny aktywów zakładają, że inwestorzy starają się zmaksymalizować oczekiwaną użyteczność, a przesłanką wspierającą to założenie jest racjonalność inwestorów. Przedstawiciele finansów behawioralnych są jednak krytyczni wobec definiowania koncepcji oczekiwanej użyteczności jako teorii opisowej. Zamiast tego proponują, aby na tego typu analizy narzucić optykę wiedzy z zakresu psychologii.

2.3.1. Teoria perspektywy

Jednym z poważniejszych zarzutów, jakie stawiane są finansom behawioralnym, jest ten, że nie stanowią one holistycznej teorii, która mogłaby zastąpić hipotezę efektywności informacyjnej²²⁶. Jest to spowodowane między innymi wielością zagadnień, które analizowane są z perspektywy behawioralnej, gdzie z jednej strony punktem wyjścia jest inwestor i jego indywidualne zachowanie, a z drugiej rynki jako całość. Niewątpliwie jednak w tym kontekście jednym z przełomowych momentów, było opublikowanie w 1979 roku przez Daniela Kahnemana i Amosa Tversky'ego artykułu *Prospect Theory: An Analysis of Decision under Risk*. W ramach swojej wieloletniej współpracy autorzy ci wskazywali, że rzeczywiste preferencje decyzyjne inwestorów naruszają aksjomaty teorii oczekiwanej użyteczności w sposób systematyczny. W związku z tym koncepcja ta nie powinna być rozpatrywana w kategorii modelu adekwatnie opisującego rzeczywistość. W zamian zaproponowali oni alternatywny model, który znany jest pod nazwą teorii perspektywy²²⁷ (ang. *prospect theory*). Aby opisać mechanizm podejmowania decyzji przez inwestorów, Kahneman i Tversky używają określenia „ramy decyzyjnej” (ang. *decision frame*). Hersh Shefrin charakteryzuje to zagadnienie w następujący sposób:

²²⁶ Fama, E. F., *Market efficiency, long-term returns, and behavioral finance*, „Journal of Financial Economics”, 1998, nr 49, s. 284.

²²⁷ Kahneman, D., Tversky, A., *Prospect Theory: An Analysis of Decision under Risk*, „The Econometric Society”, 1979, nr 47/2, s. 263.

Forma używana do opisu problemu decyzyjnego nazywana jest jego ramą. Kiedy mówię o niezależności ramy, mam na myśli to, że forma nie ma znaczenia dla zachowania. Zwolennicy tradycyjnych finansów zakładają, że ramy są przejrzyste. Oznacza to, że praktycy mogą przejrzeć wszystkie różne sposoby opisu przepływów pieniężnych. Jednak wiele ramek nie jest przezroczystych, a raczej są one nieprzezroczyste. Kiedy dana osoba ma trudności z widzeniem przez nieprzezroczystą ramę, jej decyzje zazwyczaj zależą od konkretnej ramy, której używa. Zatem różnica w formie jest jednocześnie różnicą w istocie. Zachowanie odzwierciedla zależność od ramki²²⁸.

Zgodnie z tą ideą, sposób postrzegania sytuacji, jaki przyjmie inwestor, jest częściowo zależny od tego, jak sformułowany zostanie problem decyzyjny, a częściowo od czynników wewnętrznych, takich jak jego: normy, nawyki czy cechy osobiste²²⁹.

Teoria perspektywy składa się z dwóch części. Pierwszą jest funkcja wartości, drugą natomiast funkcja wag decyzyjnych. W pierwszym przypadku cechą różnicującą propozycję Kahnemana i Tversky'ego od teorii neoklasycznej jest to, że podczas gdy teoria oczekiwanej użyteczności zakłada, że nośnikiem wartości jest ostateczny poziom bogactwa, teoria perspektywy wskazuje, że taką rolę odgrywają zyski i straty oszacowywane względem punktu odniesienia. W teorii perspektywy wyniki są wyrażane jako dodatnie lub ujemne odchylenia (zyski lub straty) od neutralnego punktu odniesienia, któremu przypisuje się wartość zerową²³⁰. Istotne jest jednak to, że ów referencyjny punkt nie jest stały. Jeśli nastąpi jego przesunięcie (a inwestorzy pogodzą się np. ze stratą), zostanie on sformułowany ponownie. Jeśli wróci on do poprzedniego poziomu, inwestor będzie traktował tę sytuację w kategoriach zysku²³¹. To, jaką perspektywę (zysków lub strat) przyjmie inwestor podczas podejmowania decyzji, zależne jest więc od umiejscowienia punktu referencyjnego²³².

Ta logika stoi w opozycji do normatywnego aksjomatu porównywalności. Zgodnie z nią, inwestorzy mogą bowiem wykazywać różne preferencje, w zależności od

²²⁸ Shefrin, H., *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing*, Oxford University Press, Oxford 2002, s. 23.

²²⁹ Tversky, A., Kahneman, D., *The Framing of Decisions and the Psychology of Choice*, „Science”, 1981, nr 211, s. 453.

²³⁰ Tversky, A., Kahneman, D., *The Framing of Decisions* ..., s. 454.

²³¹ Por. Zielonka, P., *Gięda i psychologia*..., op.cit., s. 79.

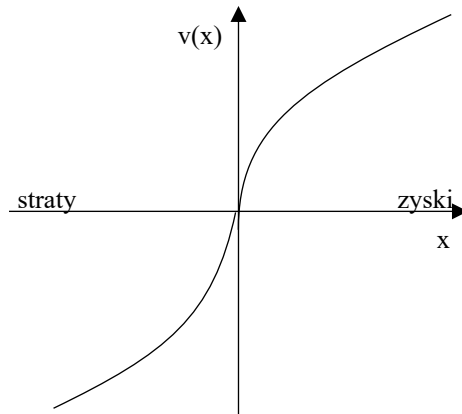
²³² Kahneman, D., Tversky, A., *Prospect Theory*..., op.cit., s. 274.

sposobu, w jaki informacja zostanie im zaprezentowana²³³. Podawana jest więc w wątpliwość ich zdolność do poprawnego interpretowania komunikatów. Nawet jeśli powtarzające się problemy decyzyjne – z obiektywnego punktu widzenia – są takie same, to ze względu na ograniczenia związane z postrzeganiem informacji, mogą być zinterpretowane jako różne, z uwagi na odmienny sposób ich sformułowania. Obserwacja ta potocznie nazywana jest mianem choroby azjatyckiej. Rozumie się przez to sytuację, w której wybór alternatyw uzależniony jest od tego, czy problem decyzyjny przedstawiony został w kontekście zysków, czy też strat²³⁴.

Bardziej precyzyjny opis powyższych mechanizmów charakteryzują wspomniane funkcje: wartości oraz wag decyzyjnych. Pierwsza z nich przyjmuje trzy właściwości:

- (1) określana jest na podstawie odchylenia od punktu referencyjnego;
- (2) w ogólnym ujęciu jest wklęsła dla zysków i wypukła dla strat;
- (3) jest bardziej stroma dla strat niż dla zysków²³⁵.

Rysunek 3. Krzywa wartości dla zysków i strat teorii perspektywy. Teoria perspektywy



Źródło: Kahneman, D., Tversky, A., *Prospect Theory: An Analysis of Decision under Risk*, „The Econometric Society”, 1979, nr 47/2, s. 279.

Pierwszy punkt wskazuje na to, że postrzegana przez inwestora wartość nie jest definiowana przez absolutną wartość bogactwa, a przez zmiany w jego poziomie –

²³³ Szyszka, A., *Behavioural Finance...*, op.cit., s. 10.

²³⁴ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 77.

²³⁵ Kahneman, D., Tversky, A., *Prospect Theory...*, op.cit., s. 279.

zarówno po stronie zysków, jak i strat²³⁶. Ściślej mówiąc, wartość powinna być traktowana jako funkcja o dwóch argumentach: pozycji aktywów, która służy jako punkt odniesienia, oraz wielkości zmiany (dodatniej lub ujemnej) od tego punktu²³⁷. Druga i trzecia charakterystyka funkcji wartości związana jest ze stosunkiem inwestorów do ryzyka. W celu objaśnienia fenomenu, do którego w tym kontekście dotarli Kahneman i Tversky, należy przytoczyć zaproponowany przez nich wzór funkcji²³⁸:

$$v(x) = \begin{cases} x^\alpha & \text{jeśli } x \geq 0 \\ -\lambda(-x)^\beta & \text{jeśli } x < 0 \end{cases}$$

$v(x)$ – funkcja wartości; x – zyski lub straty; α, β – parametry funkcji $\alpha, \beta \in (0; 1)$; λ – współczynnik awersji do strat.

W sytuacji, w której wartość definiowana jest w kategoriach ujemnych, pojawia się współczynnik awersji do strat. Środek ciężkości teorii perspektywy położony jest więc nie tyle na awersję do ryzyka, tak jak przyjmuje się w przypadku założenia racjonalności, ile właśnie na awersję do straty. Kahneman i Tversky opisując tę prawidłowość, użyli słynnego sformułowania „strata boli bardziej, niż cieszy zysk”²³⁹. Fundamentalnym elementem, pozwalającym wyjaśnić reakcje inwestorów, jest zatem ich niechęć do ponoszenia strat, ponieważ odczuwają je dotkliwiej niż radość, wynikającą z ewentualnych zysków. Tak zdefiniowana awersja opisuje stosunek do podejmowania ryzyka w sytuacjach, w których inwestorzy mogliby doświadczyć zysków, gdyby dopisywało im szczęście, ale jednocześnie mogliby ponieść straty, jeżeli potencjalnie wynik byłby niekorzystny²⁴⁰. Determinantą tego, czy inwestorzy wykazywać będą skłonność, czy awersję do ryzyka, jest więc ich perspektywa, tj. ujęcie problemu w kategoriach zysków lub strat²⁴¹. W ten sposób objaśnić można pierwszą część teorii perspektywy.

Część druga związana jest z problemem prawdopodobieństwa i przypisywanym mu wagom decyzyjnym. O ile teoria oczekiwanej użyteczności zakłada, że inwestorzy prawidłowo przypisują prawdopodobieństwa, teoria perspektywy wskazuje, że mają oni

²³⁶ Zielonka, P., *Gięda i psychologia...*, op.cit., s. 79.

²³⁷ Kahneman, D., Tversky, A., *Prospect Theory...*, op.cit., s. 277.

²³⁸ Tversky, A., Kahneman, D., *Advances in Prospect-Theory...*, op.cit., s. 309.

²³⁹ Kahneman, D., Tversky, A., *Prospect Theory...*, op.cit., s. 288.

²⁴⁰ Shefrin, H., *The behavioural Paradigm Shift*, „RAE-Revista de Administração de Empresas”, 2015, nr 55/1, s. 96.

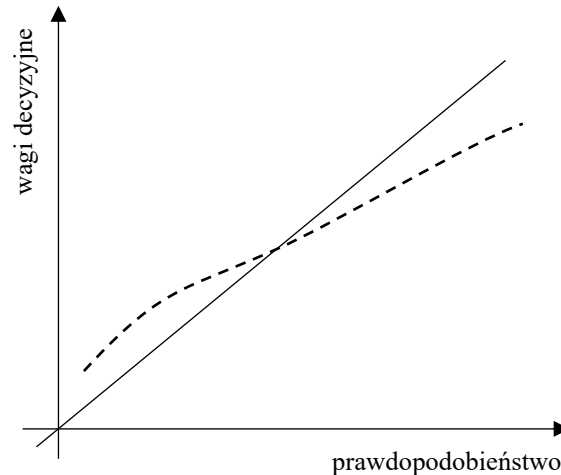
²⁴¹ Por. Zielonka, P., *Gięda i psychologia...*, op.cit., s. 78.

tendencję do przypisywania niektórym z nich nadmiernych wag, z kolei innym – zbyt małych²⁴². Chociaż wartość każdego wyniku jest mnożona przez wagę decyzji – podobnie jak w przypadku koncepcji neoklasycznej – to wagi te wywiedzione są z wyborów między perspektywami²⁴³. Zagadnienie to zdefiniować można za Kahnemanem w następujący sposób:

mając do czynienia ze zdarzeniem o niepewnym wyniku, w swojej ocenie nadajesz możliwym rezultatom różną wagę. Wagi mają oczywisty związek z prawdopodobieństwem rezultatów: szansa 50 procent na wygranie miliona jest atrakcyjniejsza niż szansa 1 procenta na wygranie takiej samej kwoty²⁴⁴

W ogólnym ujęciu, na podstawie drugiej części teorii perspektywy wysnuć można dwa wnioski. Po pierwsze inwestorzy w procesie decyzyjnym nie kierują się matematycznym poziomem prawdopodobieństwa²⁴⁵. Po drugie, mają oni tendencję do zawyżania małych prawdopodobieństw i niedoszacowywania wag prawdopodobieństw średnich i wysokich²⁴⁶. Dla niskich poziomów prawdopodobieństw funkcja wag przyjmie bowiem wartości wyższe niż obiektywne prawdopodobieństwo i odwrotnie²⁴⁷.

Rysunek 4. *Funkcja wag decyzyjnych*



Źródło: Tversky, A., Kahneman, D., *Advances in Prospect-Theory – Cumulative Representation of Uncertainty*, „Journal of Risk and Uncertainty”, 1992, nr 5/4, s. 313.

²⁴² Shefrin, H., *A Behavioral Approach to Asset Pricing*, Elsevier Inc, Burlington 2008, s. 392.

²⁴³ Kahneman, D., Tversky, A., *Prospect Theory...*, op.cit., s. 280. Należy zaznaczyć, że wagi decyzyjne nie są tożsame z prawdopodobieństwem: nie są one zgodne z aksjomatami prawdopodobieństwa i nie powinny być interpretowane jako mierniki lub stopnie przekonań.

²⁴⁴ Kahneman, D., *Pułapki myślenia. O myśleniu szybkim i wolnym*, Media Rodzina, Poznań 2011, s. 317.

²⁴⁵ Żurek, M., *Inklinacje behawioralne...*, op.cit., s. 33.

²⁴⁶ Tversky, A., Kahneman, D., *Advances in Prospect-Theory...*, op.cit., s. 312.

²⁴⁷ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 81.

Wzorce zachowań towarzyszące perspektywom obciążonym ryzykiem obrazuje krzywa wag decyzyjnych. Wynika z niej kilka prawidłowości:

- (1) inwestorzy są niewrażliwi na różnice prawdopodobieństw ze środka skali²⁴⁸;
- (2) awersja do ryzyka w sytuacji wyboru między zyskiem pewnym i prawdopodobnym wynikać może z przypisywania relatywnie niskich wag wysokim prawdopodobieństwom;
- (3) ten sam mechanizm tłumaczyć może skłonność do podejmowania ryzyka w sytuacji wyboru między pewną a prawdopodobną stratą²⁴⁹;
- (4) funkcja wag decyzyjnych dla zysków i strat kształtuje się analogicznie, jednak dla zysków jest ona bardziej zakrzywiona. W efekcie awersja do ryzyka po stronie zysków jest wyraźniejsza niż skłonność do podejmowania ryzyka po stronie strat dla umiarkowanych i wysokich prawdopodobieństw²⁵⁰.

Na podstawie powyższych rozważań zarysować można ogólny stosunek do ryzyka inwestorów, który zaprezentowany jest w tabeli 1.

Tabela 1. *Stosunek do ryzyka. Teoria perspektywy.*

	Zysk	Strata
Średnie i wysokie wartości prawdopodobieństwa	Awersja do ryzyka	Skłonność do ryzyka
Bardzo niskie wartości prawdopodobieństwa	Skłonność do ryzyka	Awersja do ryzyka

Zródło: Zielonka, P., Gielda i psychologia. Behawioralne aspekty inwestowania na rynku papierów wartościowych, CeDeWu, Warszawa 2017, s. 87

Odwołując się do znanych modeli neoklasycznych, na podstawie pierwszej i drugiej części teorii perspektywy zaproponowano model podejmowania decyzji określany maksymalizacją subiektywnej oczekiwanej wartości²⁵¹ (ang. *subjective expected utility*):

$$ev = \sum_i w(p_i)v(x_i), \text{ gdzie:}$$

ev – subiektywna oczekiwana wartość; $w(p_i)$ – wagi decyzyjne przypisane poszczególnym wartościom p ; $v(x_i)$ – funkcja wartości.

²⁴⁸ Tversky, A., Kahneman, D., *Advances in Prospect-Theory...*, op.cit., s. 312.

²⁴⁹ Ibidem, s. 316.

²⁵⁰ Ibidem, s. 312.

²⁵¹ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 85.

Kahneman i Tversky zauważyli, że odstępstwa od wyceniania stopy zwrotu z aktywów zgodnie z modelem oczekiwanej użyteczności, mają charakter powszechny, a nie jak wcześniej sądzono jednostkowy. To spostrzeżenie zaprowadziło ich z kolei do podania w wątpliwości – opisanych przez Neumanna i Morgensterna – aksjomatów racjonalności. Można jednak zauważyć, że w ramach teorii perspektywy zachowana została ogólna postać dwuliniowa, która leży u podstaw teorii oczekiwanej użyteczności. Mimo to, modele te u swoich podstaw wykazują zasadnicze różnice. Po pierwsze, w perspektywie behawioralnej wartości przypisywane są nie stanom końcowym, jak w przypadku ujęcia klasycznego, ale zmianom, jakie w tych stanach zachodzą. Po drugie model szacowania prawdopodobieństwa, chociaż z normatywnego punktu widzenia użyteczny, to – nie będąc tożsamym z wagami decyzyjnymi – nie opisuje rzeczywistego procesu decyzyjnego. Co do zasady jednak przyjąć można, że odstępstwa od racjonalnych preferencji mogą być przez inwestorów korygowane. Warunek tego procesu jest jednak taki, że muszą oni zdawać sobie z owych odstępstw sprawę. Jak twierdzą Kahneman i Tversky:

w wielu sytuacjach decydent nie ma możliwości odkrycia, że jego preferencje mogłyby naruszyć reguły decyzyjne, których pragnie przestrzegać. W tych okolicznościach oczekuje się, że wystąpią anomalie implikowane przez teorię perspektywy²⁵².

2.3.2. Inklinacje behawioralne inwestorów giełdowych

W ramach literatury przedmiotu z zakresu finansów powszechne jest klasyfikowanie inwestorów na: instytucjonalnych i indywidualnych lub też profesjonalnych i nieprofesjonalnych. Oczywiście tego typu podziały zawsze stanowią pewne uproszczenie. Sami inwestorzy instytucjonalni to grupa bardzo niejednorodna. W jej skład wpisuje się z jednej strony animatorów rynku, a z drugiej podmioty takie jak: banki, samorządy, spółki, skarb państwa czy firmy maklerskie. O ile jednak sposób klasyfikacji inwestorów nie stanowi meritum niniejszej dysertacji, to umiejscowienie go w kontekście analizowanych paradygmatów, każe zadać pytanie o powszechność ich implikacji.

²⁵² Kahneman, D., Tversky, A., *Prospect Theory...*, op.cit., s. 277.

Systematyzacja behawioralnych inklinacji inwestorów finansowych dzieli je z kolei na dwie obszerne grupy. Pierwsza związana jest ze sferą poznawczą inwestorów. Można tu wymienić efekty: nadmiernej pewności siebie, złudzenia kontroli, pułapki konfirmacji, czy też błędu atrybucji. Ponadto, w grupie tej znajdują się zjawiska określane mianem heurystyk, które Kahneman definiuje jako „proste procedury, które mogą udzielać adekwatnych, choć niedoskonałych odpowiedzi na trudne pytania”²⁵³. Grupa ta cechuje się niejednorodnością. W literaturze przedmiotu podawane są takie przykłady heurystyk inwestorów jak: błąd retrospekcji, heurystyka afektu, czy heurystyka reprezentatywności. Druga grupa związana jest ze skłonnościami motywacyjnymi inwestorów. Z jednej strony mieści się tutaj – opisywana wcześniej – teoria perspektywy, z drugiej natomiast jest to szereg inklinacji, takich jak: mentalne księgowanie, efekt utopionych kosztów, efekt dyspozycji, czy efekt *status quo*.

Charakterystyczne dla inwestorów instytucjonalnych jest to, że dysponują dwiema siłami, z którymi trudno mierzyć się mniejszym graczom, to jest nieprzeciętnym kapitałem oraz doświadczeniem analityków, którzy pracują na ich rzecz²⁵⁴. Ponadto, portfele, którymi zarządzają i generowane przez nich obroty, mogą mieć wpływ na dynamikę całego rynku²⁵⁵. Przytaczane w pierwszym rozdziale teorie wskazują, że to właśnie inwestorzy instytucjonalni mogą być uznani za „nadawców” cen. Prowadzone w tej tematyce badania, wskazują również, że inwestorzy instytucjonalni, podczas podejmowania decyzji, najczęściej kierują się wiadomościami o charakterze fundamentalnym, natomiast inwestorzy indywidualni – trendem cenowym²⁵⁶. Można by więc przypuszczać, że takie podmioty są mniej narażone na konsekwencje nieracjonalności czy ponoszenia – opisywanych wcześniej – kosztów związanych z pozyskiwaniem informacji. W istocie prowadzone w tej tematyce badania niejednokrotnie wykazywały podatność na różnego typu inklinacje behawioralne inwestorów indywidualnych, nie wskazując analogicznych mechanizmów wśród reprezentantów dużych podmiotów lub wskazując ich występowanie w znacznie

²⁵³ Kahneman, D., *Pułapki myślenia. O myśleniu szybkim i wolnym*, Media Rodzina, Poznań 2011, s. 9.

²⁵⁴ Banaszczyk-Soroka, U., *Rynek papierów wartościowych. Inwestorzy, instrumenty finansowe i metody wyceny*, C.H. Beck, Warszawa 2016, s. 50–52.

²⁵⁵ Szyszka, A., *Behavioural Finance...*, op.cit., s. 82.

²⁵⁶ Zob. Chen, J., *Information Theory and Market Behavior*, „CFAI Journal of Behavioral Finance”, 2005, nr 2/4, s. 25–45, 4 [za: Cohen, Gompers and Vuolteenaho: 2002].

mniejszym stopniu. Przykładem takiego stanu rzeczy może być efekt dyspozycji (ang. *disposition effect*), czyli tendencji do utrzymywania w portfelu papierów wartościowych, których ceny – od czasu zakupu – spadły, a pozbywanie się tych, które generują zysk²⁵⁷. Efekt ten potwierdzony został podczas badań przeprowadzanych na inwestorach indywidualnych²⁵⁸. Badania prowadzone na rynku amerykańskim pokazały jednak, że osoby zarządzające funduszami emerytalnymi (próba wynosiła 769) pod koniec każdego kwartału zachowywały w portfelu akcje przynoszące zyski, a sprzedawały te, których stopy zwrotu były ujemne²⁵⁹. Z kolei badania prowadzone w tym zakresie na rynku chińskim pokazały wprawdzie tendencję inwestorów instytucjonalnych do utrzymywania w portfelach „spółek przegranych”, ale nie do sprzedawania „spółek wygranych”²⁶⁰. Stwierdzenie więc, że inwestorzy instytucjonalni podlegają takim samym czynnikom psychologicznym, co inwestorzy indywidualni lub nieprofesjonalni, byłoby nadużyciem.

Z drugiej jednak strony, w literaturze przedmiotu z zakresu finansów behawioralnych wskazuje się, że pewne inklinacje psychologiczne dotyczą zarówno inwestorów indywidualnych, jak i instytucjonalnych. Co więcej, niektóre uznać można za wręcz charakterystyczne dla tej drugiej grupy. Przykładem jest tutaj mechanizm dynamicznej awersji do straty (ang. *dynamic loss aversion*), polegający na zmianie nastawienia do ryzyka w zależności od wyniku uzyskanego w poprzednim okresie oraz horyzontu czasowego. W praktyce oznacza to, że inwestorzy skłonni są do podejmowania większego ryzyka, po wcześniej uzyskanych dodatnich stopach zwrotu. Dzieje się tak, ponieważ – z punktu widzenia osoby zarządzającej portfelem – wcześniejsze zyski zamortyzują skutki ewentualnych strat w przyszłości, ograniczając tym samym ryzyko zamknięcia okresu w niekorzystny sposób. Dodatkowo wzrost podejmowanego ryzyka zmniejsza się pod koniec każdego kwartału, czyli w momencie, w którym menedżer

²⁵⁷ Zob. Shefrin, H., Statman, M., *The disposition to sell winners too early and ride losers too long. Theory and evidence*, „The Journal of Finance”, 1985, nr XL/3, s. 777–790.

²⁵⁸ Zob. Ibidem, s. 777–790.

²⁵⁹ Zob. Lakonishok, J., Shleifer, A., Thaler, R., Vishny, R., *Window Dressing by Pension Fund Managers*, „Behavioral Finance”, 1991, nr 81/2, s. 227–231.

²⁶⁰ Zob. Chen, G., Kim, K. A., Nofsinger, J. R., Rui, O. M., *Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors*, „Journal of Behavioral Decision Making”, 2007, nr 20, s. 425–451.

mógłby być oceniany, bądź jego wyniki mogłyby zostać porównane z wynikami uzyskanymi przez innych zarządzających portfelami²⁶¹.

Innym zagadnieniem, analizowanym przez badaczy, jest nadmierna pewność siebie (ang. *overconfidence*). Zjawisko to polega na przecenianiu przez inwestorów własnej wiedzy i umiejętności, które skutkuje nieuzasadnionym przekonaniem o trafności podejmowanych decyzji, czy wydawanych sądów²⁶². Według Adama Szyszki opisywany mechanizm przejawiać może się na kilka sposobów. Pierwszym jest efekt kalibracji (ang. *calibration effect*), opierający się na przesadnym przekonaniu o precyzji swojej wiedzy i trafności prognoz. Drugim sposobem jest zjawisko wybiórczej atrybucji (ang. *self-attribution error*), które polega na tendencji do przypisywania przyczyn sukcesu własnym kompetencjom i cechom osobowościowym, przy jednoczesnym tłumaczeniu przyczyn niepowodzeń niezależnymi czynnikami zewnętrznymi²⁶³. Podobnym zjawiskiem jest tak zwany podstawowy błąd atrybucji (ang. *fundamental attribution error*), przez który ludzie – podczas wyjaśniania zachowań innych ludzi – przypisują większe znaczenia czynnikom osobowościowym niż środowiskowym²⁶⁴. Trzecim zjawiskiem jest złudzenie kontroli (ang. *illusion of control*), przejawiające się w nieuzasadnionym przeświadczeniu o możliwości wpływania na przebieg procesów, które w istocie są od nas niezależne²⁶⁵. Z mechanizmem nadmiernej pewności siebie wiąże się również nieuzasadniony optymizm (ang. *excessive optimism*), czasem określany mianem „myślenia życzeniowego”. Jest to przekonanie, że dotyczący podmiotu przebieg zdarzeń, będzie pomyślny²⁶⁶. W kontekście nadmiernej pewności siebie często wspomina się również o błędzie retrospekcji (ang. *hindsight bias*). Błąd ten nazywany jest również efektem myślenia wstecznego i polega na tendencji do przypisywania większego prawdopodobieństwa wynikom zdarzeń, które już miały miejsce.

²⁶¹ Zob. Teo, M., O'Connell, P. G. J., *Prospect Theory and Institutional Investors*, „SSRN Electronic Journal”, 2005, nr. , s. 1–39. Oraz O'Connell, P. G. J., Teo, M., *Institutional Investors, Past Performance, and Dynamic Loss Aversion*, „The Journal of Financial and Quantitative Analysis”, 2009, nr 44/1, s. 155–188.

²⁶² Zielonka, P., *Gielda i psychologia...*, op.cit., s. 45.

²⁶³ Szyszka, A., *Finanse behawioralne...*, op.cit., s. 48–50.

²⁶⁴ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 66.

²⁶⁵ Ibidem, s. 48.

²⁶⁶ Ibidem, s. 50.

Tabela 2. *Wybrane błędy poznawcze i motywacyjne inwestorów giełdowych.*

Typ błędu	Objaśnienie	Potencjalne implikacje
Błąd confirmacji	Poszukiwanie w informacji dowodów, które potwierdziłyby słuszność wcześniej przyjętych założeń i hipotez oraz zaprzeczanie dowodom, które są z nimi niezgodne.	Zignorowanie sygnałów wskazujących na konieczność zmiany decyzji inwestycyjnej.
Błąd retrospekcji	Przeświadczenie, że można było przewidzieć bieg wypadków, który już miał miejsce. Przypisywanie większego prawdopodobieństwa znanym już wynikom zdarzeń.	Przekonanie o możliwości dokonania prawidłowej prognozy zdarzeń, których prawdopodobieństwo wynosi mniej niż 100%.
Heurystyka afektu	Sytuacja, w której pogląd kształtowany jest przez wcześniejsze upodobanie lub niechęć. Wpływ stanu emocjonalnego na ocenę zjawiska.	Pozytywne nastawienie zmniejszy szacowany poziom prawdopodobieństwa wystąpienia niekorzystnego zdarzenia. Negatywne nastawienie zwiększy szacowany poziom prawdopodobieństwa wystąpienia niekorzystnego zdarzenia.
Heurystyka reprezentatywności	Ocenianie prawdopodobieństwa na podstawie stopnia podobieństwa zdarzeń lub na podstawie relatywnie niewielkiej próby. Poszukiwanie znanych wzorców w celu rozwiązania nowych problemów.	Błędnie przypisywanie np. pozytywnych cech firmy lub jej menedżerów jako cechy inwestycji, będącej przedmiotem analizy.
Heurystyka zakotwiczenia	Sposób myślenia, w którym ocena jakiejś sytuacji wywodzona jest na podstawie wcześniejszych, mentalnie dostępnych w danej chwili, zdarzeń.	W zastosowaniach inwestycyjnych na prognozy inwestorów dotyczące przyszłej wartości może mieć wpływ odniesienie do pewnego wzorca lub wcześniejszych wyników.
Złudzenie kontroli	Poszukiwanie podczas oceniania sytuacji związków przyczynowo-skutkowych umiejscowionych we własnych kompetencjach, przy jednoczesnym ignorowaniu roli przypadku.	Przypisanie zbyt wysokich prawdopodobieństw prognozom korzystnym z punktu widzenia inwestora.

Źródło: opracowanie własne²⁶⁷.

²⁶⁷ Opracowanie własne na podstawie: Ahmed Z., Ibrahim H., Tuyon J., *Institutional investor behavioral biases: syntheses of theory and evidence*, „Management Research Review”, nr 40 (5), 2017; Griffin D., Tversky A., *The weighing of evidence and the determinants of confidence*, „Cognitive Psychology”, nr 24 (3), 1992; Kahneman D., *Pułapki myślenia. O myśleniu szybkim i wolnym*, Media Rodzina, Poznań 2012; Tversky A., Kahneman D., *Judgment under Uncertainty: Heuristics and Biases. Biases in judgments reveal some heuristics of thinking under uncertainty*, „Science”, nr 185 (4157), 1974; Zielonka P., *Giełda i psychologia. Behawioralne aspekty inwestowania na rynku papierów wartościowych*, CeDeWu, Warszawa 2017.

Przejawem takiego myślenia jest powszechnie używane twierdzenie: „to było do przewidzenia”²⁶⁸. Wielowątkowość nadmiernej pewności siebie jest jednym z powodów, przez który przypuszczalnie występuje ona powszechnie, zwłaszcza wśród inwestorów.

Zdaniem Jerzego Gajdki, przesadna wiara we własne możliwości jest właściwa zarówno graczom nieprofesjonalnym, jak i profesjonalnym, chociaż w tym drugim przypadku objawia się ona przede wszystkim w nadmiernym przekonaniu co do kompetencji związanych ze sporządzaniem prognoz, a cechę tą posiadają nie tylko inwestorzy, ale również znaczna część menedżerów²⁶⁹. Prowadzone na istniejących rynkach badania wskazują, że inwestorzy instytucjonalni i profesjonalni – wprawdzie w mniejszym stopniu niż indywidualni i nieprofesjonalni – przejawiają skłonność do nadmiernej pewności siebie. Przykładem takich badań są przeprowadzone przez Gongmenga Chena, Hennetha Kima, Johna Nofsingera i Oliviera Ruia w 2007 roku analizy na rynku chińskim, w ramach, których doszli do wniosku, że omawiany mechanizm wśród inwestorów instytucjonalnych (nazywanych przez nich wyrafinowanymi) przejawia się niedostateczną dywersyfikacją portfeli i zbyt częstym podejmowaniem decyzji handlowych²⁷⁰. W 2011 roku analogiczne badania przeprowadzono na rynku tajwańskim. W obu przypadkach wykazano, że wyższym poziomem nadmiernej pewności siebie cechują się inwestorzy indywidualni, ale nie wykazano, że inwestorzy instytucjonalni są tej cechy pozbawieni²⁷¹. Ponadto Robin Chou i Wang Yun-Yi zwrócili uwagę, że inwestorzy profesjonalni po odnotowaniu zysków, handlują w sposób bardziej agresywny niż wcześniej, co zgodne jest z hipotezą nadmiernej pewności siebie²⁷². Jednocześnie jest to również potwierdzenie mechanizmu dynamicznej awersji do straty.

W tym kontekście ciekawe wydają się wyniki, które Idil Ozlem Koc uzyskał podczas badania rynku tureckiego. Badanie to polegało nie tyle na analizie wyników

²⁶⁸ Ibidem, s. 55.

²⁶⁹ Gajdka, J., *Behawioralne finanse przedsiębiorstw...*, op.cit., s. 37–38.

²⁷⁰ Zob. Chen, G., Kim, K. A., Nofsinger, J. R., Rui, O. M., *Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors*, „Journal of Behavioural Decision Making”, 2007, nr 20, s. 425–451.

²⁷¹ Zob. Chuang, W.-I., Susmel, R., *Who Is the More Overconfident Trader? Individual vs. Institutional Investors*, „Journal of Banking and Finance”, 2011, nr 35/7, s. 1626–1644.

²⁷² Zob. Chou, R.K.; Wang, Y.-Y., *A Test of the Different Implications of the Overconfidence and Disposition Hypotheses*, „Journal of Banking and Finance”, 2011, nr 35/8, s. 2037–2046.

funduszy, ile analizie indywidualnych decyzji inwestycyjnych podejmowanych przez osoby zarządzające funduszami. O ile bowiem nadmierna pewność siebie objawiać może się przez zbyt wysoką częstotliwością dokonywanych transakcji, to wyniki tych badań pozwoliły określić profesjonalnych inwestorów jako zbyt konserwatywnych, co przejawiało się w zbyt wolnym reagowaniu na napływające informacje²⁷³. Mechanizm ten określany jest mianem heurystyki konserwatyizmu. Wynika on z niedostatecznej weryfikacji prawdopodobieństwa wystąpienia jakiegoś zdarzenia pod wpływem pojawiania się informacji²⁷⁴. Można stwierdzić, że nadmierna pewność siebie i heurystyka konserwatyizmu skonstruowane są na kanwie tego samego uprzedzenia poznawczego, związanego ze zwracaniem nieproporcjonalnej uwagi na siłę oraz wagę dowodu²⁷⁵. Dale Griffin i Amos Tversky opisują różnicę między siłą a wagą w następujący sposób:

rozróżnienie między siłą dowodu a jego wagą jest ściśle związane z rozróżnieniem między wielkością efektu (np. różnica między dwoma średnimi) a jego wiarygodnością (np. błąd standardowy różnicy). Chociaż nie zawsze jest możliwe rozłożenie wpływu dowodów na osobne wkłady siły i wagi, istnieje wiele kontekstów, w których można je wymieniać niezależnie. Silna lub słaba rekomendacja może pochodzić z wiarygodnego, lub niewiarygodnego źródła [...]²⁷⁶.

Zatem nawet jeśli do nowej informacji inwestor przypisze dużą wagę, ale informacja ta nie będzie miała wystarczającej siły, w sferze decyzyjnej może pojawić się nadmierny konserwatyzm. Z kolei, jeśli siła informacji będzie duża, a jej waga zbyt niska, inwestor może przejawiać nadmierną pewność siebie²⁷⁷.

Można powiedzieć, że w opozycji do nadmiernej pewności siebie stoi mechanizm zachowań stadnych (ang. *herding behavior*), który jest jednym z czynników, mogących sugerować niezgodność rzeczywistych działań inwestorów z modelem paradygmatu neoklasycznego. Inaczej mechanizm ten nazywany jest naśladownictwem. Na podstawie przeglądu badań stwierdzić można, że wśród inwestorów instytucjonalnych udało się go

²⁷³ Zob. Koc, J. O., *Do Institutional Investors Get Affected by Biases in Their Individual Investments*, „International Journal of Business & Management Studies”, 2022, nr 3/1, s. 54.

²⁷⁴ Zielonka, P., *Gielda i psychologia...*, op.cit., s. 68.

²⁷⁵ Griffin, D.; Tversky, A., *The weighing of evidence and the determinants of confidence*, „Cognitive Psychology”, 1992, nr 24, s. 418.

²⁷⁶ Ibidem, s. 412.

²⁷⁷ Ibidem, s. 413.

zaobserwować na rynkach o różnych charakterystykach, w tym rynkach rozwiniętych, jak: Stany Zjednoczone²⁷⁸, Niemcy²⁷⁹, czy Japonia²⁸⁰. Badania nad tym mechanizmem przeprowadzone zostały również na rynku polskim, ale ich autorom udało się potwierdzić jego występowanie jedynie wśród inwestorów indywidualnych²⁸¹. W ramach literatury przedmiotu wskazywanych jest kilka, charakterystycznych zwłaszcza dla menedżerów, przyczyn takich zachowań, związanych zarówno z czynnikami psychologiczno-społecznymi, jak i ekonomicznymi. Po pierwsze, zaobserwowano, że skłonność do naśladownictwa zwiększa się wraz ze słabnącą jakością i jednoznacznością informacji. Podejmowanie decyzji na podstawie obserwacji aktywności innych inwestorów – w takich warunkach – tłumaczone jest strachem przed utratą reputacji. Zarządzający portfelami aktywów skłonni są odstąpić od obserwowanej w branży tendencji wtedy, gdy posiadają informacje, które nazwać można poufnymi, a informacje te traktują jako pewne²⁸². Drugim istotnym czynnikiem jest sposób wynagradzania osób odpowiedzialnych za decyzje inwestycyjne, a dokładniej relacji wysokości opcjonalnej premii do kary w postaci utracenia części potencjalnego wynagrodzenia²⁸³. Takie myślenie wynikać może z – opisaną w teorii perspektywy – awersji do straty. Należy przy tym zaznaczyć, że w ramach finansów behawioralnych, przyszły przepływ pieniężny często traktowany jest jako już zaksięgowany w ludzkim umyśle, a więc będący punktem odniesienia ewentualnej straty. Ponadto, osiągnęty przez osoby zarządzające portfelami wynik, w postaci określonej stopy zwrotu, również posiada swój punkt odniesienia, w postaci zmiany wartości odpowiedniego indeksu lub uzyskanej w danej grupie funduszy średniej stopy zwrotu. Awersja do straty stanowić może więc motywację dla korzystania z mechanizmu naśladownictwa, ponieważ zmniejsza ona

²⁷⁸ Zob. Sias, W. R., *Institutional Herding*, „Review of Financial Studies”, 2004, nr 17/1, s. 165–206.

²⁷⁹ Zob. Menkhoff, L., Schmidt, U., Brozynski, R., *The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence*, „European Economic Review”, 2006, nr 50, s. 1753–1766.; Zob. Lütje T., *To be good or to be better: Asset managers' attitudes towards herding*, „Applied Financial Economics”, 2009, nr 19, s. 825–839.

²⁸⁰ Zob. Suto, M.; Toshino, M., *Behavioural biases of Japanese institutional investors: Fund management and corporate governance*, „Corporate Governance: An International Review”, 2005, nr 13/4, s. 466–477.

²⁸¹ Zob. Goodfellow, Ch., Bohl, M. T., Gebka, B., *Together we invest? Individual and institutional investors' trading behaviour in Poland*, „International Review of Financial Analysis”, 2009, nr 18, s. 212–221. Należy zaznaczyć, że okres badawczy obejmował lata 1996–2000, a więc mieścił się on w pierwszej dekadzie funkcjonowania Giełdy Papierów Wartościowych w Warszawie.

²⁸² Szyszka, A., *Finanse behawioralne...*, op.cit., s. 87.

²⁸³ Ibidem, s. 88.

ryzyko uzyskania wyniku niższego niż średnia, a tym samym utraty potencjalnej premii²⁸⁴.

Za konsekwencję zachowań stadnych – zwłaszcza wśród menedżerów zarządzających dużymi portfelami – uznaje się oddalanie się cen walorów od ich wartości fundamentalnej. Wynika to, chociażby z faktu, że w znacznie większym stopniu do podejmowania decyzji wykorzystywane będą informacje o charakterze niefundamentalnym²⁸⁵. Dodatkowo zachowania stadne przyczynić mogą się do nadreakcji rynku na napływające informacje, a w ujęciu krótkoterminowym wywoływać tendencję do kontynuowania kierunku dynamiki stóp zwrotu lub odwrotnie – skutkować nadmierną zmiennością cen. Jeżeli bowiem intencją menedżera jest odwzorowywanie składu obranego benchmarku, to zmiana wartości walorów spółki spowoduje również zmiany w strukturze benchmarku (np. wzrost lub spadek udziału danej spółki w indeksie). W konsekwencji, na rynku pojawi się popyt na akcje, znajdujące się w trendzie wzrostowym, a zmniejszy się zainteresowanie walorami, które w strukturze benchmarku zaczną zajmować mniej miejsca²⁸⁶. Pojawienie się na rynku takiego procesu, zaowocuje zmniejszeniem jego efektywności w rozumieniu hipotezy efektywności informacyjnej.

Fundamentalnym problemem, którego jak dotąd nie udało się jednoznacznie rozstrzygnąć, a który autorka starała się przybliżyć, jest adekwatny do rzeczywistości opis funkcjonowania rynków finansowych. Jak zostało scharakteryzowane, perspektywa neoklasyczna i behawioralna, kreują odrębną wizję zarówno rynku jako całości, jak i mechanizmów kierujących motywacjami inwestorów. Jeżeli hipotezę efektywności informacyjnej przyjąć za w pełni słuszną, to – jak powiedział Fama – badanie wykresów stałoby się działaniem podobnym do astrologii i dla inwestorów pozostałoby bez

²⁸⁴ Por. Szyszka, A., *Finanse behawioralne. Nowe podejście do inwestowania na rynku kapitałowym*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2009, s. 88 oraz Borowski, K., *Finanse behawioralne. Modele*, Wydawnictwo Difin, Warszawa 2014, s. 45.

²⁸⁵ Lütje T., *To be good or to be better: Asset managers' attitudes towards herding*, „Applied Financial Economics”, 2009, nr 19, s. 836–837.

²⁸⁶ Borowski, K., *Finanse behawioralne...*, op.cit., s. 46.

wartości²⁸⁷. Analogicznie sytuacja dotyczyłaby analizy fundamentalnej. Fama wykluczył bowiem możliwość osiągnięcia na jej podstawie zysków lub przynajmniej – jak pisze Burton Malkiel – postawił on pól silną efektywność w opozycji do wielu przekonań podtrzymywanych przez osoby wykorzystujące analizę fundamentalną²⁸⁸. Inaczej jednak mechanizm ten wyglądałby przy założeniu, że gracze rynkowi mają asymetryczny dostęp do informacji, obarczeni są różnego typu implikacjami natury psychologicznej, a w efekcie nie przetwarzają informacji w taki sposób, jaki byłoby zgodny z założeniami racjonalności. Przyjęcie takiej optyki może stanowić podstawy do zrozumienia anomalii związanych z autokorelacją stóp zwrotu, czy też nadmierną lub niedostateczną reakcją rynku na napływające informacje. Dlatego też w kolejnej części dysertacji autorka proponuje dokonać analizy czynników, poprzedzających zarówno mechanizmy poznawcze, jak i procesy decyzyjne inwestorów rynkowych. Chodzi tu o sposób, w jaki informacja w ogóle jest konstruowana. Z uwagi jednak na rozpatrywany problem badawczy, analiza ta dotyczyć będzie informacji publicznych, prezentowanych w postaci artykułów prasowych. Pytanie, jakie należy sobie bowiem zadać, dotyczy tego, czy otrzymywane przez inwestorów informacje wyposażone są w – opisywane w teorii perspektywy – kadry, stanowiące punkt wyjścia dla różnego typu implikacji behawioralnych, czy też traktować można je jako obiektywny przekaz wydarzeń ekonomicznych, co niewątpliwie ułatwiałoby inwestorom jednoznaczną i racjonalną ocenę sytuacji rynkowej.

²⁸⁷ Fama, E. F., *Random Walks...*, op.cit., s. 57.

²⁸⁸ Malkiel, B.G., *Błądząc po Wall Street. Dlaczego nie można wygrać z rynkiem*, WIG-Press, Warszawa 2003. s. 118.

Rozdział III – Dziennikarz jako nadawca informacji na rynku finansowym

Kontynuując rozważania nad istotą informacji, należy zdefiniować kim – zwłaszcza w kontekście rynku finansowego – jest jej nadawca. W najprostszym ujęciu można oczywiście stwierdzić, że takim nadawcą jest emitent papierów wartościowych, który kieruje informacje do swoich akcjonariuszy²⁸⁹. Warto jednak zauważyć, że w codziennej praktyce rynkowej katalog nadawców zdaje się ulegać systematycznemu rozszerzaniu, a założenie, że rolę tę pełnią jedynie spółki – lub ujmując to szerzej: zorganizowane podmioty rynkowe – jest coraz trudniejsze do podtrzymania. Przykładowo, z badań prowadzonych przez Stowarzyszenie Inwestorów Indywidualnych wynika, że większość ankietowanych do pozyskiwania informacji wykorzystuje serwisy internetowe (96,5%), serwis YouTube (73,6%), blogi inwestorskie (58,2%), gazety (57,7%), czy też fora dyskusyjne (55%)²⁹⁰. Należy więc stwierdzić, że spółki emitujące papiery wartościowe są obecnie autorami jedynie części informacji, jakie otrzymują inwestorzy. Z uwagi na to, że w przeprowadzonych przez autorkę badaniach wykorzystywanym źródłem danych tekstowych są artykuły prasowe (dostępne w formie elektronicznej), w tej części pracy uwaga zostanie skoncentrowana na nadawcy informacji w rozumieniu dziennikarza finansowego.

Komunikacja medialna jest specyficznym sposobem przekazywania informacji, odgrywającym coraz większą rolę w kontekście funkcjonowania rynku finansowego. Umieszczenie problematyki niniejszej dysertacji w tej przestrzeni, pozwala z jednej strony zadać pytanie o sam przebieg omawianego procesu, z drugiej o to, co sprawia, że wydarzenia rynkowe stają się informacjami. Tak postawiony problem daje możliwość wykazania, że przekaz biegnący od nadawcy do odbiorcy, obarczony jest wpływem czynników endo- i egzogenicznych, umiejscowionych zarówno po jednej, jak i po drugiej stronie całego procesu. Podgatunkiem komunikacji medialnej jest tzw. dziennikarstwo EBF. Akronim ten definiowany jest jako media skoncentrowane wokół treści

²⁸⁹ Będowska-Sójka, B., *Wpływ informacji na ceny instrumentów finansowych. Analiza danych śróddziennych*. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2014, s. 16 [za: Tirole, 2006].

²⁹⁰ Stowarzyszenie Inwestorów Indywidualnych, *Oto najpopularniejsze źródła informacji dla inwestorów* [OBI 2022], <https://www.sii.org.pl/15945/aktualnosci/badania-i-rankingi/oto-najpopularniejsze-zrodla-informacji-dla-inwestorow-obi-2022.html>. Dostęp: 8 mar. 2023.

ekonomicznych, biznesowych oraz finansowych (ang. *economic, business and financial journalism*). Mimo że klarowny podział na trzy podgrupy, tworzące ten typ dziennikarstwa, wydaje się być teoretycznie trafny, to w praktyce trudno wyobrazić sobie wyraźną linią oddzielającą treści biznesowe od finansowych, czy też finansowe od makroekonomicznych. Chociaż z oczywistych względów autorka w najwyższym stopniu zainteresowana jest problematyką dziennikarstwa finansowego, to w dalszej części rozprawy zwrot ten używany będzie przy pełnym zrozumieniu faktu, że informacje finansowe nie są przedstawiane w oderwaniu od pozostałych treści EBF²⁹¹. Jednocześnie należy zwrócić uwagę, że o ile istnieje dorobek badawczy w zakresie wpływu ogłoszeń emitentów na dynamikę walerów rynkowych, to analiza relacji, jaka zachodzi między narracją dziennikarską a decyzjami inwestorów, wciąż stanowi badawczą lukę.

Celem poniższego rozdziału jest scharakteryzowanie czynników mogących mieć wpływ na subiektywizację treści, prezentowanych przez dziennikarzy finansowych. Istnieje bowiem szereg czynników mogących mieć wpływ na subiektywizację przekazywanych przez dziennikarzy informacji. Czynniki te podzielić można na trzy grupy. Pierwsza związana jest ze zjawiskami zewnętrznymi wobec dziennikarzy, takimi jak oczekiwania inwestorów giełdowych oraz wyzwania związane z przemianami technologicznymi. Druga grupa to czynniki, które wydają się przez dziennikarzy internalizowane w procesie zawodowej adaptacji. Chodzi tu przede wszystkim o specyficzne definiowanie wartości informacyjnych, czyli takich wydarzeń lub treści, które z różnych względów uznawane są za godne opublikowania. Trzecia grupa czynników w największym stopniu wynika z osobistych uproszczeń, emocji i nastawienia dziennikarza. Wszystkie opisywane grupy zamykają się w problematyce określanej w literaturze przedmiotu mianem teorii selekcji informacji (ang. *gatekeeping*), która przejawia się zarówno na poziomie szeroko pojętej kultury finansowej rzeczywistości, praktyki funkcjonowania w organizacji, jak i w ramach samego procesu tworzenia przez dziennikarzy publikowanych treści. Zasadne wydaje się również przyjęcie przesłanki mówiącej o tym, że efektem procesu selekcji informacji jest jej specyficzne kadrowanie, tj. mimowolne lub celowe umiejscawianie jej w kontekście,

²⁹¹ Zob. Merrill, J. G., *The Political Content of British Economic, „Business and Financial Journalism”*, Palgrave Macmillan, Cham, 2019.

mogącym sugerować optymalny – zdaniem dziennikarzy – sposób interpretacji konkretnych wydarzeń rynkowych²⁹².

3.1. Kadrowanie jako immanentna cecha przekazywania informacji

Przegląd literatury przedmiotu pozwala stwierdzić, że istnieją powody, aby sądzić, że media finansowe odgrywają istotną rolę w kształtowaniu oczekiwań inwestorów i postrzeganiu przez nich konkretnych zdarzeń. Na mechanizm ten należy jednak patrzeć jako na wielowarstwowy i złożony system wzajemnych interakcji i powiązań²⁹³. Z jednej strony dzięki wiadomościom finansowym, poprzez które mimowolnie uzasadniana jest pewna logika rozumienia rynku, dyskurs na temat finansów jest regularnie utrwalany. Z drugiej jednak strony, należy zdawać sobie sprawę, że odzwierciedlanie opisywanej rzeczywistości nie może być rozumiane jako jej prosta reprezentacja.

Ten niejednoznaczny wpływ, jaki potencjalnie dziennikarz finansowy ma na odbiorcę informacji, ubrać można w ramy koncepcji określanej mianem *kadrowania*. W ostatnich latach koncepcja ta zyskała na popularności między innymi dzięki – opisywanym w poprzedniej części pracy – paradygmatowi finansów behawioralnych i teorii perspektywy zaproponowanej przez Daniela Kahnemana i Amosa Tverskiego. Na tym etapie rozważań, istotne wydaje się wyjaśnienie istoty tego zjawiska w kontekście nadawcy informacji, to jest dziennikarza finansowego. Jak już zastało opisane, idea kadrowania bazuje na założeniach związanych z procesami poznawczymi, to jest założeniach o istnieniu ustrukturyzowanych reprezentacji kognitywnych i kierowaniu się przez ludzi zasadami zgodnymi z teoriami przetwarzania informacji. Założenia te podzielane są przez przedstawicieli wielu dyscyplin zorientowanych na analizę procesów poznawczych, w tym psychologów i socjologów²⁹⁴.

W najprostszym ujęciu przyjąć można, że zadaniem dziennikarzy finansowych jest relacjonowanie wydarzeń rynkowych. Jeżeli spoczywające na nich obowiązki

²⁹² Zob. Entman, R. M., Framing: Toward Clarification of a Fractured Paradigm, „Journal of Communication”, 1993, nr 43/1. Zob. także. de Vreese, C.H., *News Framing: Theory and Typology*, „Information Design Journal”, 2005, nr 13/1.

²⁹³ Thompson, P. A., *Market Manipulation? Applying the Propaganda Model to Financial Media Reporting*. „Westminster Papers in Communication and Culture”, 2009, nr 6/2, s. 90.

²⁹⁴ Pan, Z.; Kosicki, G. M., *Framing analysis: An approach to news discourse*, „Political Communication”, 1993, nr 10, s. 56.

ograniczyć by jedynie do takiej roli, oznaczałoby to konieczność prezentowania przez nich wyłącznie faktów. Choć obecnie można spotkać wiele przykładów aktywności nadawców informacji ograniczających się jedynie do przedstawiania wydarzeń w postaci listy punktów, to z reguły efektem pracy dziennikarza jest artykuł lub chociażby notatka prasowa. W każdym z tych przykładów dziennikarz, świadomie lub nie, umieszcza wydarzenie w jakimś kontekście. Ten kontekst stanowi dla wydarzeń kadr, który zmienić może ich znaczenie i charakter.

W artykułach prasowych fakt nie pozostaje w izolacji od kontekstu²⁹⁵. Definicja kontekstu jest jednak w tym przypadku rozumiana bardzo szeroko. Jeżeli postrzegać tekst jako zorganizowany system elementów znaczących (lub też słów), to ich dobór sugerować może zarówno poparcie dla pewnych idei, jak i dostarczać narzędzi zachęcających odbiorców do przetwarzania informacji w określony sposób. Dzieje się tak z uwagi na to, że z jednej strony elementy te definiowane są przez odbiorców i nadawców zgodnie z pewnymi wspólnymi konwencjami, a z drugiej – ponieważ nie istnieje bezpośrednia zgodność między słowem a jego desygnatem. Taka dowolność relacyjna może być wykorzystywana przez dziennikarzy do zmaksymalizowania prawdopodobieństwa uzyskania preferowanych znaczeń²⁹⁶.

Jeśli natomiast potraktować tekst jako przejaw twórczej pracy dziennikarzy, można stwierdzić, że kadrowanie pełni cztery funkcje. Po pierwsze, jest to określanie problemu, poprzez wskazanie jego podmiotu, podejmowanych przez niego działań i związanych z nim kosztów, czy też korzyści mierzonych wspólnymi wartościami kulturowymi. Po drugie, jest to diagnozowanie przyczyn problemu, to znaczy identyfikowanie sił, które doprowadziły do jego powstania. Po trzecie, jest to dokonanie moralnego osądu tego problemu, poprzez ocenę jego czynników sprawczych oraz ich skutków, a po czwarte – sugerowanie środków mogących zaradzić problemowi i wskazanie ich prawdopodobnych konsekwencji. Niezależnie od wykorzystanych przez dziennikarza narzędzi, stwierdzić można, że istota tego zjawiska obejmuje dwa procesy: selekcję oraz znaczenie. Innymi słowy, kadrowanie jest więc selekcionowaniem przez

²⁹⁵ Butterick, K. J., *A Churnalism, Complacency and Collusion: A Critical Introduction to Business and Financial Journalism*, Pluto Press, Londyn 2015, s. 75.

²⁹⁶ Pan, Z.; Kosicki, G.M., *Framing Analysis...*, op.cit., s. 55–59. Zob. także. De Saussure, F., *Kurs językoznawstwa ogólnego*, Państwowe Wydawnictwo Naukowe, Warszawa 1961.

autorów pewnych aspektów rzeczywistości i uwypuklaniem ich w tekście, tak aby uwydatnić określony sposób definiowania problemu oraz jego interpretacji i oceny, to jest jego znaczenia²⁹⁷.

Zdaniem Roberta Entmana, w procesie komunikacji medialnej, kadry ukryte są w czterech jej elementach: narzędziu komunikacji, tekście, odbiorcy oraz kulturze. Przy użyciu każdego z tych elementów dziennikarze finansowi, w procesie przekazywania informacji, dokonują osądów, kierując się przy tym schematami, pozwalającymi na organizację ich systemu przekonań. W efekcie publikowane przez nich teksty zawierają kadry, które uwidaczniają się poprzez „obecność lub brak pewnych słów, zwrotów giełdowych, stereotypowych obrazów, źródeł informacji i zdań, które stanowią tematycznie wzmacniające zbiory faktów lub osądów”²⁹⁸. Jednocześnie, odbiorcy informacji również wyposażeni są we właściwe sobie schematy i uproszczenia poznawcze, które wpływają na ich sposób wnioskowania²⁹⁹. Dlatego właśnie konieczne jest uwzględnienie w ramach procesu komunikacyjnego percepcji jego uczestników.

Na ten problem zwrócono uwagę już w pierwszych pracach na temat selekcji informacji. W latach 60. XX w. Johan Galtung i Mari Holmboe Ruge zaproponowali model komunikacji medialnej, pokazujący, że między wydarzeniem a odbiorcą występują przynajmniej dwa momenty, w których dochodzi do zniekształcenia obrazu wydarzenia poprzez: (1) percepcję mediów; (2) percepcję odbiorcy informacji³⁰⁰. Ponieważ opisywana teoria odnosi się do relacji, jaka zachodzi między prezentowanymi w mediach kadrami a wcześniejszą wiedzą i predyspozycjami uczestników komunikacji, to należy mieć na uwadze, że zarówno w przypadku nadawców, jak i odbiorców informacji, kadrowanie może mieć wpływ na proces przyswajania wiedzy, interpretacji problemów i oceny wydarzeń.

²⁹⁷ Entman, R. M., *Framing: Toward Clarification of a Fractured Paradigm*, „Journal of Communication”, 1993, nr 43/1, s. 52.

²⁹⁸ Ibidem, s. 52–53.

²⁹⁹ Ibidem.

³⁰⁰ Zob. Galtung, J., Ruge, M. H., *The Structure of Foreign News*, „Journal of Peace Research”, 1965, nr 2/1, s. 65.

Konsekwencje kadrowania widoczne są zarówno na poziomie społecznym, jak i indywidualnym³⁰¹. Poziom społeczny rozumieć można na dwa sposoby. Po pierwsze, na dyskurs informacyjny składają się pewne nieuchwytnie, ale szeroko podzielane przekonania na temat społeczeństwa, czy też konkretnych społeczności, które mieszczą się w przestrzeni czegoś, co powszechnie nazywane jest „zdrowym rozsądkiem”, czy też mądrością w jej konwencjonalnym rozumieniu. Przekonania te określają kadry w szerokim ujęciu. To właśnie one wyznaczają krańce pewnego zamkniętego układu, wewnątrz którego wiadomości mogą być konstruowane, transmitowane i rozwijane³⁰². W tym sensie kadrowanie wykracza poza funkcję porządkowania myśli, pozwalając na budowanie dominującej perspektywy, dzięki której informacje są w ogóle zrozumiałe i mogą być komentowane w określony sposób³⁰³. Po drugie, kadry mogą przyczynić się do kształtowania procesów związanych z socjalizacją, postawami politycznymi i działaniami zbiorowymi. Konsekwencją tego mechanizmu na poziomie indywidualnym może być tymczasem zmiana nastawienia poszczególnych osób do jakiejś sprawy poprzez wystawienie ich na ekspozycję konkretnych kadrów³⁰⁴. Na gruncie finansów egzemplifikacją tego zjawiska jest – opisywane już – kadrowanie sytuacji w kategoriach potencjalnej straty lub potencjalnego zysku w odniesieniu do punktu referencyjnego³⁰⁵. O ile jednak badacze z zakresu finansów systematycznie poszerzają wiedzę na temat wpływu kadrowania informacji na decyzje inwestycyjne, to analiza powodów, narzędzi i mechanizmów tworzenia kadrów, wydaje się stanowić niszowe pole zainteresowania.

Michael Fleischer zauważa, że media jedynie umożliwiają budowanie konstruktów, stojącym za nimi „systemom świadomościowym”, pozwalają na „wytwarzanie fikcji konsensusu”³⁰⁶. Zwrot ten wydaje się nader adekwatny dla poruszanej problematyki. Po pierwsze, proces przedstawiania treści informacyjnych jest

³⁰¹ de Vreese, C.H., *News Framing: Theory and Typology*, „Information Design Journal”, 2005, nr 13/1, s. 52.

³⁰² Pan, Z.; Kosicki, G. M., *Framing analysis...*, op.cit., s. 57.

³⁰³ Piao, J., *Financial Media, Globalisation, and China's Economic Integration: Comparing Narrative Construction of The Economist and Caijing*. Rozprawa doktorska. University of Westminster 2015, s. 128.

³⁰⁴ de Vreese, C. H., *News framing...*, op.cit., s. 52.

³⁰⁵ Wiseman, R. M., Gomez-Mejia, L. R., *A behavioral agency model of management risk taking*, „Academy of Management Review”, 1998, nr 23/1, s. 136.

³⁰⁶ Fleischer, M., *Zarys ogólnej teorii komunikacji*. [w:] red. nauk. Graszewicz, M., Jastrzębski, J. *Teorie komunikacji i mediów*, Oficyna Wydawnicza ATUT, Wrocław 2009, s. 39.

w pewnym sensie ich produkcją. Po drugie, owa produkcja rozgrywa się w przestrzeni konsensusu, co implikuje, że nie wszystkie wydarzenia rynkowe w ów konsensus mają szansę się zmieścić. Po trzecie, konsensus ten jest jedynie iluzoryczny, bowiem nie sposób uzyskać symetrię przy wyborze publikowanych wydarzeń, ani obiektywizmu podczas tworzenia newsów.

W konsekwencji powyższego stwierdzenia naturalnym wydaje się zadanie pytania o to, jakie czynniki decydują o tym, które wydarzenia stają się informacją – tj. w tym kontekście: *newsem* – a które nie. Kto jest odpowiedzialny za tę selekcję i jakimi kieruje się przesłankami? Czy przesłanki te mają charakter wewnątrzorganizacyjny, czy są względem tej organizacji zewnętrzne? Czy ich lista jest stała, czy dynamiczna? I wreszcie, czy są immanentne dla wydarzenia, czy to nadawcy wyposażają je w cechy godne upublicznienia³⁰⁷? Jak sugeruje wielu badaczy, media wybierają wydarzenia, które chcą relacjonować, według pewnego zestawu kryteriów użyteczności publikacyjnej³⁰⁸. Sam *gatekeeping* zdefiniować można w ramach dwóch ujęć. W ujęciu wąskim należy rozumieć go niemal dosłownie jako pewną liczbę bramek-strażników, przez które musi przejść artykuł, aby finalnie trafić do publikacji. W tym sensie proces selekcji informacji można rozpatrywać w kategoriach linearnych, podobnie jak dzieje się to w przypadku transmisyjnego ujęcia komunikacji. Zaczyna się on już na poziomie samego emitenta, gdzie funkcję taką pełnić mogą specjaliści, tacy jak: prawnicy, działy *public relations*, analitycy finansowi, audytorzy, czy też agencje ratingowe. Ich zadaniem jest wspieranie emitentów podczas konstruowania różnego typu dokumentów przekazywanych do zewnętrznego użytku, jak na przykład sprawozdania finansowe, czy też strategiczne decyzje zarządów³⁰⁹. Selekcja wyraźnie widoczna jest również po stronie podmiotów publikujących informacje wtórnie wobec emitentów, gdzie strażnikami mogą być sami

³⁰⁷ Palczewski, M., *News values w amerykańskich i europejskich teoriach selekcji informacji*. [w:] red. nauk. Graszewicz, M., Jastrzębski, J. *Teorie komunikacji i mediów*. Oficyna Wydawnicza ATUT Wrocław, 2009. s. 200–201.

³⁰⁸ Zob. Fowler, R., *Language in the news: Discourse and ideology in the press*, Routledge, Nowy Jork 1991, s. 13.

³⁰⁹ Borden, M. J., *The Role of Financial Reporting in Corporate Governance*, „Fordham Journal of Corporate & Financial Law”, 2007, nr XII, s. 321–322.

redaktorzy, jak również wydawcy, którzy podejmują decyzje o charakterze dystrybucyjnym, czy też marketingowym³¹⁰.

W szerokim ujęciu istotą tego procesu jest jednak kreowanie społecznej rzeczywistości. Za Pamelą Shoemaker wyróżnić można tutaj pięć poziomów selekcji informacji. Pierwszy związany jest z indywidualnymi cechami „strażników”, reprezentowanymi przez nich wartościami, doświadczeniem zawodowym. Drugi poziom opiera się na rutynie związanej z komunikacją oraz pracą zawodową. Przejawiają się w nim zwyczajowe wartości informacyjne (ang. *news values*), które zostaną szerzej omówione w dalszej części rozdziału. Kolejno mechanizm selekcji zaobserwować można na poziomie organizacyjnym. Poziom ten różni się od poprzedniego tym, że „czynniki newsotwórcze” podzielane są tutaj w ramach wąskiej grupy osób zatrudnionych w danej organizacji. Wartości informacyjne natomiast akceptowane są na szerszej płaszczyźnie kulturowej pracy dziennikarskiej. Poziom czwarty wykracza poza rzeczywistość stricte dziennikarską i przenosi problematykę na grunt instytucjonalny. Gatekeeping widoczny jest tutaj w źródłach informacji, reklamodawcach, szeroko pojętych grupach interesu, konkurencji rynkowej. Ostatni, poziom systemu społecznego, ma podłoże kulturowe i ideologiczne³¹¹. Gatekeeping w szerokim znaczeniu odnosi się więc nie tylko do samego procesu wydawania lub niewydawania produktów, będących informacjami. Chodzi również o możliwość udzielania lub odmawiania dostępu do różnych głosów w społeczeństwie. Tym samym proces ten często staje się przestrzenią konfliktu³¹². Na konstruowanie oraz publikowanie wiadomości należy więc patrzeć jako na wieloetapowy, złożony proces, podlegający oddziaływaniu czynników o charakterze społecznym, kulturowym, ekonomicznym, jak również psychologicznym³¹³.

Rozwój literatury związanej z analizą mediów zaowocował propozycjami modeli, poprzez które autorzy starają się wyjaśnić tego typu procesy komunikacyjne.

³¹⁰ McQuail, D., *McQuail's Mass Communication Theory*, SAGE Publications, Londyn 2010, s. 403.

³¹¹ Palczewski, M., *News values...*, op.cit., s. 213–214; [za Shoemaker, P. J., *Gatekeeping. Communication Concepts*, Sage Publications, Londyn 1991, s. 3].

³¹² McQuail, D., *McQuail's...*, op.cit., s. 403.

³¹³ Merrill, G. J., *The revolution must wait: Economic, business and financial journalism beyond the 2008 crisis*, „Ethical Space: The International Journal of Communication Ethics”, 2012, nr 9/1, s. 3. Zob. także. Merrill, G. J., *The Political Content of British Economic, Business and Financial Journalism*, Palgrave Macmillan, Cham, 2019. s. 10.

Przykładowo Zhongdang Pan i Gerald M. Kosicki opracowali schemat, który nazwali „procesem dyskursu w mediach informacyjnych”. Autorzy wykazali niebywałą złożoność relacji zachodzących między instytucjami, będącymi źródłami informacji, a budowaniem przekazu medialnego, w ramach którego dziennikarz ograniczony jest czynnikami organizacyjnymi, konwencjami zawodowymi, czy też mimowolną koniecznością przewidywania odbioru publikowanych treści. W zbudowanym przez nich schemacie, proces inicjowany jest w momencie, w którym dochodzi do wydarzenia związanego ze źródłem informacji, które to wydarzenie charakteryzowane jest jako „posiadające wartości informacyjne” lub gdy sam dziennikarz poszukuje informacji o danym problemie. Proces ten przyjmuje strukturę kołową, a jego elementy opierają się o endogenne pętle. Pokazuje to, że wszyscy uczestnicy komunikacji – w związku z pełnionymi przez siebie rolami – projektują, konstruują i transmitują informacje³¹⁴. Inny model, sugerujący również wstępną klasyfikację czynników kadrowania, zaproponowany został przez Claesa H. de Vreese. Autor zwraca uwagę, że komunikacja jest procesem dynamicznym, na który składają się dwa elementy. Pierwszy związany jest z tworzeniem kadrów, czy też sposobem ich wyłaniania się na kanwie strukturalnych czynników pracy dziennikarskiej (ang. *frame-building*). Drugi opiera się na „ustawianiu kadrów” (ang. *frame-setting*), polegającym na wzajemnym oddziaływaniu między przedstawianymi w mediach kadrami a predyspozycjami odbiorców³¹⁵.

Na podstawie przedstawionych modeli zauważyć można, że zrozumienie koncepcji kadrowania wymaga wykroczenia poza samą treść publikowanych informacji. Wydaje się, że zagadnienia związane z ich selekcją sklasyfikować można w trzech grupach. Po pierwsze, są to kwestie niezależne od dziennikarza, czyli takie, które może potraktować on jako zastane, przystępując do pracy w określonych warunkach organizacyjnych, czy też społecznych. W ramach tej grupy wyróżnić można natomiast uwarunkowania zewnętrzne wobec organizacji, w której pracuje dziennikarz oraz uwarunkowania wewnętrzne. Te pierwsze oparte są o relacje między nadawcą a innymi podmiotami jak: państwo, odbiorcy, konkurencja czy źródła danych. Uwarunkowania wewnętrzne z kolei oparte są o zamknięty krwiobieg organizacji, tj. właścicieli, czy też

³¹⁴ Pan, Z., Kosicki, G. M., *Framing Analysis...*, op.cit., s. 57.

³¹⁵ de Vreese, C. H., *News Framing...*, op.cit., s. 51.

innych pracowników³¹⁶. Po drugie, są to czynniki, których istnienie jest w dużej mierze od dziennikarza niezależne, ale w trakcie edukacji i pracy zawodowej mogą być przez niego zinternalizowane. Przykładem takich czynników są wspomniane już „wartości informacyjne” (ang. *news values*). Trzecią grupę stanowią czynniki umiejscowione w samym dziennikarzu, takie jak jego światopogląd, postawa i emocje. W dalszej części autorka postara się scharakteryzować te grupy, a tym samym wykazać, że przekazywana przez dziennikarzy informacja nie jest i nie powinna być rozumiana jedynie jako obiektywne referowanie wydarzeń, mających miejsce na rynku finansowym.

3.2. Zewnętrzne czynniki selekcji informacji

Czynniki zewnętrzne wobec dziennikarza potraktować można jako swoisty kontekst, w którym dziennikarz finansowy realizuje swoją rolę. Wydaje się, że nakreślić można tutaj dwie płaszczyzny, będące owymi kontekstami. Pierwsza związana jest z szeroko pojętą misją społeczną. W jej ramach dziennikarstwo finansowe rozpatrywać można przez pryzmat przemian gospodarczych, potencjalnego wpływu na nastroje społeczne, czy też relacji łączącej je z rzeczywistością polityczną. Druga płaszczyzna pozwala spojrzeć na rolę dziennikarzy finansowych w wąskiej perspektywie, tj. wysnuwanych względem nich oczekiwań związanych z ich miejscem na rynku finansowym. Obie te płaszczyzny potraktować można jako element mechanizmu selekcji informacji. Z jednej strony bowiem dziennikarze mogą odczuwać presję związaną z odpowiadaniem na tego typu oczekiwanie. Z drugiej natomiast na drodze do ich realizacji stać mogą obiektywne, jak również subiektywne bariery, ograniczające lub zniechęcające dziennikarzy do realizowania ich misji.

Z uwagi na tematykę dysertacji autorka skoncentruje się na drugiej płaszczyźnie, w ramach której dziennikarz finansowy to osoba objaśniająca i interpretująca występujące na rynku zdarzenia³¹⁷. Jego aktywność jest wynikiem istnienia luki „między tym, co społeczeństwo wie, co chce wiedzieć, a tym, co musi wiedzieć [...], dzięki czemu jest ono na bieżąco z aktualnymi wydarzeniami, a tym samym jest w stanie pociągnąć

³¹⁶ Mrozowski, M., *Media masowe. Władza, rozrywka i biznes*, Oficyna Wydawnicza ASPRA-JR, Warszawa 2001, s. 233.

³¹⁷ Tambini, D., *What are financial journalists for?*. „Journalism Studies”, 2010, nr 11/2, s. 161–162.

system finansowy do odpowiedzialności”³¹⁸. Takie określanie roli dziennikarzy EBF sugeruje, że bliżej im do inwestorów, o których interesy pośrednio dbają, niż do zarządów spółek, których ewentualne nadużycia mogliby demistyfikować³¹⁹. Tym samym miejsce dziennikarstwa finansowego rozpatrywać można na tle szeroko pojętego systemu ładu korporacyjnego. Realizacja roli dziennikarstwa EBF przejawia się tutaj poprzez ujawnianie wydarzeń wskazujących na nieefektywne zarządzanie spółką lub wykorzystywanie w jej działalności nielegalnych praktyk. W tym kontekście dziennikarze finansowi mogą pełnić na rynku dwie funkcje. Po pierwsze, mogą wykrywać i nagłaśniać nadużycia finansowe, ostrzegając uczestników, a tym samym inicjując reakcję rynku. Po drugie, ujawnianie przez nich niewłaściwych zachowań skutkować może wszczynaniem postępowań przez instytucje regulacyjne, co w optymistycznym wariancie skłonić może zarządy spółek do zaprzestania stosowania nieuczciwych praktyk³²⁰.

Oba przypadki wymagają od dziennikarzy wiedzy i świadomości nie tylko z zakresu finansów, ale również zarządzania lub prawa. Wydaje się jednak, że w kierunku dziennikarstwa finansowego – szczególnie po kryzysie przełomu lat 2007/2008 – pada wiele oskarżeń, dotyczących powierzchowności, braku wystarczających kompetencji, czy też woli działania, które umożliwiłyby odkrycie rosnącej bańki w odpowiednim momencie. Oczywiście zarzuty te mają sens w sytuacji, w której skłonni jesteśmy osiągnąć konsensus, co do społecznej roli dziennikarzy finansowych. Zwłaszcza że opinie na ten temat – nawet wśród samych zainteresowanych – są podzielone³²¹. Nietrudno bowiem zauważyć, że natura rynku staje się coraz bardziej skomplikowana, a opłacalność finansowania dziennikarstwa specjalistycznego pozostaje wątpliwa. Należy mieć na uwadze, że media nie pozostają w izolacji od przemian społecznych oraz prawideł ekonomicznych. Uczestniczą w tym samym wyścigu gospodarczym, co każda inna organizacja. Tym samym podlegają mechanizmom konkurencyjności, konieczności optymalizacji kosztów i odpowiadania na oczekiwania odbiorców oraz partnerów

³¹⁸ Knowles, S., Philips, G., Lidberg, J., *Reporting The Global Financial Crisis: A longitudinal tri-nation study of mainstream financial journalism*, „Journalism Studies”, 2017, nr 18/3, s. 3.

³¹⁹ Butterick, K. J., *A Churnalism...*, op.cit., s. 76.

³²⁰ Zob. Borden, M. J., *The Role of Financial Reporting in Corporate Governance*, „Fordham Journal of Corporate & Financial Law”, 2007, nr XII, s. 323–332.

³²¹ Zob. Tambini, D., *What are financial...*, op.cit.

biznesowych. W przypadku mediów, partycypowanie w tym systemie niesie jednak osobliwe niebezpieczeństwo. Związane z nim czynniki mogą bowiem w specyficzny sposób wpływać na jakość efektów pracy. W literaturze przedmiotu jako przykłady takich czynników wymienia się między innymi: konieczność wydawania tekstów w formie online, czy też zwiększające się tempo pracy całych redakcji. Z drugiej jednak strony, jak wnioskuje Damian Tambini: „być może sami dziennikarze biznesowi i finansowi nie uważają się za zaangażowanych w raportowanie w *interesie publicznym* w taki sam sposób, jak robią to dziennikarze polityczni”³²².

Nie sposób podejmować się jednak próby analizy powyższych zagadnień, nie zwracając uwagi na problem kosztów. Autorka proponuje spojrzeć na nie z dwóch punktów widzenia. Po pierwsze, chodzi o koszty pracy dziennikarzy, których pewną częścią są koszty dotarcia do źródeł informacji. Po drugie, chodzi o stosunek owych kosztów do potencjalnych przychodów oraz źródeł ich pozyskiwania. Podstawowym zadaniem dziennikarzy jest „wiedzieć jak najwięcej”. Wiedza jest polem, na którym konkurują ze sobą dziennikarze, zatrudnieni w różnych redakcjach i na tym to polu budują oni swoją przewagę. Pozyskiwanie wnikliwej wiedzy na temat wszystkich firm działających w danej branży jest jednak aktywnością wyjątkowo kosztowną i czasochłonną. Dlatego też konieczne wydaje się zawężenie obszaru zainteresowań do ograniczonej liczby podmiotów. Sądzić można, że występują dwa czynniki, mające kluczowy wpływ na ich dobór. Pierwszy związany jest z potencjalną liczbą odbiorców prezentowanych treści. Wydaje się, że właściwy dobór spółek to taki, który pozwala przyciągnąć uwagę zarówno inwestorów indywidualnych, jak i instytucjonalnych³²³. Takie działanie daje możliwość zwiększenie liczby odbiorców tekstu, co będzie miało przełożenie na zainteresowanie reklamodawców. Drugi związany jest z poziomem wnikliwości wiedzy na temat poszczególnych spółek. Mniej kosztochłonne jest oczywiście korzystanie z publicznie dostępnych danych niż zdobywanie ich u źródła lub w inny bezpośredni sposób. Organizacja medialna zorientowana na zysk i funkcjonująca w konkurencyjnym środowisku może chcieć obniżać koszty produkcji. Dotyczy to również sposobu pozyskiwania danych. Z tego punktu widzenia wysyłanie reportera

³²² Ibidem, s. 159–160.

³²³ Butterick, K. J., *A Churnalism...*, op.cit., s. 79.

w celu przeprowadzenia śledztwa jest mniej opłacalne niż na przykład pozyskanie informacji od agencji prasowej³²⁴. Prawdopodobnie tę – w badaniach przeprowadzonych przez Damiana Tambiniego – potwierdza jeden z wydawców angielskiego czasopisma finansowego, który twierdzi, że „skierowanie dwóch lub trzech osób do projektu na miesiąc, pod koniec którego mogą nie uzyskać, jest czymś, co musielibyśmy bardzo intensywnie przemyśleć, ponieważ jest to drogie... Kiedyś mieliśmy małe jednostki dochodzeniowe, tak naprawdę już ich nie mamy”³²⁵. W efekcie rozsądnym wyborem ze strony dziennikarzy jest śledzenie aktywności dużych spółek. Przedsiębiorstwa o niższej kapitalizacji, mogą cieszyć się bowiem mniejszym zainteresowaniem wśród potencjalnych czytelników. Ponadto, pozyskanie na ich temat szczegółowych danych wymaga większych nakładów pod względem kosztów i czasu pracy³²⁶.

Kolejną kwestią są oczekiwania związane z tempem pracy. Czas konieczny do wyprodukowania artykułu z oczywistych względów przekłada się na liczbę opublikowanych newsów. Presja wydajności wymusza na dziennikarzach publikowanie większej liczby artykułów w czasie znacznie krótszym niż kilka dekad temu. Narzucanie krótkich terminów realizacji materiałów, zdaniem samych dziennikarzy, wpływa zarówno na poziom nadzoru redakcyjnego oraz podtrzymywania pewnego standardu weryfikacyjnego, jak i również na możliwość pozyskania dodatkowych informacji z nieszablonowych źródeł³²⁷. Dodatkowo, trend optymalizacji kosztów skutkować może ograniczaniem liczby etatów. W konsekwencji mniejsza liczba dziennikarzy wyprodukować musi większą liczbę artykułów, a taki model działania wpływać może na jakość i rzetelność wyników pracy³²⁸.

U podstaw konieczności optymalizacji kosztów pracy redakcji stoi spadek rentowności czasopism i magazynów. Można zauważyć, że w model biznesowy przemysłu mediowego niemal od początku wpisany jest czynnik reklamowy. Gazety z jednej strony zawsze były przestrzenią publikacyjną, z drugiej zaś platformą

³²⁴ Cernat, M., *The Ethics of financial journalism. A critical perspective*, „Cogito-Multidisciplinary Research Journal”, 2014, nr VI/3, s. 61.

³²⁵ Tambini, D., *What are financial...*, op.cit., s. 169.

³²⁶ Butterick, K. J., *A Churnalism...*, op.cit., s. 79.

³²⁷ Tambini, D., *What are financial...*, op.cit., s. 165.

³²⁸ Butterick, K. J., *A Churnalism...*, op.cit., s. 69.; Zob. także. Strauß, N. *Financial journalism in today's high-frequency news and information era*, „Journalism”, 2019, nr 20/2.

reklamodawców, którzy stanowili jedno z kluczowych źródeł ich finansowania. W obliczu rozwoju Internetu spadające nakłady tradycyjnych gazet zmniejszyły ich atrakcyjność reklamową, co nieuchronnie musiało wpłynąć na poziom przychodów. W odpowiedzi na to zjawisko, dominująca część gazet, o ile nie przeniosły się w całości do przestrzeni internetowej, prowadzi działalność publikacyjną zarówno w formie drukowanej, jak i elektronicznej. Należy jednak zauważyć, że obie te płaszczyzny stawiają przed redakcjami inne wymagania pod względem technicznym, przenoszą je do innego środowiska konkurencyjnego i nakładają inne obwarowania czasowe. W efekcie dziennikarze obciążeni są koniecznością odpowiadania na podwójne wymagania, co dodatkowo zwiększa ich obciążenie³²⁹.

Nadie Strauß sugeruje, że pojawienie się wiadomości online w dziennikarstwie finansowym złożyło się w czasie z odstępowaniem od aktywności zawodowej bardziej doświadczonych dziennikarzy starszego pokolenia, jak również wzrostem liczby dziennikarzy niezależnych, zajmujących się tematyką finansową, co związane było z procesami redukcji etatów w redakcjach czasopism. Wprawdzie na rynku wciąż dostępne są serwisy informacyjne, które charakteryzują się wysokimi standardami etycznymi, kładąc nacisk na jakość publikowanych treści, ale obawy dotyczące wpływu wskazywanych mechanizmów na rzetelność, wartość, etykę i dokładność publikowanych treści, nie są pozbawione podstaw. Wzrost liczby blogów o tematyce finansowej oraz rozpowszechnianych w Internecie alternatywnych źródeł informacji, mogą te obawy wręcz potwierdzać³³⁰. Konkludując, stwierdzić można, że sposób pracy dziennikarzy finansowych – w konsekwencji przemian społecznych i gospodarczych w ostatnich latach – ewoluował, czego przykładem jest zwiększone tempo ich pracy, konieczność odpowiadania na specyfikę różnych kanałów komunikacyjnych, co nie zawsze idzie w parze z możliwością docierania do różnorodnych źródeł informacji. Dodać można, że złożoność rynku finansowego wymaga jednak coraz bardziej specjalistycznej wiedzy i wysokich kompetencji analitycznych³³¹.

³²⁹ Butterick, K. J., *A Churnalism...*, op.cit., s. 68–69.

³³⁰ Strauß, N., *Financial journalism...*, op.cit., s. 288.

³³¹ Tambini, D., *What are financial...*, op.cit., s. 166.

3.3. Internalizowane czynniki selekcji informacji

Drugą grupą czynników mających wpływ na mechanizm selekcji informacji, są te, które uznać można za zewnętrzne wobec dziennikarza, ale potencjalnie przez niego internalizowane w procesie edukacji zawodowej i zdobywania doświadczenia. W kontekście dziennikarstwa finansowego zwraca się szczególną uwagę na trzy tego typu czynniki, tj. etykę zawodową, rutynę związaną z wypełnianiem obowiązków oraz niezbędną dziennikarzowi wiedzę³³². Należy zwrócić uwagę, że efektem oddziaływania tych czynników może być zbudowanie w świadomości dziennikarza zestawu tak zwanych wartości informacyjnych (ang. *news values*), czyli właściwych każdemu wydarzeniu cech, które stanowiąc będą przesłankę, skłaniającą autora lub wydawcę do opublikowania informacji.

Jednym z czynników mogących mieć wpływ na treść informacji, jak również na to, czy w ogóle zostaną one opublikowane, jest rutyna. Jej właściwością jest silne zakorzenienie w kulturze organizacyjnej. Kontekst organizacji odnosi się bowiem do reguł, konwencji oraz różnego typu procedur, związanych z codzienną pracą, warunkujących funkcjonowanie mediów w ogóle. Rutyna w pracy dziennikarskiej wydaje się odciskać swoje piętno w dwóch obszarach. Po pierwsze, wysuwa się ona na pierwszy plan podczas decydowania, które elementy przekazu powinny wyjść z kanału informacyjnego, a które z nich powinny zostać odrzucone. Proces selekcji jako taki składa się bowiem z szeregu procedur, a co za tym idzie, jego charakter jest silnie zrutynizowany. Jeśli na dziennikarza spojrzymy jako na uczestnika tego procesu, to nie pozbawiając jego twórczości znamion indywidualizmu, zauważyć można, że do jakiegoś stopnia jest ona zawsze efektem wykonywania zestawu schematycznych procedur³³³. Tym samym dostrzec można drugi obszar pracy dziennikarza, w jakim przejawia się kwestia rutyny – tj. jej przyswajanie i internalizacja. Dziennikarz, tak jak pracownik każdej innej organizacji, nabywa wiedzę o sposobie pracy w organizacji w procesie przekazywania jej przez bardziej doświadczonych pracowników. Zdaniem Gary’ego

³³² Merrill, G. J., *The revolution must wait: Economic, business, and financial journalism beyond the 2008 crisis*. „Ethical Space: The International Journal of Communication Ethics”, 2012, nr 9/1, s. 43.

Zob. także. Merrill, G. J., *The Political Content of British Economic, Business and Financial Journalism*, Palgrave Macmillan, Cham, 2019, s. 10.

³³³ Shoemaker, P.J., Vos, T., *Gatekeeping Theory*, Routledge, Nowy Jork 2009, s. 51.

Merrillina ów niewypowiedziany redakcyjny *modus operandi*, może okazać się na tyle silny, że większość osób z trudnością przeciwstawi się dominującej kulturze organizacyjnej, pośrednio przyjmując ją za swoją³³⁴. W oczywisty sposób tak pojęta rutyna przejawiać może się na poziomie językowym. Redakcje czasopism wyróżniają się jednak nie tylko zakresem swoich zainteresowań tematycznych, ale również właściwym sobie stylem, wyrażającym się w języku, czy nawet konkretnych, powtarzających się sformułowaniach. Stanowić to może przesłankę do postawienia hipotezy, że to samo wydarzenie, przekształćcane w procesie wytwarzania informacji, będzie opisane w inny sposób w zależności od konkretnego medium³³⁵.

3.3.1. Etyka

Kolejny, analizowany w literaturze przedmiotu czynnik, to standardy etyczne, które obecnie normowane są przez szereg kodeksów. Część z nich znajduje odzwierciedlenie w skodyfikowanych przepisach prawa, natomiast część przyjmuje postać zalecanych praktyk. W ich ramach najczęściej powtarzającymi się pryncypiami są: (1) prawdziwość przedstawianych informacji; (2) ich klarowność; (3) stanie na straży praw publicznych; (4) odpowiedzialność za kształtowanie opinii publicznej; (5) podtrzymywanie standardów gromadzenia i prezentowania informacji; (6) poszanowanie integralności źródeł informacji³³⁶. Wydaje się więc, że w ramach literatury przedmiotu panuje konsensus związany z definiowaniem obowiązków etycznych dziennikarzy finansowych w kontekście ich odpowiedzialności względem odbiorców informacji. Z drugiej jednak strony, o czym autorka już wspominała, postrzeganie roli dziennikarzy w przestrzeni nadzorczej nie zawsze idzie w parze z ich wyobrażeniem na temat wykonywanego zawodu. Damian Tambini wskazuje, że paradoksalnie tam, gdzie dziennikarze definiują siebie przez pryzmat dostarczycieli istotnych dla inwestorów informacji, tam mniej odczuwalne staje się poczucie odgrywania roli w systemie ładu korporacyjnego: „ten funkcjonalny, systemowy pogląd na rolę dziennikarzy finansowych może zostać odrzucony przez dziennikarzy, którzy powołują się na wąskie lub rynkowe pojęcie swoich obowiązków. [...] Ci etyczni minimaliści uważali, że ich ostateczną

³³⁴ Merrill, G. J., *The revolution must...*, op.cit., s. 48.

³³⁵ Butterick, K. J., *A Churnalism...*, op.cit., s.76.

³³⁶ McQuail, D., *McQuail's Mass Communication Theory*. SAGE Publications. Londyn, 2010. s.231. [za Laitila, 1995].

odpowiedzialnością jest szanowanie prawa i służenie akcjonariuszom ich spółek, a nie wypełnianie luk w systemie nadzoru korporacyjnego”³³⁷.

Powyższe stwierdzenie wskazuje na jeden z kluczowych problemów etycznych, jakim jest konflikt interesów. Konflikt ten uwydatnia się w dwóch obszarach. Pierwszy dotyczy relacji między dziennikarzami a źródłami informacji. Skuteczność pracy analitycznej jest uzależniona od dostępu do danych. Michael Borden zwraca uwagę, że jeśli opinie dziennikarzy działać będą sprzecznie z interesami danej korporacji, ryzykują oni pogorszeniem relacji z ich przedstawicielami i potencjalnym ograniczeniem dostępu do wiedzy na temat spółek, która wykraczałaby poza tę zawartą w sprawozdaniach finansowych lub innych oficjalnych dokumentach³³⁸. Jest to jeden z powodów, dla których media mogą wykazywać mniej negatywny ton narracji, kiedy *de facto* opisywany przez nich sektor czy spółka przeżywa kryzys³³⁹.

Pokłosiem tego problemu jest zauważalne przesunięcie znaczenia usług *public relations* do obszaru doradztwa strategicznego. Kontrolując przekaz wydostający się poza mury spółek, tego typu organizacje zaczęły stanowić „punkty dostępu” do informacji. Zarządy spółek są z oczywistych względów zainteresowane informacjami publikowanymi na temat prowadzonych przez nich przedsiębiorstw. O ile ich bezpośredni wpływ na treść przedstawianych przez dziennikarzy wiadomości jest niewielki, to wciąż mogą pośrednio pomóc w kształtowaniu opinii i odczytywaniu wydarzeń w oczekiwany przez siebie sposób, zapewniając im odpowiednie tło, przedstawiając informacje poparte wizualizacjami i propagując ich bardziej neutralną wizję niż wskazywałyby na to fakty³⁴⁰. W konsekwencji, swobodny przepływ informacji zostaje niejako przerwany, a sama informacja, którą otrzymuje dziennikarz EBF może być w pewien sposób wcześniej „wyjałowiona”³⁴¹.

Druga płaszczyzna, gdzie dostrzec można występujący konflikt interesów, to relacja między dziennikarzami finansowymi a inwestorami. Chodzi tu przede

³³⁷ Tambini, D., *What are financial...*, op.cit., s. 160–161.

³³⁸ Borden, M. J., *The Role of Financial Reporting in Corporate Governance*. „Fordham Journal of Corporate & Financial Law”, 2007, nr XII, s. 367.

³³⁹ *Media Coverage of Banking and Financial News*, University of Oxford. Reuters Institute for the Study of Journalism, 2014, s. 20.

³⁴⁰ Ibidem, s. 28.

³⁴¹ Zob. Butterick, K. J., *A Churnalism*, op.cit.; [za: Tambini, 2010: 22].

wszystkim o ewentualność, w której autor publikacji posiada papiery wartościowe analizowanej przez siebie spółki. Prowadzić może to do – znanej w literaturze przedmiotu – pokusy nadużycia. Handel na podstawie nieupublicznionej przez dziennikarza informacji lub upublicznienie jej w nieodpowiadający rzeczywistości sposób, bliski jest problemowi *insider tradingu* i naturalnie odbiega od standardów etycznych³⁴². Co więcej, sprzeczne jest z zasadami prezentowanymi w kodeksach polskich organizacji dziennikarskich. Przykładowo Stowarzyszenie Dziennikarzy Polskich w pkt. 19. i 20. Kodeksu Etyki Stowarzyszenia Dziennikarzy Polskich stwierdza, że „jakakolwiek kryptoreklama lub zatajanie informacji w celu uzyskania własnych korzyści jest naganne oraz że „dziennikarzowi nie wolno wykorzystywać we własnym interesie nieujawnionych publicznie informacji uzyskanych w działalności zawodowej, zwłaszcza w dziennikarstwie finansowym i ekonomicznym”³⁴³. Podobnie, Izba Wydawców Prasy w Kodeksie Dobrych Praktyk Wydawców Prasy stwierdza, że wprowadzie materiały, których twórcami są agencje PR lub działy promocji, mogą zostać wykorzystane w publikacji, ale muszą zostać poddane krytycznej analizie i skonfrontowane z innymi źródłami³⁴⁴. Ponadto, w Polsce obowiązuje Ustawa z dnia 26 stycznia 1984 r. Prawo prasowe. Zgodnie z jej art. 12: „Dziennikarzowi nie wolno prowadzić ukrytej działalności reklamowej wiążącej się z uzyskaniem korzyści majątkowej bądź osobistej od osoby lub jednostki organizacyjnej zainteresowanej reklamą”³⁴⁵. Zdaniem Damiana Tambiniego większość uznanych redakcji prowadzi również wewnętrzną politykę, zgodnie z którą wszelkie, posiadane przez dziennikarzy papiery wartościowe, ujawniane są menedżerowi lub redaktorowi naczelnemu, którzy mogą monitorować i reagować na występujący konflikt interesów³⁴⁶.

³⁴² Zob. Justia U.S. Supreme Court Center: *Carpenter v. United States*, 484 U.S. 19 (1987). No. 86–422. <https://supreme.justia.com/cases/federal/us/484/19>, Dostęp: 20 lip. 2022.

Zob. Także. Bowers, S., *Three face jail over Mirror share fraud*,

<https://www.theguardian.com/media/2005/dec/08/pressandpublishing.mirror1>, Dostęp: 20 lip. 2022.

Zob. Także. BBC News, *City Slickers guilty of tip scam*, http://news.bbc.co.uk/2/hi/uk_news/4507448.stm, Dostęp: 20 lip. 2022.

³⁴³ Zob. Kodeks Etyki Stowarzyszenia Dziennikarzy Polskich. Stowarzyszenie Dziennikarzy Polskich.

³⁴⁴ Zob. Kodeks Dobrych Praktyk Wydawców Prasy. Izba Wydawców Prasy.

³⁴⁵ Ustawa z dnia 26 stycznia 1984 r. Prawo prasowe. (Dz. U. 1984 Nr 5 poz. 24).

³⁴⁶ Tambini, D., *What are financial...*, op.cit., s. 162–163.

3.3.2. Wartości informacyjne

Jednym z zagadnień, stojącym niemalże u podstaw teorii selekcji informacji, są tak zwane wartości informacyjne (ang. *news values*). W tych bowiem wartościach badacze upatrują przyczyn podejmowanych decyzji, związanych z oceną wydarzenia jako wartego przekształcenia w treść publikacyjną (ang. *newsworthy*)³⁴⁷. W szerokim rozumieniu wartości informacyjne określić można jako głębokie struktury lub mapy kulturowe, z których korzystają dziennikarze, aby móc zrozumieć świat³⁴⁸. Marek Palczewski definiuje wartości informacyjne w następujący sposób: „*News values* dostarczają kryteriów w rutynowej praktyce dziennikarskiej, umożliwiając podejmowanie decyzji. Mimo że nieskodyfikowane, są jednak szeroko podzielane wśród dziennikarzy i pracowników mass mediów. Kształtują profesjonalne zasady, wpływają na socjalizację, praktykę i ideologię ludzi mediów”³⁴⁹. Innymi słowy, ich wpływ na aktywność publikacyjną wynika z tego, że istnieją w codziennej praktyce i są elementem wiedzy zdobywanej przez dziennikarzy podczas pracy³⁵⁰. Wartość informacyjna ma wpisany w siebie teleologiczny charakter, to ona wszak „czyni zdarzenie materiałem interesującym dla odbiorcy”³⁵¹.

W latach 60. Johan Galtung i Mari Ruge wskazali dwanaście czynników, których występowanie zwiększało prawdopodobieństwo przekształcenia wydarzenia w news. Były to: (1) częstotliwość; (2) przekroczenie progu³⁵²; (3) jednoznaczność; (4) podobieństwo kulturowe; (5) współbrzmienie³⁵³; (6) nieoczekiwanie; (7) ciągłość³⁵⁴; (8)

³⁴⁷ Palczewski, M., *News values w amerykańskich i europejskich teoriach selekcji informacji* [w:] red. nauk. Graszewicz, M.; Jastrzębski, J. *Teorie komunikacji i mediów*, Oficyna Wydawnicza ATUT, Wrocław 2009, s. 200.

³⁴⁸ Harcup, T., O'Neill, D., *What Is News? Galtung and Ruge revisited*. „Journalism Studies”, 2010, nr 2/2, s. 265; za: Hall, S., Critcher, Ch., Jefferson, T., Clarke, J., Roberts, B., *Policing the Crisis: Mugging, the State, and Law and Order*, Londyn 1978, Macmillan.

³⁴⁹ Palczewski, M., *News values...*, op.cit., s. 209.

³⁵⁰ Harcup, T., O'Neill, D., *What Is News...*, op.cit. s. 261.

³⁵¹ McQuail, D., *Teoria komunikowania masowego*, Wydawnictwo Naukowe PWN, Warszawa 2007, s. 311.

³⁵² Przekroczenie progu: zdarzenia muszą przekroczyć próg, zanim zostaną w ogóle zarejestrowane. Im większa intensywność, tym większy wpływ na postrzeganie osób odpowiedzialnych za wiadomości.

³⁵³ Współbrzmienie: selektor wiadomości może przewidzieć, że coś się wydarzy, tworząc w ten sposób mentalny „obraz” wydarzenia, co z kolei zwiększa jego szansę na stanie się wiadomością.

³⁵⁴ Ciągłość: gdy wydarzenie stanie się nagłówkiem wiadomości, przez jakiś czas pozostaje w centrum uwagi mediów – nawet jeśli jego amplituda została znacznie zmniejszona – ponieważ stało się znane i łatwiejsze do interpretacji. Kontynuacja relacji ma również na celu usprawiedliwienie uwagi przyciągniętego wydarzenia.

kompozycja³⁵⁵; (9) odniesienie do elitarnych narodów; (10) odniesienie do ludzi z elit; (11) odniesienie do osób; (12) odniesienie do czegoś negatywnego³⁵⁶.

Gaultung i Ruge wskazali w swoim badaniu trzy hipotezy. Po pierwsze, stwierdzili, że im bardziej wydarzenie spełnia wymienione kryteria, tym większe jest prawdopodobieństwo, że zostanie ono utrwalone w postaci newsa. Po drugie, gdy dana cecha wydarzenia została wyselekcjonowana, co sprawiło, że wydarzenie to mogło zostać opublikowane, cecha ta zostanie zaakcentowana (co określane jest mianem błędu zniekształcenia). Po trzecie, zarówno proces selekcji, jak i proces zniekształcenia będą odbywać się na wszystkich etapach łańcucha od zdarzenia do czytelnika (proces ten określony został mianem replikacji)³⁵⁷. Lista stworzona przez Gaultung i Ruge, choć częstokroć krytykowana, dała początek rozważaniom nad czynnikami „newsotwórczymi”. W późniejszych latach była ona niejednokrotnie uzupełniana i przekształcana. Przykładowo w 2001 roku Tom Hurcup i Deirde O’Neill zaproponowali następujące wartości informacyjne: (1) elita władzy; (2) sława; (3) rozrywka; (4) niespodzianki; (5) złe wiadomości; (6) dobre wiadomości; (7) wielkość³⁵⁸; (8) istotność; (9) kontynuacja³⁵⁹; (10) agenda gazety³⁶⁰.

Prowadzone dotychczas badania w zakresie produkcji informacji sektora EBF ujawniły występujące w nich, powtarzające się wzorce. Stanowiąc to może przesłankę do stwierdzenia, że również w tym typie dziennikarstwa mechanizm selekcji nie jest kwestią przypadkową, a odbywa się według pewnych nieskodyfikowanych kryteriów³⁶¹. Zauważyć można jednak, że w literaturze przedmiotu odznaczają się w tej sprawie dwa sprzeczne podejścia. Część autorów zakłada, że media EBF kierują się analogicznymi

³⁵⁵ Kompozycja: wydarzenie może zostać uwzględnione jako wiadomość nie tyle ze względu na jego wartość wewnętrzną, ile dlatego, że pasuje do ogólnej agendy medium.

³⁵⁶ Harcup, T., O’Neill, D., *What Is News...*, op.cit., s. 263.

³⁵⁷ Galtung, J., Ruge, M. H., *The Structure of Foreign News*. „Journal of Peace Research”, 1965, nr 2/1, s. 71.

³⁵⁸ Wielkość: historie, które są postrzegane jako wystarczająco istotne, zarówno pod względem liczby zaangażowanych osób, jak i potencjalnego wpływu.

³⁵⁹ Kontynuacja: opowieści na tematy, które już wcześniej pojawiły się w wiadomościach.

³⁶⁰ Agenda gazety: historie, które podtrzymują agendę informacyjną gazety. Harcup, T., O’Neill, D., *What Is News...*, op.cit., s. 279.

³⁶¹ Butterick, K. J., *A Churnalism...*, op.cit., s. 73.

wartościami informacyjnymi, co pozostałe rodzaje mediów³⁶². Natomiast inni autorzy zwracają uwagę, że jednoznaczne przełożenie powszechnych wartości informacyjnych na dziennikarstwo EBF nie jest możliwe. Sytuacja bowiem komplikuje się, kiedy zauważymy, że wiadomości o charakterze ekonomicznym pojawiają się zarówno w przekazach głównego nurtu, jak i specjalistycznych.

Zdaniem Gillian Doyle, właśnie to rozróżnienie determinować może wybór i decyzje o publikacji treści wiadomości. Sugeruje ona bowiem, że dziennikarze pracujący w mediach głównego nurtu obarczeni są koniecznością odpowiadania na oczekiwania szerokiej publiczności. Wybierając i opisując temat, muszą jednocześnie zastanowić się, czy będzie on zrozumiały dla odbiorcy o dużo mniejszym poziomie zaangażowania niż inwestorzy. Odpowiada więc on na pytanie, czy historia może przyciągnąć niespecjalistyczną publiczność. Dostrzec można, że tego typu media koncentrują się przede wszystkim na spółkach o dobrze rozpoznawalnych markach oraz na kilku tematach, które mogą być zrozumiałe i wzbudzić zainteresowanie masowego odbiorcy. W przypadku mediów koncentrujących się na dostarczaniu informacji inwestorom, zrozumienie tego „co jest warte opublikowania” wydaje się mieć zgoła inne fundamenty. Dzielone między dziennikarzami przeświadczenie o wartości informacyjnej opiera się na odpowiedzi na pytanie, czy dane zdarzenie może mieć wpływ na krajobraz inwestycyjny. Dziennikarze muszą mieć też na uwadze, że swoje artykuły kierują do bardziej wymagającego odbiorcy. W efekcie, poprzestawanie na powierzchownym relacjonowaniu zdarzeń czy komunikatów, może nie być wystarczające. W tym przypadku konieczne będzie zbudowanie szerszej koncepcji tekstu, która umożliwi na przykład wskazanie pojawiających się wzorców działań spółek, opisanie potencjalnych efektów aktualnych wydarzeń lub analizę nadzwyczajnych sytuacji, mogących mieć wpływ na ceny akcji³⁶³. Na konieczność rozróżnienia wartości informacyjnych wiadomości finansowych ze względu na typ mediów, uwagę zwracają również Mark Boukes i Rens Vliegthart. Klasyfikując gazety według klucza: (1) popularne; (2)

³⁶² Zob. Butterick, K. J., *A Churnalism...*, op.cit., s. 73; zob. także. Boukes, M., Vliegthart, R., *A general pattern in the construction of economic newsworthiness? Analyzing news factors in popular, quality, regional, and financial newspapers*, „Journalism”, 2020, nr 21/2, s. 279–300.

³⁶³ Doyle, G., *Financial news journalism. A post-Enron analysis of approaches towards economic and financial news production in the UK*, „Journalism”, 2006, nr 7/4, s. 448.

jakościowe; (3) regionalne; (4) finansowe, wskazują, że część z nich – podczas opisywania wydarzeń finansowych – kieruje się wartościami informacyjnymi analogicznymi do tych, funkcjonujących w ramach mediów głównego nurtu, a część kieruje się indywidualnymi czynnikami określającymi wartość informacyjną³⁶⁴.

Mimo to, analizując w tym zakresie literaturę przedmiotu, trudno natrafić na propozycje list wartości informacyjnych, które wywodziłyby się z badań nad tematyką dziennikarstwa EBF. Dlatego też autorka, kierując się logiką zaproponowaną przez wspomnianych autorów, proponuje przyjąć za punkt wyjścia popularne wartości informacyjne przedstawiane w literaturze z zakresu dziennikarstwa i na ich tle wykazać specyfikę sektora finansowego. Wartościami tymi będą: personifikacja; negatywizm, elitarność, potencjalne znaczenie (wpływ), kontrowersja, bliskość geograficzna i ciągłość³⁶⁵.

Personifikacja

Tony Harcup i Deirdre O'Neill, analizując problem skłonności do personifikacji, stwierdzają: „wiadomości mają tendencję do przedstawiania wydarzeń jako działań nazwanych ludzi, a nie jako wynik sił społecznych. Personifikacja wykracza poza historie o *ludzkim interesie* i może odnosić się do kulturowego idealizmu, zgodnie z którym człowiek jest panem własnego losu, a wydarzenia mogą być postrzegane jako wynik aktu wolnej woli”³⁶⁶. U podstaw takiej skłonności leżeć może kilka czynników. Po pierwsze, wydaje się, że skoncentrowanie się na ludziach jest łatwiejszym sposobem na przyciągnięcie czytelników niż koncentrowanie się na opisie skomplikowanych trendów technologicznych, strukturze branży lub objaśnianiu mechanizmów przyczynowo-skutkowych³⁶⁷. Po drugie, skupienie się przykładowo na dyrektorsze generalnym buduje przeświadczenie, że to on jest odpowiedzialny za losy firmy. Oczywiście zniekształca to obraz rzeczywistości, sugerując, że w przypadku słabych wyników, usunięcie prezesa zarządu bezpośrednio przełoży się na poprawę efektów działalności spółki. Strukturalne

³⁶⁴ Zob. Boukes, M., Vliegthart, R., *A general pattern...*, op.cit., s. 279–300.

³⁶⁵ Boukes, M., Vliegthart, R., *A general pattern...*, op.cit., s. 282; [za: Eilders C, 2006: *News factors and news decisions. Theoretical and methodological advances in Germany*, „Communications” 31(1): 5–24].

³⁶⁶ Harcup, T., O'Neill, D., *What Is News...*, op.cit., s. 263.

³⁶⁷ Doyle, G., *Financial news journalism. A post-Enron analysis of approaches towards economic and financial news production in the UK*, „Journalism”, 2006, nr 7/4, s. 437.

(a więc bardziej skomplikowane) problemy firmy, mogą w tej narracji nie mieć już tak wielkiego znaczenia.

W tym kontekście interesujące badania przeprowadziło w 2008 roku FTI Consulting. Stwierdzono w nich, że reputacja dyrektora generalnego jest kluczowym czynnikiem, podczas podejmowania przez inwestorów decyzji o zakupie lub sprzedaży akcji spółek, a blisko jedna trzecia decyzji inwestycyjnych podejmowana jest na podstawie opinii prezesów zarządów. Co ciekawe, to samo badanie wykazało, że media mają relatywnie ograniczony wpływ na inwestorów, gdyż tylko 27% ankietowanych stwierdziło, że podjęło decyzję inwestycyjną na podstawie rekomendacji mediów. Jak na ironię, poglądy inwestorów na temat kompetencji i zdolności prezesów zarządów kształtowane są właśnie przez media³⁶⁸. Badania przeprowadzone przez Marka Boukesa oraz Rens Vliegthart wykazały, że obecność czynnika newsotwórczego w postaci personifikacji w informacjach finansowych jest właściwa gazetom popularnym oraz tym o zasięgu regionalnym. W dużo mniejszym stopniu czynnik ten widoczny był w przypadku gazet finansowych oraz jakościowych, przy czym te ostatnie koncentrowały się na personifikacji 1,5 raz częściej niż gazety finansowe³⁶⁹.

Elitarność

Inną wartością informacyjną, analizowaną w ramach literatury przedmiotu, jest tak zwana obecność elit. Anglojęzyczne sformułowanie *elitness* wydaje się jednak lepiej oddawać istotę sprawy. W zależności od badaczy wartość ta obejmuje: status, władzę, czynniki atrybucji, celebrytów, rozgłos, zasługi lub po prostu elitarność. Tym samym określenie to stosowane jest do podmiotów o niejednorodnym charakterze, takich jak: różnego typu instytucje, wydarzenia, czy też ludzie³⁷⁰. W najprostszym ujęciu chodzi o obecność w treści informacji osób lub innego typu podmiotów o dużej sile oddziaływania społecznego³⁷¹. W codziennej praktyce dziennikarstwa biznesowego może mieć to związek, chociażby z wielkością firm, o których dziennikarze zdecydują się

³⁶⁸ Butterick, K. J., *A Churnalism...*, op.cit., s.74.

³⁶⁹ Boukes, M., Vliegthart, R., *A general pattern...*, op.cit., s. 290–295.

³⁷⁰ Zob. Bednarek, M., Caple, H., *The Discourse of News Values. How news organizations create newsworthiness*, Oxford University Press, Nowy Jork 2017, s. 58.

³⁷¹ Boukes, M., Vliegthart, R., *A General Pattern...*, op.cit., s. 282; [za: Eilders C., 2006: *News factors and news decisions. Theoretical and methodological advances in Germany*, „Communications” 31(1): 5–24].

pisać. Przykładowo, Kennath R. Ahern i Danis Sosyura przeprowadzili badania, w których sprawdzali zasady publikowania informacji dotyczących fuzji i przejęć. Wyniki wskazywały na istnienie pewnej specyfiki, dotyczącej wielkości firm mierzonej ich wartością księgową. Okazało się, że w przypadku największych spółek dziennikarze skłonni byli przekazywać informacje na temat tego typu zdarzeń, wskazując, że rozmowy na temat fuzji są np. w fazie wstępnej. Innymi słowy, chociaż dziennikarze próbowali wskazać czytelnikom, że informacja ma charakter raczej spekulatywny, to i tak wielkość firmy decydowała o ich wysokiej wartości informacyjnej. Autorzy badania wprost nazywają takie treści plotkami, ponieważ dowiedli, że w rzeczywistości dokładność przewidywań co do opisywanych zdarzeń, była relatywnie niewielka³⁷².

Istnieje jednak pewne uzasadnienie dla skupiania się przez dziennikarzy na dużych przedsiębiorstwach. Takie podmioty, na przykład poprzez działy *public relations*, mogą włożyć więcej zasobów i wysiłku w komunikację medialną. Zdaniem Keitha J. Buttericka w obecnych czasach dziennikarze są w pewnym sensie uzależnieni od tego typu źródeł danych, a sytuacja ta jest tak samo błogosławieństwem, jak i przekleństwem. Z jednej strony w oczywisty sposób ułatwia to wykonywanie zadań dziennikarskich. Z drugiej jednak rodzi również pytania o zasadność artykułu, który w dużej mierze pochodzi ze źródeł korporacyjnych, a zatem prawdopodobnie zawiera pogląd, który firma chce promować³⁷³. Problem ten zostanie szczerzej opisany w dalszej części rozdziału.

Kontrowersje i ciągłość

Jak wynika z badania Aherna i Sosyura w przypadku dużych spółek dziennikarze są bardziej skłonni upubliczniać informacje o niskim stopniu pewności. Może mieć to również związek z inną wartością informacyjną, która w literaturze przedmiotu definiowana jest jako kontrowersja. W praktyce pod hasłem tym kryje się próba opisywania lub sugerowania istnienia jakiejś niezgody między – istotnymi z punktu widzenia rynku – podmiotami, takimi jak organizacje, partie polityczne, czy też osoby³⁷⁴. Czynniki ten wydaje się na tyle zasadniczy, że Mark Boukes, Natalie Jones i Rens Vliegthart określają go jako jedną z dwóch, obok elitarności, najbardziej wpływowych

³⁷² Ahern, K. R., Sosyura, D., *Rumor Has It: Sensationalism in Financial Media*. „The Review of Financial Studies”, 2015, nr 28/7, s. 2088.

³⁷³ Butterick, K. J., *A Churnalism...*, op.cit., s. 65.

³⁷⁴ Por. Boukes, M., Vliegthart, R., *A general pattern...*, op.cit., s. 289.

wartości informacyjnych w mediach³⁷⁵. Z drugiej jednak strony, duże organizacje posiadają zasoby, które – mniej lub bardziej skutecznie – hamować mogą niekontrolowane wydostawanie się wiedzy o istnieniu konfliktów poza mury przedsiębiorstwa³⁷⁶.

Reakcja inwestorów na tego typu informacje może mieć związek z charakterystyką gazety, w jakiej się ona pojawia. Ahern i Sosyur wskazują, że informacja spekulacyjna, pojawiająca się w starszych gazetach o większych nakładach, generuje wyższe stopy zwrotu w dniu zerowym, niż ta sama informacja opublikowana w młodszej gazecie o niższym nakładzie³⁷⁷. Oznaczać to może, że w krótkim terminie relacja, jaka zachodzi między informacją a ceną aktywów, może różnić się w zależności od tego, w jakiej gazecie jest ona publikowana³⁷⁸. Z kolei Boukes i Vliegenthart, podejmując się przekrojowej analizy sposobów konstruowania informacji ekonomicznych (zgodnie z typologią gazet), zauważyli obecność czynnika kontrowersji w próbie gazet regionalnych (35,1%), dzienników opiniotwórczych (34,8%) oraz popularnych (32,9%). Ponownie jednak wyraźną różnicę pod tym względem można było zauważyć w przypadku gazet finansowych (24,5%)³⁷⁹. Wydaje się, że taki stan rzeczy może mieć kilka przyczyn.

Po pierwsze, należy mieć na uwadze, że upublicznianie kontrowersyjnych informacji niesie za sobą szereg konsekwencji. Informacje takie mogą mieć bowiem wpływ na wycenę aktywów. Jeśli mają one charakter spekulacyjny lub powodowane są raczej pobudkami sprzedażowymi gazety, a nie dokładnością wersyfikacyjną, to zaklasyfikować można je raczej do kategorii szumu informacyjnego. W efekcie, ewentualna dynamika cen akcji opisywanej w ten sposób spółki – nawet krótkotrwała – daleka będzie od oceny jej wartości fundamentalnej³⁸⁰. Po drugie, konsekwencją pogoni za kontrowersją jest również, opisywane przez Roberta Shillera zjawisko „przeciążenia rekordowego” (ang. *record overload*). Chodzi tu o budowanie wrażenia nieustannego

³⁷⁵ Zob. Boukes, M., Jones, N. P., Vliegenthart, R., *Newsworthiness and story prominence: How the presence of news factors relates to upfront position and length of news stories*, „Journalism”, 2020, s. 15.

³⁷⁶ Manning, P., *Financial journalism, news sources and the banking crisis*, „Journalism”, 2012, s. 181–182.

³⁷⁷ Ahern, K. R., Sosyura, D., *Rumor Has It...*, op.cit., s. 2074.

³⁷⁸ Ibidem, s. 2054.

³⁷⁹ Boukes, M., Vliegenthart, R., *A general pattern...*, op.cit., s. 292.

³⁸⁰ Ahern, K. R., Sosyura, D., *Rumor Has It...*, op.cit., s. 2054.

ustanawiania nowych i znaczących rekordów, co tak naprawdę pogłębia tylko odczucie związanego z gospodarką zamętu oraz utrudnia rozpoznanie przez odbiorców sytuacji, w których dzieje się coś naprawdę istotnego lub nowego³⁸¹.

Można podać jednak w wątpliwość czy dziennikarze gazet finansowych, których targetem są inwestorzy, chcą być kojarzeni z szumem informacyjnym. Taka konotacja odebrać mogłaby im prawo do alarmowania inwestorów w sytuacjach, które faktycznie będą tego wymagać. Dodatkowo merytoryczna interpretacja konkretnego konfliktu wymaga przestrzeni. Taką przestrzeń zapewnić mogą czasopisma opiniotwórcze. Dziennikarze gazet finansowych zakładać mogą, że ich odbiorcy zorientowani są w poruszanej tematyce. Być może więc – dla tej grupy dziennikarzy – nie tyle kontrowersja jest kluczową wartością informacyjną, ile jest nią ciągłość³⁸². Jak twierdzi Shiller „media informacyjne są naturalnie zainteresowane rynkami finansowymi, ponieważ rynki przynajmniej dostarczają ciągłych wiadomości w postaci codziennych zmian cen. Nic nie przebije giełdy z powodu samej częstotliwości potencjalnie interesujących wiadomości³⁸³”.

Konsekwencje ekonomiczne

To stwierdzenie prowadzi do kolejnego pytania, o wartość informacyjną horyzontów czasowych wiadomości finansowych lub – innymi słowy – o wartość informacyjną, którą badacze nazywają aktualnością. Kluczowe pytanie dotyczy tutaj tego, w jaki sposób dane wydarzenie definiowane jest jako czasowo istotne z punktu widzenia odbiorcy. W tym kontekście zdarzenia klasyfikowane mogą być jako niedawne, nowe, wciąż trwające (informacyjna wartość natychmiastowości) lub mające miejsce w najbliższej przyszłości (wartość nieuchronności). W szerokim horyzoncie czasowym wydarzenie może zostać skadowane również jako odnoszące się do aktualnych trendów lub wydarzeń o charakterze sezonowym³⁸⁴. Wydaje się, że dziennikarze domyślnie operują nadawaniem wagi poszczególnym wydarzeniom zgodnie z ich typologią czasową. Z tego punktu widzenia wydaje się to rzeczą naturalną – czas niemalże

³⁸¹ Shiller, R. J., *Irrational Exuberance*, Princeton University Press, New Jersey 2015, s. 104–105.

³⁸² Por. Boukes, M., Jones, N. P., Vliegthart, R., *Newsworthiness...*, op.cit, s. 15.

³⁸³ Shiller, R.J., *Irrational Exuberance...*, op.cit., s. 102.

³⁸⁴ Bednarek, M., Caple, H., *The Discourse of News Values. How news organizations create newsworthiness*, Oxford University Press, Nowy Jork 2017, s. 65.

wbudowany jest w definicję wiadomości. Obecna możliwość publikowania wiadomości w formie online nie tyle skróciła okres, który upłynąć musi od momentu zajścia wydarzenia do momentu opublikowania materiału, ile właściwie pokonała barierę czasu³⁸⁵.

Ogólne prawidłowości związane z tą wartością informacyjną oparte są na zasadzie bliskości (ang. *news value of proximity*). Oznacza to, że im bliżej czasu publikacji znajduje się opisywany punkt odniesienia, tym bardziej aktualne, a w konsekwencji wartości opublikowania staje się samo wydarzenie³⁸⁶. Co więcej, jak zauważa Nikki Usher, dziennikarze finansowi rzadko koncentrują się na dostrzeganiu i opisywaniu systemowych problemów rynku. Może wynikać to z faktu, że zarówno oni, jak i organizacja tworząca ramy ich pracy, nadmiernie skupieni są na inwestorach, którzy z kolei wykazują największe zainteresowanie informacjami związanymi z bieżącymi wydarzeniami³⁸⁷. W kontekście badań prowadzonych w dyscyplinie finansów, stanowi to przesłankę do podejmowania badań dotyczących relacji zachodzących między publikacjami a zmianami cen walorów w krótkim terminie. Podążając za tokiem rozumowania Roberta Shillera, media finansowe niejako zmuszone są do dotrzymywania kroku częstotliwości wydarzeń rynkowych. Chociaż, zgodnie z wiedzą autorki, literatura przedmiotu nie dostarcza badań, które analizowałyby tę cechę przez pryzmat mediów związanych z tematyką finansową, to domniemywać można, że – również w tym przypadku – aktualność jest jedną z priorytetowych wartości informacyjnych.

Powyższa charakterystyka związana była z zagadnieniem bliskości w rozumieniu czasowym. Czynniki ten widoczny jest jednak również w innych kontekstach, np. geograficznym oraz kulturowym. Wydaje się, że dziennikarze kierują się nim z uwagi na wartość informacyjną określaną jako „wpływ i znaczenie” (ang. *influence and relevance*). W uproszczeniu powiedzieć można, że podstawowe pytanie, na jakie dziennikarz musi odpowiedzieć, brzmi: czy dane wydarzenie będzie miało swoje konsekwencje i kto te konsekwencje odczuje. Mówiąc inaczej – czy dla danej grupy

³⁸⁵ McQuail, D., *McQuail's...*, op.cit., s. 415.

³⁸⁶ Bednarek, M., Caple, H., *The Discourse...*, op.cit., s. 65.

³⁸⁷ Usher, N., *Making business news: A production analysis of The New York Times*, „International Journal of Communication”, 2017, nr 11, s. 367.

odbiorców, opisywana historia, będzie znacząca, w takim sensie, że będą oni mogli utożsamić się z jej implikacjami.

W literaturze przedmiotu opisywanych jest kilka kryteriów realizacji tej wartości. Christiane Eilders wymienia w tym kontekście dwie. Pierwsza odnosi się do liczby osób, na jakie dane wydarzenie mogło mieć wpływ. Chodzi tu o trafność, czy też zasięg danej historii. Druga związana jest z jakością owego wpływu. Chodzi bowiem o to, czy konotacje, jakie niesie ze sobą wydarzenie, mają dla tej grupy charakter szkody czy sukcesu³⁸⁸. W perspektywie ekonomicznej badacze wskazują na istnienie tzw. ramy konsekwencji ekonomicznych (ang. *economic consequences-frame*). W ten sposób opisywana jest sytuacja, w której artykuł jednoznacznie „odnosi się do ekonomicznego wpływu, jaki historia ma, miała lub może mieć na osoby, firmy, grupę, region lub kraj³⁸⁹”. Częstokroć ten potencjalny wpływ analizowany jest pod kątem geograficznej i kulturowej bliskości pomiędzy opisywanymi zdarzeniami a odbiorcami. Eilders zauważa, że wydarzeniom, które mają miejsce w pobliżu domu odbiorcy, przypisywana jest wyższa wartość konsekwencji ekonomicznych, niż wydarzeniom geograficznie odległym³⁹⁰.

Z punktu widzenia autorów informacji, takie reguły wydają się również logiczne. Opisywanie wydarzeń, mających miejsce w niewielkiej odległości, do których można mieć bezpośredni dostęp, lub relatywnie szybko móc dotrzeć do jego świadków, jest oczywiście łatwiejsze niż przejście tego samego procesu, w miejscu od niego oddalonym. Kennath R. Ahern i Danis Sosyura sugerują nawet, że nie tyle przestrzeń kraju jest tutaj decydująca, ile wręcz miasta. Ich zdaniem dziennikarze finansowi pracujący w Nowym Jorku są zwykle dokładniejsi niż ich koledzy, pracujący w innych częściach Stanów Zjednoczonych. Oczywiście może to mieć związek z przypuszczeniem, że do ulokowanych tam redakcji, trafiają najlepsi dziennikarze. Równie dobrze jednak wynikać

³⁸⁸ Eilders, Ch., *News factors and news decisions. Theoretical and methodological advances in Germany*. „Communications”, 2006, nr 31/3, s. 15.

³⁸⁹ Boukes, M., Vliegthart, R., *A General Pattern...*, op.cit., s. 289.; Por. Boukes, M., Jones, N. P., Vliegthart, R., *Newsworthiness...*, op.cit., s. 10.

³⁹⁰ Eilders, Ch., *News Factors...*, op.cit., s. 15.

może to z faktu, że mają oni lepsze kontakty z osobami aktywnymi na giełdzie na Wall Street³⁹¹.

Mimo to, wyniki badań Marka Boukesa i Rens Vliegenthart wskazują, że przy użyciu tej logiki nie można w pełni scharakteryzować mediów zajmujących się stricte tematyką finansową. Na tym tle przodują bowiem gazety regionalne (62%) oraz gazety popularne (61,7%). W dalszej kolejności wartość ta dotyczy dzienników opiniotwórczych (53,8%), a w najmniejszym stopniu związana jest ona z gazetami finansowymi (42,6%). Oznacza to, że w mediach specjalistycznych większość wiadomości dotyczyła wydarzeń zagranicznych³⁹². Próba objaśnienia tych wyników wciąż jednak odwołuje do odpowiedzi na pytanie: kto – zdaniem dziennikarzy – ponosi konsekwencje opisywanych wydarzeń. Wydaje się więc, że dziennikarze EBF nie tylko są świadomi, że głównymi odbiorcami ich wiadomości są inwestorzy, ale wręcz kultywują podtrzymywanie paradygmatu budowania wartości dla akcjonariuszy. Aeron Davis, analizując próbę 425 artykułów związanych z tematyką biznesową, wskazał, że słowa „akcjonariusz” oraz „inwestor” pojawiały się w 63,5% z nich, podczas gdy słowa „pracownik” i „konsument” – jedynie w 7,1% oraz w 4,2%³⁹³. Można więc domniemywać, że – tak jak u podstaw działań menedżerów stoi przekonanie o konieczności zwiększania wartości dla akcjonariuszy – tak u podstaw decyzji dziennikarskich, związanych z selekcją informacji, stać może przeświadczenie o potrzebie ciągłego odpowiadania na oczekiwania inwestorów³⁹⁴.

Taka świadomość, podczas odpowiadania na pytanie o możliwość utożsamienia się przez odbiorcę z danym problem, rodzić może kilka implikacji. Po pierwsze, biorąc pod uwagę tendencje związane globalizacją przepływu kapitału, założenie, że inwestorzy zainteresowani będą jedynie lokalnymi wydarzeniami gospodarczymi, czy też jedynie lokalnym rynkiem kapitałowym, może nie być do końca trafne. Dodatkowo należy zwrócić uwagę, że tendencja ta dotyczy nie tylko przepływów realizowanych poprzez instrumenty finansowe, ale również samych przedsiębiorstw. Jako przykład można podać

³⁹¹ Ahern, K. R., Sosyura, D., *Rumor Has It...*, op.cit., s. 2071–2072.

³⁹² Boukes, M., Vliegenthart, R., *A General Pattern...*, op.cit., s. 292.

³⁹³ Zob. Davis, A., *Public relations, business news and the reproduction of corporate elite power*, „Journalism”, 2000, nr 1/3, s. 296.

³⁹⁴ Por. Butterick, K. J., *A Churnalism...*, op.cit., s. 75.

tutaj spółki typu offshore³⁹⁵. Ten fakt nie tylko tworzy sposobność do inwestowania poza horyzontami własnego kraju, ale również uzbraja inwestorów w świadomość, że wydarzenia gospodarcze w jednym kraju, mogą wpłynąć na sytuację na rynkach w innym. Po drugie, w opisywanym kontekście uwypukla się znaczenie Internetu dla aktywności inwestycyjnej. Nie chodzi tu jednak jedynie o współczesny sposób zawierania transakcji, ale również o sposób pozyskiwania informacji, a może przede wszystkim o sposoby jej wymiany, a poprzez to budowania swoistej wspólnoty o charakterze kulturowym i językowym. Jak twierdzi bowiem Natalie Fenton: “Internet, opierając się na *sieci sieci*, może wspierać tożsamość zbiorową i wzmacniać solidarność. Innymi słowy, bierze udział w procesie konstruowania znaczeń. Charakter i zakres technologii wpływają nie tylko na sposób, w jaki ruch komunikuje swoje cele i zadania, ale także na jego skalę geograficzną, strukturę organizacyjną i tożsamość zbiorową”³⁹⁶.

Być może więc bastion ograniczeń geograficznych upada, by w to miejsce powstała nowa tożsamość inwestorów budowana na gruncie międzynarodowym. Nawet jeśli jednak powyższe, dość pompatyczne stwierdzenie, uznać za prawdziwe, to należy pamiętać, że w ramach czynnika bliskości, poruszany jest również problem bliskości kulturowej, która już wcześniej stała w sprzeczności z ograniczeniami geograficznymi. Nie zawsze wszak tekst o wydarzeniach ułożonych bliżej pod względem odległości był tym, który okazywał się dla czytelnika najbardziej interesujący. Jako przykład podać można tutaj spostrzeżenia Moniki Bednarek i Helen Capli dotyczące rynku mediowego Australii i Oceanii. Ich zdaniem pozornie przyjąć można, że w artykułach publikowanych w gazetach australijskich, relatywnie często pojawiać będą się treści dotyczące wydarzeń mających miejsce w Papui-Nowej Gwinei, Nowej Zelandii czy Indonezji. Niewątpliwie wszystkie te kraje charakteryzują się bliskością geograficzną. Okazuje się jednak, że bardziej prawdopodobne jest, że odbiorca wiadomości zainteresowany będzie wydarzeniami, mającymi miejsce w Wielkiej

³⁹⁵ Zob. Cameron, A., Palan, R., *The imagined economies of globalization*, SAGE Publications, Londyn 2004.

³⁹⁶ Kantola, A., *On the dark side of democracy: the global imaginary of financial journalism*, [w:] (red.) Cammaerts, E., Carpentier, N., *Reclaiming the Media. Communication Rights and Democratic Media Roles*, European Communication Research and Education Association, Bristol 2007, s. 236.

Brytanii, co oznacza, że przedkłada on bliskość kulturową nad bliskość w rozumieniu geograficznym³⁹⁷.

Powyższa problematyka, związana z odpowiadaniem na pytanie o to, czy dane wydarzenie może być odebrane przez czytelnika jako mające dla niego znaczenie lub niosące ze sobą konsekwencje ekonomiczne, jest widoczna wśród wszystkich typów gazet zajmujących się tematyką ekonomiczną. Jak wynika z badania Boukesa i Vliegentharta, analiza występowania wartości informacyjnej – zdefiniowanej przez nich jako „wpływ i znaczenie” – w kontekście tematyki gospodarczej, nie wykazała znaczących różnic między gazetami popularnymi, opiniotwórczymi, regionalnymi i finansowymi (odpowiednio: 78,4%, 78,3%, 79,8%, 76,0%)³⁹⁸. Należy jednak zaznaczyć, że ich rezultaty uzyskane były na podstawie odpowiedzi na pytanie: „Czy artykuł wyraźnie odnosi się do ekonomicznego wpływu”. W związku z tym stwierdzić należy, że dziennikarze w istocie mogą kierować się opisywaną zasadą, ale stwierdzenie to ma zastosowanie jedynie do sytuacji, w których fakt ten zaznaczony jest bezpośrednio w artykule. Innymi słowy, dziennikarze muszą być świadomi istnienia tej wartości albo jako immanentnej dla samego wydarzenia, albo równie świadomie kadrować wydarzenie na takie, które ową wartość posiada. Powyższe spostrzeżenie nie odnosi się więc do czynników, które – mogąc mieć wpływ na sposób doboru i konstruowania informacji przez dziennikarzy – nie występują na świadomym lub też wyrażonym *explicite* poziomie. Istnieją bowiem przesłanki sugerujące, że proces selekcji informacji związany jest nie tylko z regułami – wynikającymi z zasad funkcjonowania rynku prasowego czy organizacji, w jakiej zatrudniony jest dziennikarz – ale również z autonomicznie przyjmowanej przez dziennikarza logiki zachowania.

3.4. Wewnętrzne czynniki selekcji informacji

Trzecim, opisywanym w literaturze przedmiotu typem czynników mogących wpływać na kształt treści informacji oraz mechanizm jej selekcji, są czynniki związane z samym dziennikarzem. W tym kontekście kluczowe znaczenie mają uproszczenia poznawcze i emocje. Z punktu widzenia tematyki dysertacji czynniki te stanowią dla autorki

³⁹⁷ Zob. Bednarek, M., Caple, H., *The Discourse of News Values. How news organizations create newsworthiness*, Oxford University Press, Nowy Jork 2017, s. 63.

³⁹⁸ Boukes, M., Vliegenthart, R., *A General Pattern...*, op.cit., s. 292.

przedmiot szczególnego zainteresowania. Zarówno bowiem z teoretycznego punktu widzenia, jak i w wyniku prowadzonych w tej tematyce badań, domniemywać można, że nadawcy informacji pozostawiają w niej ślady własnego nastawienia względem opisywanego zdarzenia.

Zdaniem Leszka Tymiakina wszelkie treści językowe związane są triadą ekspresji, emocji oraz wartościowania. Ekspresja dotyczy pragmatycznej funkcji języka. Opiera się na możliwości wyrażania przeżyć. Emocje, w kontekście komunikacyjnym, zakorzenione są w odpowiedzi na pytanie, czy dane doświadczenie jest dla autora dobre czy złe. Ostatnim elementem triady jest natomiast nie tyle skłonność człowieka do wartościowania, ile właściwość człowieka polegająca na wartościowaniu. Zdaniem Tymiakina, w procesie komunikacji do czynności wartościowania dochodzi niemal bez przerwy. Może mieć ono charakter ilościowy lub jakościowy, bezrefleksyjny lub racjonalny, pośredni lub bezpośredni, bardziej lub mniej subiektywny. Zawsze jednak dokonywane jest ono w obrębie siatki aksjologicznej wypracowanej i obowiązującej w danej kulturze³⁹⁹.

Można stwierdzić, że opisywana koncepcja trafnie odzwierciedla dominujące podejście do emocji w ramach badań nad analizą sentymentu oraz jego przejawów w tekstach dziennikarzy finansowych. Zanim jednak analizie poddane zostaną wyniki dotychczasowych badań, istotne jest uchwycenie teoretycznego tła tej problematyki poprzez osadzenie jej w perspektywie teorii selekcji informacji. Już bowiem wstępny przegląd literatury stanowi przesłankę do stwierdzenia, że publikowane treści obarczone są czynnikami emocjonalnymi, które mogą mieć wpływ na odbiór tekstu. Można zauważyć, że w tym kontekście dominująca klasyfikacja emocji obejmuje podział na pozytywne i negatywne, przy czym ta druga kategoria wydaje się stanowić przedmiot większego zainteresowania. Takie podejście może być wynikiem tego, że pozytywne emocje są zarówno trudniejsze do zaobserwowania, jak i do późniejszego zinterpretowania. Jak twierdzi jeden z respondentów badań prowadzonych przez Alyta Damstra oraz Knuta De Stewarta: „negatywne zmiany są gwałtowne i łatwe do

³⁹⁹ Zob. Tymiakin, L., *O triadzie komunikacyjnej: wartościowanie – emocje – ekspresja (On the Communication Trifecta: Valuating – Emotions – Expression)*, „Annales Universitatis Mariae Curie-Skłodowska Lublin – Polonia”, 2017, nr II, s. 199–216.

interpretacji. Natomiast pozytywne trendy są trwałe. Takie zmiany są bardzo trudne do uchwycenia, zwłaszcza w przypadku napiętych terminów⁴⁰⁰”. Z drugiej jednak strony, dla dziennikarzy włączanie się w pozytywne długofalowe trendy może być łatwiejsze niż koncentrowanie się na pojedynczych negatywnych zdarzeniach.

Ten wysiłek, na co zwracają uwagę Elitan Goldman, Jordan Martel i Jan Schneemeier, związany jest z treściami, które przekazywane są przez same spółki. Według nich kluczowe jest zrozumienie specyficznej, endogenicznej relacji, zachodzącej między dziennikarzem a jego podstawowym źródłem danych, czyli analizowaną organizacją. Tok swojej analizy oparli o proste założenie: spółkom zależy na tym, żeby rozpowszechniać pozytywne informacje i nie rozpowszechniać informacji negatywnych. W związku z tym zarządy przedsiębiorstw dokonują czegoś, co badacze określają mianem „zaciemniania” (ang. *obfuscation*) niekorzystnych przekazów. Robią to np. poprzez stosowanie w raportach niejasnych sformułowań, czy też nadmierne podkreślanie pozytywnych treści. Istotne jest to, że dziennikarze zdają sobie z tego sprawę. W ten sposób budowane jest w dziennikarzach skrzywienie poznawcze, które autorzy zdefiniowali jako „uprzedzenie *ex ante*” (ang. *ex ante bias*). Pomimo więc faktu, że nadejście dobrych i złych wiadomości jest równie prawdopodobne, to skłonność do ich publikowania zmienia się w zależności od zabarwienia emocjonalnego. Okazuje się, że dobre wiadomości publikowane są właściwie zawsze z dodatnim prawdopodobieństwem, skrajnie złe wiadomości wprawdzie również przedstawiane są z dodatnim prawdopodobieństwem, ale ze wskazaniem, że jest ono niskie, natomiast wiadomości, których wydźwięk byłby nieznacznie negatywny, nie są publikowane w ogóle. Taki stan rzeczy wynika z kilku czynników. Najbardziej prozaicznym jest ten, że dziennikarze mają świadomość konieczności odkodowania „zaciemnionych” treści, co może być czasochłonne. Spółki w ten sposób próbują więc wpływać na motywacje dziennikarzy do podejmowania konkretnych tematów. Co więcej, mechanizm ten objawia się tym silniej, im bardziej wykwalifikowani są dziennikarze, z którymi spółka ma do czynienia. Dodatkowo autorzy każdorazowo muszą odpowiedzieć na pytanie o wartość informacyjną, jaką ich publikacja przyniesie odbiorcom. W związku z tym im mniej jasna

⁴⁰⁰ Damstra, A., De Swert, K., *The making of economic news: Dutch economic journalists contextualizing their work*, „Journalism”, 2020, nr 22/12, s. 3091.

i jednoznaczna jest treść, jaką przekazuje spółka, tym niżej oceniona zostanie jej przydatność⁴⁰¹.

Przyjmując perspektywę historyczną, zauważyć można, że na brak aktywności dziennikarzy EBF w okresach załamań gospodarczych – a więc w sytuacjach, w których przekaz informacyjny powinien przyjmować sentyment negatywny – uwagę zwracano już wcześniej. Na przykład Gary James Merrill zauważył, że teza o udziale mediów finansowych w podtrzymywaniu nastrojów „boomu” i pośredniego przyczyniania się do powstawania baniek spekulacyjnych, znajduje swoje potwierdzenie w badaniach, czego przykładem jest kryzys finansowy 2007/2008⁴⁰². Podobne zdanie prezentuje Robert Shiller, opisując Wielki Kryzys (ang. *Great Depression*), który rozpoczął się w 1929 roku oraz krach giełdy w roku 1987 (tzw. Czarny poniedziałek)⁴⁰³. Na analogiczne zjawisko wskazuje Jocelyn Pixley w przypadku banki Dotcom, po której ekonomiści obwiniali media finansowe za wzniesienie optymizmu rynkowego na irracjonalnie wysoki poziom⁴⁰⁴.

Istnieją również przesłanki rynkowe, które skłaniają dziennikarzy do tonowania negatywnych emocji. W tym kontekście mówi się o problemie „samospełniającej się przepowiedni” (ang. *self-fulfilling prophecy*). Obawa ta wydaje się wynikać wprost z założenia, że emocje prezentowane przez dziennikarzy finansowych, mogą wywoływać reakcje inwestorów. W latach 90. na łamach *Washington Post* Howard Kurtz opublikował artykuł, w którego nagłówku zadał pytanie *Czy gospodarka cierpi z powodu choroby medialnej?* (ang. *Is the economy suffering from media malady?*). W artykule przytoczył treść newslettera jednej z agencji nieruchomości: „uważamy, że ciągły napływ negatywnych artykułów w The Post przekształcił *miękkie lądowanie* w znacznie poważniejsze spowolnienie [...]. Tydzień po tygodniu mieliśmy w gazecie źle zbadany, ponury nastrój i fatum, które wpłynęło na zaufanie ludzi do nieruchomości

⁴⁰¹ Zob. Goldman, E., Martel, J., Schneemeier, J., *A Theory of Financial Media*, „Journal of Financial Economics”, 2022, nr 145/1, s. 239–258.

⁴⁰² Zob. Merrill, G. J., *The Political Content of British Economic, Business and Financial Journalism*, Palgrave Macmillan, Cham, 2019, s. 5.

⁴⁰³ Zob. Shiller, R. J., *Irrational...*, op.cit., s. 110–119.

⁴⁰⁴ Pixley, J., *Emotions in Finance: Booms, Busts and Uncertainty*, Cambridge University Press, Cambridge, 2012, s. 70.

jako inwestycji [...]. The Post tworzy wiadomości, a nie je raportuje⁴⁰⁵”. Wielokrotnie w historii doszło zresztą do sytuacji, w których polityczni i ekonomiczni decydenci obwiniali prasę o brak odpowiedzialności, twierdząc, że gdyby tylko dziennikarze skupili się na sile gospodarki, optymizm zostałby przywrócony. Jeden z nagłówków australijskiego dziennika z czasów recesji w latach 60. brzmiał *Nie pozwól mi słyszeć, że nie śpiewasz* (ang. *Don't let me hear you not singing*), co było nawiązaniem do obwiniania przez rząd Roberta Menziesa prasy za osłabienie gospodarki⁴⁰⁶.

Założenie o tym, że emocje odgrywają istotną rolę na rynku finansowym, jest na tyle silne, że o ile sytuacja nie jest tragiczna, jak na przykład w 2008 roku, przywódcy polityczni zawsze wzywają ekspertów do „rozmawiania” o gospodarce, co w praktyce rozumieć należy jako zachętę do łagodzenia tonu wypowiedzi i niewykazywania negatywnych emocji. Z punktu widzenia profesji dziennikarskiej wydaje się, że przesadny pesymizm postrzegany jest również jako zmienna obniżająca jakość informacji. W tym sensie jest on niepożądany, ponieważ może być budowany kosztem czytelnego i atrakcyjnego produktu informacyjnego⁴⁰⁷. Takie myślenie znajduje potwierdzenie w badaniu Marka Boukesa i Rens Vliegthart. W artykułach ekonomicznych wartość informacyjna określona jako „negatywność” widoczna była przede wszystkim gazetach popularnych (13,7%) oraz regionalnych (12,7%). W mniejszym stopniu odznaczała się ona w gazetach opiniotwórczych (10%) oraz finansowych (9,2%)⁴⁰⁸.

Z drugiej jednak strony niektórzy badacze sugerują, że media w pewnym stopniu zmieniły swoje podejście, w szczególności do raportowania aktywności sektora bankowego, a od czasu wybuchu kryzysu 2007/2008 odsetek negatywnych informacji pozostaje wyższy niż przed kryzysem⁴⁰⁹. Weuhua Huang twierdzi nawet, że pozytywne interpretacje wydarzeń mogą zaszkodzić reputacji dziennikarzy, niezależnie od tego,

⁴⁰⁵ Kurtz, H., *Is the economy suffering from media malady?*, <https://www.washingtonpost.com/archive/business/1990/10/28/is-the-economy-suffering-from-media-malady/52d9122b-792d-480d-9658-483e334f31e2>. Dostęp: 9 mar. 2023.

⁴⁰⁶ Pixley, J., *Emotions in Finance...*, op.cit., s. 70.

⁴⁰⁷ Damstra, A., De Swert, K., *The making of economic news...*, op.cit., s. 3090.

⁴⁰⁸ Boukes, M., Vliegthart, R., *A General Pattern...*, op.cit., s. 292.

⁴⁰⁹ Reuters Institute for the Study of Journalism: Media Coverage of Banking and Financial News, University of Oxford, 2014. s. 29–30.

w jakim stopniu adekwatnie opowiadają o rzeczywistości. Jest to wynikiem nieufności inwestorów i ich milczącego założenia o swoistej zмовie między dziennikarzami a spółkami. W konsekwencji oportunistyczni dziennikarze mogą mieć skłonność do publikowania niekorzystnych analiz. W dłuższej perspektywie jednak, co zaznacza Huang, ci sami dziennikarze potrafią opracowywać strategie losowego publikowania interpretacji pozytywnych i negatywnych⁴¹⁰.

Naturalnie kryzys finansowy nie jest jedynym czynnikiem przemawiającym na rzecz publikowania treści o negatywnym nacechowaniu. Istotną rolę odgrywają opisywane już oczekiwania, dotyczące pracy dziennikarzy, które wzmacniać mogą skłonność do negatywizmu, zwłaszcza w sytuacjach, w których na rynkach dominuje niepewność co do rzeczywistej kondycji i stanu gospodarki. Zdaniem samych dziennikarzy finansowych w okresach trudnej sytuacji gospodarczej odbiorcy zgłaszają większe zapotrzebowania na informację oraz interpretację wydarzeń. Można więc mniemać, że w odpowiedzi na to, autorzy będą starali się podążać za nastrojami odbiorców, co wydaje się mechanizmem naturalnym⁴¹¹.

Niezależnie jednak od tego, czy dziennikarze EBF będą skłonni kierować swoją uwagę na wydarzenia o wydźwięku negatywnym, czy też pozytywnym, przytoczone powyżej powody dalekie są od czynników bezpośrednio związanych z problemem efektywności informacyjnej. Mówiąc inaczej, istnieje szereg przesłanek, wskazujących na to, że decyzje składające się na mechanizm selekcji informacji, nie zawsze powodowane są ideą efektywności. Jak twierdzi Jocelyn Pixley: „to tyle, jeśli chodzi o ortodoksyjne idee *doskonałej informacji* i przejrzystości na wolnym rynku. Kluczowa rola dziennikarzy, jaką jest informowanie opinii publicznej [...] – rzadko może ujrzeć światło dzienne⁴¹²”.

⁴¹⁰ Huang, W., *The Role of Financial Media in Corporate Financing*, „SSRN Electronic Journal”, 2010, s. 3.

⁴¹¹ Damstra, A., De Swert, K., *The making of economic news...*, op.cit., s. 3086–3092.

⁴¹² Pixley, J., *Emotions in Finance...*, op.cit., s. 70.

Rozdział IV – Sentyment tekstowy i jego zastosowanie w finansach

Celem niniejszego rozdziału jest przybliżenie istoty analizy tekstu oraz możliwości jej zastosowania w dziedzinie finansów. Motywacją do podjęcia rozważań nad tym problemem jest pojawienie się w ostatnich latach nowego typu danych, określanych mianem *Big Data*. W najprostszym ujęciu stwierdzić można, że ich istnienie jest wynikiem codziennej ludzkiej aktywności, polegającej na systematycznym i rutynowym wykorzystywaniu urządzeń elektronicznych, między innymi w celach przekazywania i odbierania informacji. Wydaje się, że rynki finansowe w niemal naturalny sposób włączyły się w ten proces. Z jednej strony widoczne jest to poprzez zmiany, jakie zachodzą w sposobie transmisji danych o obrocie instrumentami finansowymi i zawierania transakcji handlowych, z drugiej natomiast objawia się poprzez mechanizmy komunikacji, jakie widoczne są w ramach praktyki rynkowej. Cechy charakterystyczne tego rodzaju danych, takie jak: wielkość, zmienność, różnorodność, prędkość czy wielowymiarowość, sprowokowały natomiast badaczy i analityków do poszukiwania nowych narzędzi, umożliwiających wydobywanie z tych nieustrukturyzowanych zbiorów danych wartości – to jest ich eksplorowania.

Na kanwie powyższych obserwacji, przedstawiciele różnych nurtów badawczych związanych z ekonomią i finansami, zaczynają wykazywać zainteresowanie wykorzystaniem tego typu danych w celach zarówno analitycznych, jak i predykcyjnych. Jednym z zagadnień, związanych z *Big Data*, jest tematyka określana mianem eksploracji tekstu. Tekst, co szczególnie zauważalne jest przy stosowaniu analizy fundamentalnej, stanowi nieodłączny – a nawet dominujący – element przekazu informacyjnego. Wykorzystanie potencjału, jaki niesie za sobą ten typ badań, możliwe jest dzięki zastosowaniu narzędzi informatycznych, w tym programowania i dedykowanych słowników dziedzinowych. Jak można zauważyć, eksploracja tekstu jest zagadnieniem interdyscyplinarnym, które łączy w sobie pozornie sprzeczne nauki, takie jak: lingwistyka i informatyka, a w przypadku niniejszej dysertacji, również finanse, ze szczególnym uwzględnieniem paradygmatu behawioralnego.

Poniższy rozdział składa się z dwóch części. W pierwszej szczegółowo omówione zostało zagadnienie *Big Data* oraz ich wpływu na potencjalny proces badawczy.

Kolejno przedstawiony został teoretyczny model eksploracji, stosowany – co do zasady – na dużych zbiorach danych, jak również objaśnione zostały podstawowe pojęcia związane z analizą tekstu, w tym podjęta została próba zdefiniowania sentymentu. W drugiej części, autorka przybliżyła najpopularniejsze sposoby aplikowania tego typu analiz do badań o tematyce finansowej. Na koniec dokonany został przegląd dotychczasowych badań, związanych z wykorzystaniem eksploracji sentymentu na rynku kapitałowym.

4.1. Big Data – ewolucja danych a proces badawczy

Rozważania na temat analizy tekstu warto zacząć od charakterystyki danych, jakimi obecnie operuje się w rzeczywistości naukowej i biznesowej. W tym celu konieczne jest przybliżenie dwóch pojęć, które – choć nie definiują *explicite* – to jednak oddają sens zmian w tej dziedzinie. Mowa tu o *Big Data* oraz ewolucji, jaka zachodzi w obszarze danych jakościowych. XXI wiek określany jest mianem *epoki Big Data*⁴¹³. Otacza nas niezliczona liczba sensorów, generujących niespotykaną dotąd liczbę danych, a duża część z nich może być, lub w jakiś sposób jest już, przetwarzana⁴¹⁴. W chwili pisania niniejszej dysertacji truizmem wydaje się stwierdzenie, że między człowiekiem a nowymi technologiami zachodzi nieustająca interakcja. Oczywiście wydaje się również to, że owa interakcja przejawia się przynajmniej na dwa sposoby. Z jednej strony człowiek jest biernym odbiorcą zapisywanej cyfrowo – i rozumianej w tym przypadku bardzo szeroko – informacji. Z drugiej strony natomiast – bardziej lub mniej świadomie – jest również jej kreatorem. Już w 2015 roku przeprowadzone na obywatelach Stanów Zjednoczonych badania wykazały, że niemal 1/4 osób znajduje się w trybie online „prawie na stałe”⁴¹⁵. W analogicznych badaniach przeprowadzonych sześć lat później wykazano, że wynik ten

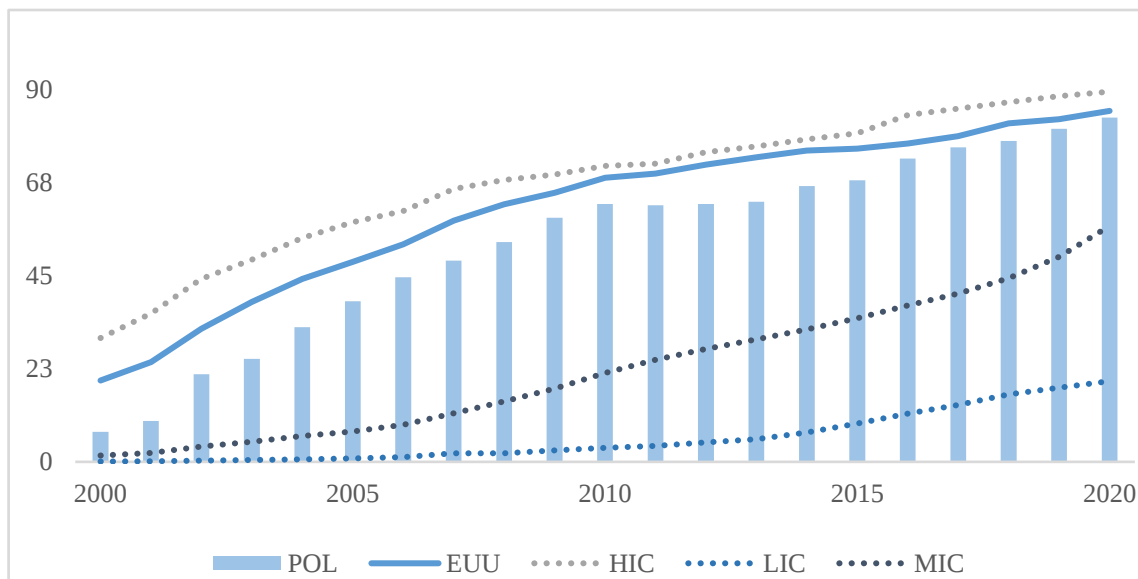
⁴¹³ Zob. Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data: A survey*, „Computer Science Review”, 2015, nr 17, s. 70–81.

⁴¹⁴ Chen, P., Chun-Yang, Z., *Data-intensive applications, challenges, techniques, and technologies: A survey on Big Data*, „Information Sciences”, 2014, nr 275, s. 316.

⁴¹⁵ Zob. Digital Marketing Community: *How People Use Their Devices: What Marketers Need to Know*, 2016 | *Think With Google*, <https://www.digitalmarketingcommunity.com/researches/how-people-use-their-devices-what-marketers-need-to-know-sep-2016-think-with-google/>. Dostęp: 1 kw. 2023.

zbliżył się już do 1/3. W ujęciu światowym liczba osób korzystających z Internetu w ciągu dekady wzrosła ponad dwukrotnie z poziomu 2,17 do 4,9 miliarda osób⁴¹⁶.

Wykres 1. *Osoby korzystające z Internetu (% populacji)*



POL – Polska; EUU – Unia Europejska; HIC – Kraje o wysokich dochodach; LIC – Kraje o niskich dochodach; MIC – Kraje o średnich dochodach. Źródło: Bank Światowy: World Development Indicators, <https://databank.worldbank.org/reports.aspx?source=2&series=IT.NET.USER.ZS&country=#>, Dostęp: 1 kw. 2023.

Z danych przedstawianych przez Główny Urząd Statystyczny wynika, że w Polsce już w latach 2012–2016 dostęp do Internetu w domu posiadało 80% obywateli⁴¹⁷. W 2022 roku poziom ten wzrósł do 93%⁴¹⁸. Z raportu przedstawionego przez zespół Kepios, *Digital 2023: Poland*, wynika natomiast, że w Polsce użytkownikami Internetu jest ponad 36 milionów obywateli, co stanowi 88% całej populacji. Dziennie z Internetu (w dowolny sposób) korzystamy średnio przez 6 godzin i 42 minuty, przy czym niemal 1,5 godziny spędzamy na przeglądaniu prasy. Ponadto, dla 70% osób głównym powodem użytkowania Internetu jest poszukiwanie informacji, a dla ponad 42% jest to zarządzanie swoimi finansami i oszczędnościami. Niewiele mniej, bo 39,6%, posługuje się urządzeniami elektronicznymi z dostępem do Internetu w celu korzystania z systemów

⁴¹⁶ Zob. Flynn, J., *How Many People Use The Internet? [2023]: 35 Facts About Internet Usage In America And The World*, <https://www.zippia.com/advice/how-many-people-use-the-internet/>. Dostęp: 1 kw. 2023.

⁴¹⁷ Zob. *Spółeczeństwo informacyjne w Polsce. Wyniki badań statystycznych za lata 2012–2016*, Główny Urząd Statystyczny, 2016.

⁴¹⁸ Zob. *Spółeczeństwo informacyjne w Polsce w 2022 r.*, Główny Urząd Statystyczny, 2022.

bankowych, inwestycyjnych oraz ubezpieczeniowych⁴¹⁹. W konsekwencji takiej intensywności, poprzez tworzenie, przechwytywanie, kopiowanie lub wykorzystywanie w inny sposób, codziennie generujemy około 328,77 milionów terabajtów danych (0,33 zettabajtów). Szacuje się, że 90% światowych zasobów danych wygenerowanych zostało tylko w ciągu ostatnich dwóch lat. W ciągu 13 lat natomiast, liczba ta wzrosła blisko 60-krotnie i można spodziewać się, że zjawisko to będzie jedynie narastało⁴²⁰.

Tymczasem każda taka aktywność pozostawia ślad w postaci danych, które stały się *spiritus movens* tego, co najczęściej określa się mianem Big Data⁴²¹. Można by powiedzieć, że termin ten cechuje naszą rzeczywistość tak dalece, jak dalece różne dziedziny życia są w niego zaangażowane. Obecnie zaangażowanie to zauważalne jest zarówno w globalnej ekonomii, państwowej administracji, jak również w życiu codziennym⁴²². Czym zatem są te rosnące zasoby?

Oryginalnie termin Big Data – jak sama nazwa wskazuje – określać miał po prostu duże zbiory danych. Z czasem jego definicja uległa jednak rozszerzeniu. Aktualnie nie dotyczy więc jedynie rodzaju danych charakteryzujących się dużą objętością, ale immanentnie związany jest z zachodzącymi procesami technologicznymi⁴²³. Owe procesy wydają się na tyle istotne, że stawiają praktyków i teoretyków rynków finansowych przed koniecznością dostosowania lub nawet zmiany dotychczas wykorzystywanych metod badawczych. W tym kontekście spotkać można nawet twierdzenia, opisujące Big Data jako takie dane, które – z uwagi na swoją specyfikę – nie mogą być przetworzone lub przeanalizowane przy użyciu tradycyjnych procesów oraz narzędzi⁴²⁴.

⁴¹⁹ Kepios Pte. Ltd., *Digital 2023: Poland. The essential guide to the latest connected behaviours*, <https://indd.adobe.com/view/f37abde9-e4ac-4df0-8eef-3733deaa1cc2?allowFullscreen=true&wmode=opaque>. Dostęp: 1 kw. 2023.

⁴²⁰ Zob. Duarte, F., *Amount of Data Created Daily (2023)*, <https://explodingtopics.com/blog/data-generated-per-day#how-much>, Dostęp: 1 kw. 2023.;

Por. *Where have you been all my life? How the financial services industry can unlock the value in Big Data*, PwC, 2013.

⁴²¹ *The age of analytics: competing in a data-driven world*, McKinsey Global Institute 2016, s. 22.

⁴²² Zob. Chen, P.; Chun-Yang, Z., *Data-intensive applications...*, op.cit., s. 315.

⁴²³ *Big Data: How it can become a differentiator*, Deutsche Bank 2014, s. 4.

⁴²⁴ Zikopoulos P. C., Eaton, Ch., deRoos, D., Deutsch, T., Lapis, G., *Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data*, McGraw-Hill, Nowy Jork, 2011, s. 3.;

Por. Goswami, P., Madan, S., *A Survey on Big Data & Privacy Preserving Publishing Techniques*, „Advances in Computational Sciences and Technology”, 2017, nr 10/3, s. 396.

W literaturze przedmiotu *Big Data* definiowane są poprzez przypisywane im cechy. Najczęściej przywoływaną tu koncepcją jest *3V*, przez którą rozumieć należy: wielkość danych (ang. *volume*), ich prędkość (ang. *velocity*) oraz różnorodność (ang. *variety*)⁴²⁵. Zagadnienie wielkości odnosi się bezpośrednio do ilości danych, z jakimi mamy do czynienia podczas przetwarzania Big Data⁴²⁶. Ich aktualny wolumen – co zostało już zaznaczone – jest z jednej strony efektem rosnącej liczby instrumentów, służących do ich kreowania (np. media społecznościowe, aplikacje telefoniczne, czujniki danych), z drugiej natomiast rosnącej zdolności różnego typu urządzeń do ich gromadzenia i przetwarzania⁴²⁷. Wielkość owych danych do niedawna określana była za pomocą terabajtów. Obecnie mówi się już raczej o petabajtach, przy czym ewolucja w stronę standardu zettabajtów wydaje się nieunikniona⁴²⁸. Taki stan rzeczy sprawia, że ich zbieranie, przechowywanie, jak również zarządzanie nimi oraz ich analizowanie, nie są możliwe przy pomocy tradycyjnych narzędzi, co wśród niektórych badaczy pozostaje kluczowym kryterium definiującym termin Big Data⁴²⁹. Należy jednak zwrócić uwagę, że kryterium to jest wyjątkowo niejednoznaczne. Dlatego dla zgłębienia tematu, wyjaśniania wymagają pozostałe cechy, które charakteryzują omawiane zagadnienie.

W ogólnym rozumieniu cecha prędkości związana jest z tempem przetwarzania danych⁴³⁰. Dokładniej rzecz ujmując, chodzi o szybkość ich tworzenia, przechowywania, analizowania oraz wizualizowania. Można zauważyć, że obecnie wiele z tych elementów odbywa się właściwie w czasie rzeczywistym. Coraz więcej urządzeń przesyła informacje już w momencie ich utworzenia. Przykładowo, co minutę na platformę YouTube przesyłanych jest około 100 godzin filmów, w tym samym czasie wysyłanych jest ponad 200 milionów e-maili oraz 300 000 tweetów, a poprzez wyszukiwarkę Google rejestrowanych jest około 2,5 miliona wyszukiwań⁴³¹. Chaikh Emani, Nadine Cullot oraz

⁴²⁵ Autorem koncepcji jest Doug Laney.

⁴²⁶ Por. Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies*, John Wiley & Sons Inc, New Jersey 2013, s. 16.

⁴²⁷ *Report on Big Data AAPOR Big Data Task Force*, American Association for Public Opinion Research: AAPOR, 2015, s. 8.

⁴²⁸ Zikopoulos P. C., Eaton, Ch., deRoos, D., Deutsch, T., Lapis, G., *Understanding Big Data Analytics...*, op.cit., s. 7.

⁴²⁹ Zob. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, McKinsey Global Institute 2011, s. 1.

⁴³⁰ Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies...*, op.cit., s. 16.

⁴³¹ Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s. 397.

Christophe Nicolle twierdzą wręcz, że zagadnienie szybkości wskazuje na konieczność patrzenia na Big Data jako na dane znajdujące się w ciągłym ruchu. W tym kontekście środek ciężkości kładziony jest bowiem na tempo pętli sprzężenia zwrotnego (ang. *feedback loop*). Przez pojęcie to rozumieć można czas potrzebny do przejścia pełnego procesu od załadowania danych do podjęcia decyzji, która przy ich pomocy może zostać wygenerowana⁴³².

Właściwością Big Data jest również ich wewnętrzna różnorodność⁴³³. Odnosi się do różnych typów danych, które podlegać mogą przetwarzaniu⁴³⁴. Niektóre z nich mają charakter ustrukturyzowany, podczas gdy inne są owej struktury pozbawione⁴³⁵. Zauważyć można, że Big Data rzadko jednak występują w uporządkowanej, gotowej do przetwarzania formie. Jest to bezpośrednio spowodowane źródłami, z jakich są one pozyskiwane. Należy wiedzieć, że dane te z reguły nie są tworzone dla potrzeb analitycznych, a są jedynie wynikiem aktywności tysięcy lub milionów, rozproszonych w sieci autorów (twórców filmów, zdjęć, dźwięków, czy też tekstów)⁴³⁶. Problem wewnętrznej różnorodności pozwala podzielić dane na trzy klasy: ustrukturyzowane (ang. *structured data*), częściowo ustrukturyzowane (ang. *semi-structured data*) oraz nieustrukturyzowane (ang. *unstructured data*).

Z pierwszym typem danych mamy do czynienia wówczas, gdy pobieramy je z tak zwanych relacyjnych baz danych, czyli takich, w których przechowywane elementy są ze sobą powiązane. Jeżeli każdy wiersz – znajdującej się w bazie – tabeli jest rekordem, możliwym do zidentyfikowania poprzez unikatowy klucz, a kolumny, zawierające atrybuty danych, opisują wartości poszczególnych rekordów, to możliwe jest ustalenie pewnego układu, określanego mianem modelu relacyjnego. Tego typu bazy opisać można jako proste, intuicyjne dla użytkowników, a przez to bardzo powszechne⁴³⁷. Tym samym dokonywanie za ich pomocą procesów przetwarzania danych możliwe jest na większości

⁴³² Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data...*, op.cit., s. 72.

⁴³³ Chen, P., Chun-Yang, Z., *Data-intensive applications...*, op.cit., s. 314.

⁴³⁴ Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies...*, op.cit., s. 16.

⁴³⁵ Ibidem, s. 9.

⁴³⁶ Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data...*, op.cit., s. 71–72 [za: I. O'Reilly Media, *Big Data Now: 2014 Edition*, O'Reilly Media, 2014].

⁴³⁷ Zob. Oracle Polska: *Czym jest relacyjna baza danych (RDBMS)?*

<https://www.oracle.com/pl/database/what-is-a-relational-database/>. Dostęp: 1 kw. 2023.

urządzeń elektronicznych. Źródła danych częściowo ustrukturyzowanych stanowią szerszą grupę. Mogą pochodzić zarówno z dzienników sieciowych (tj. internetowych), mediów społecznościowych, czujników danych, jak i poczty elektronicznej. Należy jednak zauważyć, że w tym przypadku, nawet jeżeli sam sposób organizacji przechowywania danych posiadać będzie pewną strukturę, to danych jako takich nie można uznać za ustrukturyzowane. Przykładowo, dane pobierane z serwisów społecznościowych, będą wyposażone w powtarzające się elementy, takie jak data publikacji, czy *nick* autora, ale ich kluczowy element w postaci wpisu czy komentarza, będącego tekstem, już takiej struktury posiadać nie będzie⁴³⁸. Warto w tym miejscu zaznaczyć, że obecnie blisko 80% światowych zasobów danych przyjmuje postać nieustrukturyzowaną, lub ustrukturyzowaną jedynie częściowo⁴³⁹. To w tej przestrzeni mieszczą się właśnie dane, definiowane jako Big Data. Można więc zauważyć, że formy, w jakich występują, charakteryzują się wyjątkową niejednorodnością i złożonością. Puneet Goswami i Suman Madan twierdzą nawet, że w cyfrowej rzeczywistości ich różnorodność jest wręcz niewyobrażalna⁴⁴⁰. Można dojść do wniosku, że istnieją przynajmniej trzy poważne konsekwencje tej konstatacji. Po pierwsze, wydaje się, że dla dalszego objaśniania rzeczywistości, do procesów badawczych należy zacząć włączać dane cechujące się częściową strukturą oraz dane nieposiadające struktur⁴⁴¹. Po drugie, chęć wykorzystywania nieustrukturyzowanych danych, przez co rozumieć należy wyodrębnianie z nich istotnych dla badaczy informacji, oznacza jednocześnie konieczność ciągłego zwiększania mocy obliczeniowej urządzeń, na których owe procesy badawcze są przeprowadzane⁴⁴². I wreszcie, po trzecie, niezbędne wydaje się zweryfikowanie dotychczas wykorzystywanych narzędzi badawczych oraz ich zrewidowanie pod kątem możliwości zastosowania do pracy z Big Data, jak również poszukiwania nowych narzędzi i metod, które taką pracę by umożliwiały.

⁴³⁸ Twitter wprawdzie posiada strukturę w formacie JSON – ale rzeczywisty tekst nie jest zorganizowany na poziomie, który uznać można za satysfakcjonujący. (Zikopoulos P. C., Eaton, Ch., deRoos, D., Deutsch, T., Lapis, G., *Understanding Big Data Analytics...*, op.cit., s. 8)

⁴³⁹ Ibidem, s.8.

⁴⁴⁰ Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s. 396–397; Por. *AAPOR Report on Big Data...*, op.cit., s. 8–10.

⁴⁴¹ Zikopoulos P. C., Eaton, Ch., deRoos, D., Deutsch, T., Lapis, G., *Understanding Big Data Analytics...*, op.cit. s. 8.

⁴⁴² Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s. 396–397.

Dokonując przeglądu literatury, zauważyć można, że niektórzy badacze wskazują, że o ile koncepcja 3V istotnie opisuje właściwości przynależne Big Data, to na obecnym etapie rozwoju wiedzy o tego typu danych, staje się niewystarczająca. W związku z tym ulega ona systematycznemu rozszerzaniu poprzez dodawanie do niej kolejnych cech, pomagających zrozumieć omawiany termin. Do takich cech należą między innymi: nieprecyzyjność (ang. *vagueness*), zmienność (ang. *variability*), prawdziwość (ang. *veracity*), ważność (ang. *validity*), czy też wartość (ang. *value*)⁴⁴³. Wydaje się, że większość z tych właściwości mieści się w problemie, który określić można jako wiarygodność. Biorąc pod uwagę, że mowa tu o dużych zbiorach różnorodnych danych, coraz większego znaczenia nabiera kwestia ich przejrzystości, to znaczy sposobu pomiaru, klasyfikowania i integrowania⁴⁴⁴, jak również ich dokładności i precyzji⁴⁴⁵. Bezsporne jest, że Big Data nie spełnią swojej funkcji, jeśli będą nieprecyzyjne lub niekompletne, zwłaszcza że przetwarzane są one w sposób zautomatyzowany, z czym mamy do czynienia w przypadku programowania i algorytmów. W najprostszym ujęciu wiarygodność danych oznacza, że są one sprawdzone i zgodne z rzeczywistością. Z jednej strony zagadnienie to odwołuje do początkowego etapu pracy analitycznej, czyli procesu oczyszczania danych, który musi być przeprowadzony przed właściwą analizą, w celu zminimalizowania, wygenerowanych poprzez formę danych, różnego typu współczynników szumu⁴⁴⁶. Z drugiej strony, wiarygodność w oczywisty sposób dotyczy otrzymywanych rezultatów i wyników badań. Innymi słowy, określić można ją jako możliwość zaufania, że dane są dokładne i trafne⁴⁴⁷. Dzięki temu, opisane poprzez pewien stopień prawdopodobieństwa wyniki badań, mają szansę odzwierciedlać analizowaną rzeczywistość⁴⁴⁸. Trudności temu zadaniu przydaje fakt, że – jak zostało już wspomniane – transfer tego typu danych coraz częściej odbywa się w czasie rzeczywistym. Jest to jedna z przyczyn nieustannej zmienności treści zawartej w ich

⁴⁴³ Zob. Tsai, C-W., Lai, C-F., Chao, H-C., Vasilakos, A.V., *Big data analytics: a survey*, „Journal of Big Data”, 2015, nr 2/21, s. 2.

⁴⁴⁴ Por. *Big Data in Capital Markets: At the Start of the Journey*, Aite Group, 2014, s. 12.

⁴⁴⁵ Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data...*, op.cit., s. 72.

⁴⁴⁶ Por. Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s. 397.

⁴⁴⁷ Por. *AAPOR Report on Big Data...*, op.cit., s. 9.

⁴⁴⁸ Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data...*, op.cit., s. 72.

zbiorach⁴⁴⁹. Oznacza to, że Big Data charakteryzują się diachroniczną nieciągłością, co odciska swój ślad na metodzie badawczej, na przykład w ramach analizy tekstu.

Należy wreszcie zastanowić się jak określić ogólny cel wykorzystywania nieustrukturyzowanych danych do analiz i różnego typu procesów badawczych. Międzynarodowa Korporacja Danych definiuje pracę z Big Data jako „nową generację oprogramowania i architektury, która ma na celu ekonomiczne wyodrębnianie wartości z bardzo dużej liczby różnorodnych danych [...]”⁴⁵⁰. Można więc mniemać, że istotą rzeczy jest cecha określana jako wartość⁴⁵¹. Zdaniem Johna Gantza i Davida Reinsela ostateczna decyzja, co do wykorzystania Big Data, oceniana jest na podstawie przynajmniej jednego z następujących kryteriów. Po pierwsze należy zastanowić się, czy dostarczą one bardziej przydatnych informacji niż standardowe zbiory danych. Po drugie, czy uzyskana za ich pomocą informacja będzie cechowała się wyższą wiarygodnością. Po trzecie, czy poprawi to terminowość otrzymania odpowiedzi na zadany problem⁴⁵². Skoro włączenie do analiz zbiorów typu Big Data wymagać będzie wykorzystania nowego oprogramowania, zwiększenia mocy obliczeniowych, a być może również zastosowania nowych metod badawczych, to uzyskana w ten sposób wartość uzasadniać powinna, związany z nią nakład pracy⁴⁵³. Proces wydobywania, czy też odkrywania wartości znajdujących się w danych, określany jest mianem ich eksploracji (ang. *data mining*) i stanowi przedmiot rozważań dalszej części rozdziału.

4.2. Eksploracja danych

Warto zadać sobie pytanie, jak wiele wiedzy udaje nam się wydobyć z danych, które obecnie pozyskujemy, kolekcjonujemy i przetwarzamy. Lub inaczej – jaką wiedzę i na jakim poziomie jesteśmy w stanie z tych danych uzyskać⁴⁵⁴? Procesem badawczym, w ramach którego badacze zajmujący się Big Data starają się odpowiedzieć na to pytanie,

⁴⁴⁹ Por. Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s. 397.

⁴⁵⁰ International Data Corporation: *Asia/Pacific Big Data and Cognitive/Artificial Intelligence Systems and Content Analytics*. https://www.idc.com/getdoc.jsp?containerId=IDC_P24232. Dostęp: 20 lip. 2022.

⁴⁵¹ Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data...*, op.cit., s. 72.

⁴⁵² Gantz, J., Reinsel, D., *IDC Iview. Extracting Value from Chaos*, 2011, s. 8.

⁴⁵³ Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s. 397.

⁴⁵⁴ Larose, D. T., *Discovering Knowledge In Data. An Introduction to Data Mining*, John Wiley & Sons Inc, New Jersey 2005, s. 4.

jest właśnie eksploracja danych⁴⁵⁵. Termin ten nieodłącznie związany jest z podejściem interdyscyplinarnym. Stosowane do niego narzędzia są często pochodną celu i dziedziny, w jakiej porusza się badacz. W związku z tym, autorzy ów proces definiują na wiele sposobów⁴⁵⁶. Najprościej eksplorację danych określić można jako wydobywanie wiedzy z ogromnych zbiorów danych, przy czym słowa „wiedza” oraz „informacja” używane są tutaj wymiennie⁴⁵⁷. Jedną z częściej spotykanych w literaturze przedmiotu definicji omawianego zagadnienia jest ta, zaproponowana przez spółkę Gartner, której pracownicy opisują eksplorację danych jako „proces odkrywania znaczących korelacji, wzorców i trendów przez *przesiewanie* dużych ilości danych, przechowywanych w repozytoriach”. Ci sami autorzy dodają, że w procesie tym wykorzystuje się technologie rozpoznawania wzorów, jak również metody statystyczne i matematyczne⁴⁵⁸.

Warto zwrócić uwagę, że powyższe stwierdzenie sugeruje jedynie cel pośredni całego procesu. Chociaż odkrywanie korelacji lub trendów jest elementem, jak również efektem pracy z Big Data, to wciąż zadania te realizowane mogą być przy użyciu standardowych danych, przechowywanych na przykład w bazach relacyjnych. W ramach innych definicji zaznacza się jednak, że eksploracja danych ma na celu „odnalezienie *niepodejrzewanych wcześniej związków* i podsumowanie danych na nowe sposoby, tak aby stały się one zrozumiałe i przydatne dla ich właścicieli⁴⁵⁹”. Takie wymagania, stawiane analizie danych, rodzą niebagatelne konsekwencje. Warto zauważyć, że Big Data nie są tworzone z myślą o potencjalnych korzyściach analitycznych, naukowych czy też biznesowych. Praca z tego typu danymi nie jest zatem pochodną ich istnienia jako takiego, ale związanych z nimi oczekiwań. Innymi słowy, źródła ich zbierania i sposoby przetwarzania zależą bezpośrednio od oczekiwanych przyszłych korzyści.

Można mieć zresztą wątpliwości co do tego, czy zwrot „eksploracja danych” oddaje w pełni istotę komponentów składających się na ten proces oraz jego teleologię. Na przykład Jiawei Han, Micheline Kamber oraz Jian Pei proponują określenie, które sprawia wrażenie bardziej precyzyjnego, to jest „wydobywanie wiedzy z danych”

⁴⁵⁵ Można spotkać się z terminami zastępczymi: ekstrakcja danych, wydobywanie danych.

⁴⁵⁶ Han, J., Kamber, M., *Data Mining Concepts and Techniques*, Elsevier, San Francisco 2006, s. 5.

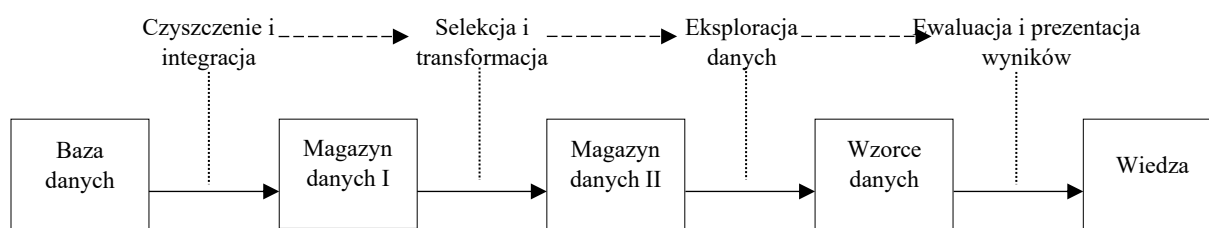
⁴⁵⁷ Zob. *Data Mining data pattern evaluation*, Tutorials Point, 2017, s. 14.

⁴⁵⁸ Gartner, Inc., *Data Mining*, <http://www.gartner.com/it-glossary/data-mining>. Dostęp: 20 lip. 2022.

⁴⁵⁹ Larose, D. T., *Discovering Knowledge In Data...*, op.cit., s. 2 [za: Hand et al.: 5].

(ang. *knowledge mining from data*)⁴⁶⁰. Inni autorzy z kolei zachęcają do stosowania terminu „odkrywanie wiedzy” (ang. *knowledge discovery*)⁴⁶¹. Chociaż w literaturze przedmiotu spotkać można opracowania, które sugerują tożsamość wszystkich tych określeń, to istnieją również takie, które definiują je na różne sposoby⁴⁶². W ramach tego drugiego podejścia, eksploracja danych traktowana jest jako kluczowy, ale tylko jeden z kilku etapów całego procesu, którego ostatecznym celem jest wydobywanie wiedzy⁴⁶³, co zaprezentowane zostało na poniższym rysunku.

Rysunek 5. Eksploracja danych jako element procesu odkrywania wiedzy



Źródło: opracowanie własne na podstawie Han, J., Kamber, M., *Data Mining Concepts and Techniques*, Elsevier, San Francisco, 2006, s. 6–8.

Eksploracja danych, traktowana jako fragment większego przedsięwzięcia, skłania nie tylko do odpowiedzi na pytanie o cel, ale również o sposób takiego działania, co w kontekście tego tematu jest jednym z najszczerzej dyskutowanych problemów. Idąc bowiem tym tropem – za Philipem Chenem i Chun-Yang Zhangiem – powiedzieć można, że eksploracja danych to „zbiór technik służących odkrywaniu wartościowych informacji (wzorców) z danych”⁴⁶⁴. Wartość ta określana będzie poziomem wysnuwanych wobec nich oczekiwań. Z kolei stosowane techniki obejmować będą

⁴⁶⁰ Han, J., Kamber, M., *Data Mining Concepts and Techniques*, Elsevier, San Francisco 2006, s. 6.

⁴⁶¹ Zob. Nikfarjam, A., Emadzadeh, E., Muthaiyah, S., *Text mining approaches for stock market prediction*, 2010, The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010 Singapore.; Zob. także. Chun, S. H., Kim, S. H., *Data mining for financial prediction and trading: Application to single and multiple markets*, „Expert Systems with Applications, 2004, nr 26/2, s.131–139. Zob. także. Gaikwad, S.V., Chaugule, A., Patil, P., *Text Mining Methods and Techniques*, „International Journal of Computer Applications”, 2014, nr 85/17, s. 42–45.; Zob. także. Vijayarani, S., Ilamathi, J., Nithya, S., *Preprocessing Techniques for Text Mining – An Overview*, „International Journal of Computer Science & Communication Networks”, 2015, nr 5/1, s. 7–16.

⁴⁶² Zob. *Data Mining data pattern evaluation*, Tutorials Point, 2017.

⁴⁶³ Zob. Han, J., Kamber, M., *Data Mining Concepts...*, op.cit., s. 6.

⁴⁶⁴ Chen, P., Chun-Yang, Z., *Data-intensive applications...*, op.cit., s. 322.

metody pochodzące z dziedzin: uczenia maszynowego, statystyki, rozpoznawania wzorców, zarządzania bazami danych, czy też wizualizacji⁴⁶⁵.

4.3. Eksploracja danych a analiza tekstu

Zostało już wspomniane, że jeśli poddać by analizie większość występujących na świecie danych, to okaże się, że znacznie częściej spotkać można się z danymi pozbawionymi struktury, niż z takimi, które ową strukturę posiadają. Ponadto, jedynie niewielka część z nich ma charakter numeryczny. Dominujące są natomiast dane tekstowe⁴⁶⁶. Z punktu widzenia biznesu i finansów wynikać może to z dwóch zjawisk. Po pierwsze, sposób komunikacji we współczesnych organizacjach w dużej mierze opiera się na rozbudowanej infrastrukturze dokumentów, wymianie maili, stronach internetowych oraz mediach społecznościowych⁴⁶⁷. Sam fakt wykonywania codziennej pracy, przez zatrudnione w danej firmie osoby, będzie więc generował jakąś ilość danych nienumerycznych. Po drugie, wymagane prawnie lub stanowiące standard komunikacyjny, udostępniane zewnętrznym odbiorcom, raporty w coraz większej mierze opierają się również na tekście. Prawdopodobnie taką zauważyć można zarówno w przepisach krajowych, jak i europejskich, co zostanie opisane w dalszej części tekstu. Warto jednak pamiętać, że również w tej dziedzinie dane oraz ich różnorodność stale rosną⁴⁶⁸. Zrozumienie tych zjawisk przekłada się na rozszerzenie pola zainteresowań naukowców i analityków, którzy do badań opartych na danych ilościowych, włączając zaczęli te o charakterze jakościowym. Tim Loughran i Bill McDonald – jedni z czołowych badaczy omawianej tematyki – twierdzą bowiem, że analizę tekstu traktować można jako część większej dziedziny, tj. analizy jakościowej⁴⁶⁹. Stąd też dla przedstawicieli finansów wynikać może wiele trudności, związanych z takim procesem badawczym. Zastosowanie analizy tekstu do badań predykcyjnych, analitycznych, czy też

⁴⁶⁵ Larose, D. T., *Discovering Knowledge In Data...*, op.cit., s. 2. [za: Evangelos Simoudis w Cabena i wsp.: 6].

⁴⁶⁶ Das, S. R., *Text and Context: Language Analytics in Finance. Foundations and Trends in Finance*, now Publishers Inc, Hanover 2014., s. 3.

⁴⁶⁷ Chen, H., Chiang, R., Storey, V., *Business intelligence and analytics: from big data to big impact*, „MIS Quarterly”, 2012, nr 36/4, s. 1175.

⁴⁶⁸ Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies...*, op.cit., s. 153.

⁴⁶⁹ Loughran, T., McDonald, B., *Textual Analysis in Accounting and Finance: A Survey*, „Journal of Accounting Research”, 2016, nr 54/4, s. 2.

finansowych w ogóle, przenosi okrzepłą metodologicznie dziedzinę na nowy grunt. Przede wszystkim dane jakościowe, w tym dane tekstowe, cechuje wysoki stopień nieprecyzyjności. Oczywiście w ramach dyskusji akademickich niejednokrotnie podnoszona była wątpliwość dotycząca badawczej i interpretacyjnej jednoznaczności danych ilościowych, jednak przyznać można, że na tym polu nauka osiągnęła już pewną dojrzałość⁴⁷⁰.

Analiza tekstu jest tematyką o wyjątkowo eklektycznym charakterze. Przenika przez wiele dyscyplin naukowych i z wielu z nich czerpie metodologiczne i teoretyczne podstawy. Zdaniem Armana Nassirtoussi'a, Saeeda Aghabozorgi'a, Teha Waha i Davida Ngo'a jej zastosowanie wymaga włączenia do procesu badawczego wiedzy, pochodzącej przynajmniej trzech innych dziedzin: lingwistyki, uczenia maszynowego i ekonomii behawioralnej⁴⁷¹. O ile w kontekście tematyki niniejszej dysertacji określenie „ekonomia” nadmiernie rozszerza przestrzeń badawczą i można stwierdzić, że lepszym określeniem byłyby finanse behawioralne, to na ogólnym poziomie taka charakterystyka wydaje się trafna.

Próbę zdefiniowania zagadnienia analizy tekstu rozpocząć można od ogólnego stwierdzenia, że jest to proces badania – pozbawionego struktury – tekstu, przekształcania go w dane ustrukturyzowane i wydobywania z niego istotnych informacji, które w dalszej kolejności mogą być wykorzystane na różne sposoby⁴⁷². Powyższa definicja wskazuje, że między analizą a eksploracją tekstu występuje analogiczna różnica, którą dostrzec można było podczas rozważań nad eksploracją danych oraz wydobywaniem z nich wiedzy. Analiza tekstu ostatecznie doprowadzić musi do określonej wartości danych, co rozumieć można poprzez wiedzę, mogącą na przykład wspierać procesy decyzyjne. Również analizę tekstu z kilku powodów uznać można za pojęcie mylące.

Po pierwsze, zwrot ten jest szeroko stosowany przez lingwistów, których przedmiotem badawczym może być język sam w sobie. Choć oczywiście należy oddać

⁴⁷⁰ Ibidem, s. 5.

⁴⁷¹ Nassirtoussi. A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D. C. L., *Text mining for market prediction: A systematic review*, „Expert Systems with Applications”, 2014, nr 41, s. 7654.

⁴⁷² Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies...*, op.cit., s. 155.

im pierwszeństwo w tej dziedzinie, to jednak finansiści nie są i prawdopodobnie nigdy nie będą w takim samym stopniu zainteresowani językiem jako takim. Warto jednak zaznaczyć, że nauki humanistyczne stosują również analizę tekstu do objaśniania procesów społeczno-kulturowych. W literaturze medioznawczej odnaleźć można następującą definicję analizy tekstu:

jest to jedna ze ścieżek zbierania informacji o tym, w jaki sposób ludzie nadają sens słowom. Jest to metodologia – proces zbierania danych – która służy tym badaczom, którzy chcą zrozumieć jak członkowie różnych kultur i subkultur nadają sens temu, kim są [...] ⁴⁷³.

Autorka niniejszej dysertacji pragnie zwrócić uwagę, że na obecnym etapie rozwoju wiedzy, nie można już traktować płaszczyzny językowej jako linii demarkacyjnej nauk humanistycznych i ekonomicznych. Za sprawą rozwoju ekonomii społecznej, czy finansów behawioralnych, straciła ona swój ultymatywny charakter. Co więcej, z teleologicznego punktu widzenia, proces eksploracji tekstu ma charakter deskryptywny, co jeszcze bardziej przybliża go do, dynamicznie rozwijającego się w ramach finansów, paradygmatu behawioralnego, który uzupełnia (a niektórzy powiedzą, że wręcz neguje) klasyczne normatywne podejście, co opisane zostało we wcześniejszych rozdziałach.

Po drugie, zwrot „analiza tekstu” używany jest również przez badaczy niebazujących na metodach, wymagających zwiększonej mocy obliczeniowej. Początki analizy tekstu w finansach wiązały się z relatywnie prostymi metodami badawczymi, polegającymi na przykład na określaniu stopnia czytelności tekstów finansowych (ang. *readability*). Narzędziem, często wykorzystywanym w tego typu analizach, był tak zwany Indeks Czytelności FOG (ang. *Fog Index*). Na otrzymany wynik indeksu składały się trzy zmienne: liczba słów w dokumencie, liczba zdań i liczba słów określonych jako długie. Im wyższą wartość przyjmował wskaźnik, tym wyższy poziom skomplikowania analizowanego tekstu. Takie badania prowadzone były również w dziedzinie finansów. Jako przykłady podać można tutaj badania Fenga Li z 2008 roku ⁴⁷⁴, Briana Millera

⁴⁷³ McKee, A., *Textual Analysis: A Beginner's Guide*, SAGE, London 2003, s. 1.; tłumaczenie własne.

⁴⁷⁴ Zob. Li, F., *Annual report readability, current earnings, and earnings persistence*. „Journal of Accounting and Economics”, 2008, nr 45, s. 221–247.

z 2010 roku⁴⁷⁵, czy Alastaira Lawrence’a z 2013 roku⁴⁷⁶. Wydaje się, że dotychczasowym przedmiotem zainteresowań w ramach problemu szacowania poziomu czytelności, były głównie przedstawiane przez spółki raporty. Zauważalne jest również, że badacze dochodzili w tym kontekście do skrajnie różnych wniosków. Gary Biddle, Hilary Gilles i Verdi Rodrigo – w badaniach z 2009 roku – stwierdzili, że Indeks FOG może być definiowany jako miernik czytelności sprawozdań finansowych⁴⁷⁷. Z kolei Tim Loughran i Bill McDonald, prowadząc badania kilka lat później, wskazali, że nie jest on odpowiednią miarą, za pomocą której, można by szacować poziom czytelności dokumentów finansowych⁴⁷⁸.

Analizy związane z obliczaniem poziomu czytelności sugerować mogą, że jedna z trzech dziedzin, składających się na analizę tekstu, tj. informatyka, czy też programowanie, w istocie nie są niezbędne. Policzenie liczby słów czy zdań w dokumencie, jest możliwe nawet bez użycia urządzeń elektronicznych. Przegląd literatury w tym zakresie wskazuje jednak, że badacze zajmujący się rynkami finansowymi i analizą tekstu, obrali kierunek związany z koniecznością zastosowania pracy algorytmicznej. Nawet jeśli wykonanie stawianego sobie przez badacza zadania było możliwe bez użycia jakichkolwiek narzędzi wspomagających, to biorąc pod uwagę dostępną obecnie liczbę danych, rezygnacja z aparatu wiedzy informatycznej, oznaczałaby zasadnicze wydłużenie procesu badawczego. Niemalże immanentną właściwością tego typu badań jest bowiem analizowanie dużych zbiorów tekstowych⁴⁷⁹. Takie podejście wymusza zaangażowanie do prac narzędzi umożliwiających pobieranie, przechowywanie i ich przetwarzanie. W literaturze przedmiotu spotkać można się nawet ze sformułowaniem *Big Text*, co jednoznacznie wskazuje, że analiza tekstu znajduje swoje miejsce w nurcie badań z obszaru Big Data⁴⁸⁰.

⁴⁷⁵ Zob. Miller, B., *The effects of reporting complexity on small and large investor trading*, „Accounting Review”, 2010, nr 85, s. 2107–2143.

⁴⁷⁶ Zob. Lawrence, A., *Individual investors and financial disclosure*, „Journal of Accounting & Economics”, 2013, nr 56, s. 130–147.

⁴⁷⁷ Zob. Biddle, G., Hilary, G., Verdi, R. S., *How does financial reporting quality relate to investment efficiency?*, „Journal of Accounting and Economics”, 2009, nr 48/2–3, s. 112–131.

⁴⁷⁸ Zob. Loughran, T.; McDonald, B., *Measuring Readability in Financial Disclosures*, „The Journal of Finance”, 2014, nr LXIX/4, s. 1643–1671.

⁴⁷⁹ Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data...*, op.cit., s. 74.

⁴⁸⁰ Zob. Das. S. R., *Text and Context: Language Analytics in Finance...*, op.cit., s. 3–4.

Z uwagi na powyższe, wydaje się, że bardziej adekwatnym terminem opisującym zastosowanie analizy tekstu w dziedzinie finansów, jest określenie proponowane przez Sanjiva Dasa, tj. eksploracja tekstu (ang. *text mining*). Autor pojęcie to definiuje jako: „automatyczne przetwarzanie na dużą skalę czystego języka zawartego w tekście w cyfrowej formie w celu eksploracji danych, które są przekształcane w użyteczne informacje ilościowe lub jakościowe”⁴⁸¹. Chociaż powyższa definicja – zdaniem autorki – koncentruje się na wszystkich kluczowych aspektach omawianego procesu, to wymaga jednak dodatkowego objaśnienia. Co bowiem rozumieć przez określenie „czysty język”? Bez jasnego zdefiniowania tego pojęcia, trudno będzie zrozumieć, na czym miałyby polegać mechanizmy jego automatycznego przetwarzania. Dlatego w tym miejscu przybliżony zostanie problem języka naturalnego.

4.4. Język naturalny a eksploracja tekstu

Zostało już zaznaczone, że eksploracja tekstu czerpie z wiedzy wypracowanej na gruncie przynajmniej trzech innych dyscyplin naukowych. Należy jednak zaznaczyć, że każda z nich to bogaty zbiór właściwych im paradygmatów i subdyscyplin. I tak w polu zainteresowań lingwistyki mieszczą się takie obszary badawcze, jak chociażby: lingwistyka obliczeniowa, wyszukiwanie informacji, analiza treści, stylometria czy przetwarzanie języka naturalnego⁴⁸². Analitycznym fundamentem dla każdej z nich jest tak zwany język naturalny. Przez to pojęcie rozumie się język, będący podstawą komunikacji między ludźmi. Oczywiście zagadnienie to stanowi przedmiot zainteresowania zarówno nauk humanistycznych, jak i społecznych, a wokół niego zgromadzony został bogaty dorobek naukowy i publikacyjny. Wydaje się jednak, że zgłębianie tej przestrzeni wiedzy, choć niewątpliwie byłoby interesujące, wykraczałoby poza zakres tematyczny niniejszej dysertacji. Z tego punktu widzenia istotne jest takie zdefiniowanie pojęcia, aby możliwe było zarysowanie potencjału badawczego, jaki tworzy analiza języka naturalnego, w dziedzinie finansów oraz wskazanie kluczowych problemów, jakie się z nią wiążą. Jak pisze Ewa Nowicka:

⁴⁸¹ Ibidem, s. 5.

⁴⁸² Loughran, T., McDonald, B., *Textual Analysis in Accounting...*, op.cit., s. 1187.

języki ludzkie operują mniejszą lub większą liczbą symboli, która jest jednak na tyle duża, że pozwala na przekazanie potrzebnego zakresu emocji i myśli i dobrze służy funkcjonowaniu społeczeństwa we wszystkich dziedzinach życia. Różnicuje je nie tylko fonetyka, ale także bogactwo słownictwa i nacisk na opis różnych sfer rzeczywistości. Jednakże sam język jest bodaj najistotniejszym przejawem wspólnoty ogólnoludzkiej na poziomie kultury⁴⁸³.

Z jednej strony stwierdzić więc można, że symboliczny charakter języka przesądza o jego społecznie i kulturowo fundamentalnym znaczeniu, poprzez umożliwianie przekazywania myśli i emocji. Z drugiej jednak strony – z badawczego punktu widzenia – generuje właściwe sobie trudności.

Problem ten odwoływać może do koncepcji języka opisanej przez Ferdinanda de Saussure'a. Jego zdaniem tworzywem języka są znaki, na które składają się dwa elementy: *znaczony* oraz *znaczący*. Element *znaczony* nie powinien być utożsamiany z rzeczą, ale raczej z pojęciem, które odpowiada konkretnemu wyrazowi, czyli elementowi *znaczącemu*. Wszystkie znaki w swojej naturze mają identyczną funkcję, którą jest odwoływanie elementu *znaczącego*, znajdującego się wewnątrz systemu językowego, do elementu *znaczzonego*, znajdującego się poza tym systemem⁴⁸⁴. Istnienie języka, rozumianego jako system znaków, warunkowane jest natomiast istnieniem jakiejś „zbiorowości mówiącej”. W tym sensie język definiować należy jako fakt społeczny⁴⁸⁵. Tę w gruncie rzeczy oczywistą prawdę interpretować można na dwa sposoby. Po pierwsze, konieczne jest istnienie zbiorowości, w obrębie której panować będzie pewna zgodność co do relacji zachodzących wewnątrz znaków. Innymi słowy, musi istnieć grupa osób, która konkretny element *znaczący* będzie odwoływała do tego samego elementu *znaczzonego*. W szerokim ujęciu może być to grupa posługująca się wspólnym językiem. Ten sam mechanizm wydaje się zachodzić jednak w przypadku wąskiej grupy, funkcjonującej w obrębie wspólnego systemu dyskursywnego. Po drugie, należy zwrócić uwagę na to, że aby relacja między elementem *znaczonym* a *znaczącym* języka mogła się społecznie urzeczywistniać, konieczny jest indywidualny akt mówienia. Ferdinand de

⁴⁸³ Nowicka, E., *Świat człowieka – świat kultury*, Wydawnictwo Naukowe PWN, Warszawa 2006, s. 392.

⁴⁸⁴ de Saussure, F., *Kurs językoznawstwa ogólnego*, Państwowe Wydawnictwo Naukowe, Warszawa 1961, s. 79.

⁴⁸⁵ Ibidem, s. 87–88.

Saussure określa to mianem mowy jednostkowej (*parole*). Jej właściwością jest to, że stanowi

indywidualny akt woli i inteligencji, w którym należy wyróżnić: a) połączenia, za pomocą których osoba mówiąca posługuje się kluczem języka w celu wyrażenia swojej własnej myśli; b) mechanizm psychofizyczny pozwalający jej uzewnętrznić te połączenia⁴⁸⁶.

Oznacza to, że autor, podczas mówienia, przekształca znaki językowe w jednostkowy akt semantyczny, nadając w ten sposób wypowiedzi cech indywidualnych. Tym samym mowa jednostkowa w pewnym sensie wymyka się z ram systemu znaków, tworząc przestrzeń dla ewolucji języka. Załączek owej ewolucji znajduje się zawsze w mowie jednostkowej⁴⁸⁷.

Z tej perspektywy język naturalny rozumieć można jako system znaków i symboli, które podlegają ciągłej ewolucji, a występujące między nimi reguły są trudne do klarownego sprecyzowania. Dlatego też założenie, że ich desygnaty będą charakteryzowały się jednoznacznością, można uznać za nadużycie, a nawet naiwność. Jednak ludzie rzadko myślą znaczenia. Mają bowiem świadomość istnienia różnych kontekstów, w których występują symbole danego słowa, oraz przynajmniej ogólną wiedzę o świecie rzeczywistym. W tę świadomość nie jest natomiast wyposażony komputer⁴⁸⁸. W opozycji do języka naturalnego znajduje się więc język, który w literaturze przedmiotu, określany jest jako sztuczny. To w nim mieszczą się na przykład wszelkie języki programowania, czy też notacje matematyczne⁴⁸⁹. Na styku języka naturalnego oraz sztucznego znajduje się problem eksploracji tekstu, którego rozwiązanie nazywane jest Przetwarzaniem Języka Naturalnego (ang. *Natural Language Processing* – *NLP*). W szerokim ujęciu pod tym hasłem kryją się właściwie wszystkie komputerowe procesy i transformacje, które dokonywane są na języku naturalnym⁴⁹⁰. Intencją badaczy, zajmujących się tym obszarem, jest zbieranie wiedzy na temat

⁴⁸⁶ Ibidem, s. 29.

⁴⁸⁷ Ibidem, s. 34.

⁴⁸⁸ Jackson, P., Moulinier, I., *Natural Language Processing for Online Applications: Text Retrieval, Extraction, and Categorization*, John Benjamins Publishing, Amsterdam 2007, s. 5–7.

⁴⁸⁹ Por. Bird, S., Klein, E., Loper, E., *Natural Language Processing with Python*, O'Reilly Media, Sebastopol 2009, s. IX.

⁴⁹⁰ Por. Ibidem, s. IX.

rozumienia i używania języka przez ludzi, w celu opracowania odpowiednich narzędzi i technik umożliwiających „rozumienie” i przetwarzanie języków naturalnych poprzez systemy komputerowe, tak aby wykonywać za ich pomocą pożądane zadania⁴⁹¹. Można też powiedzieć, że NLP to zbiór metod obliczeniowych obejmujących szereg technik, służących analizowaniu lub też konstruowaniu tekstów na różne sposoby, po to, aby możliwe było wyodrębnienie leżących u ich podstaw informacji⁴⁹². Mówiąc inaczej, przetwarzanie języka naturalnego to obszar badań, który analizuje, w jaki sposób komputery mogą być używane do rozumienia tekstu (pisanego, mówionego, dźwiękowego), przedstawionego w języku naturalnym i przetwarzania go w celu wypracowania, przydatnych dla użytkowników odpowiedzi lub aplikacji⁴⁹³.

4.5. Sentyment tekstowy

Przegląd literatury, dokonany w niniejszej pracy, wykazał, że terminy takie jak „analiza tekstu”, „eksploracja tekstu”, „eksploracja opinii” czy też „analiza sentymentu” bywają przez autorów stosowane wymiennie. Z jednej strony wskazuje to na pewną metodologiczną niedojrzałość omawianej dziedziny, z drugiej jednak zwraca uwagę na esencjonalność problemu sentymentu w tego typu badaniach. Można bowiem stwierdzić, uchwycenie nastrojów, zawartych w przekazach medialnych, jest na tym polu traktowane jako kwestia zasadnicza i tym samym stanowi jedno z głównych wyzwań badawczych w obszarze eksploracji tekstu⁴⁹⁴.

Do niedawna sentyment inwestorów rynkowych najczęściej mierzony był za pomocą narzędzi statystycznych, skoncentrowanych na porównywaniu spadków i wzrostów cen instrumentów finansowych. W ostatnim czasie jednak widoczny jest trend, w którym zwiększanie potencjału analitycznego, jest równoznaczne z wykorzystaniem metod uwzględniających Big Data⁴⁹⁵. Próba uchwycenia emocji, zawartych w języku naturalnym, odpowiada w szczególności jednej z cech Big Data,

⁴⁹¹ Chowdhury, G. G., *Natural language processing*, „Annual Review of Information Science and Technology”, 2003, nr 37/1, s. 51.

⁴⁹² Blazquez, D., Domenech, J., *Big Data sources and methods for social and economic analyses*, „Technological Forecasting & Social Change”, 2017, nr 130, s. 104.

⁴⁹³ Chowdhury, G. G., *Natural language processing...*, op.cit., s. 51.

⁴⁹⁴ Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., Chen, Y., *The effect of news and public mood on stock movements*, „Information Sciences”, 2014, nr 278, s. 827.

⁴⁹⁵ *Big Data in Capital Markets: At the Start of the Journey*, Aite Group, 2014, s. 28.

to jest zmienności, która odnosi się do permanentnej ewolucji danych oraz zawartych w nich znaczeń. Nowe desygnaty słów wypierają stare, a ich nacechowanie emocjonalne podlega transformacji, co sprawia, że analiza sentymentu staje się niepowtarzalnym wyzwaniem, polegającym na dekodowaniu dużych zbiorów danych tekstowych, charakteryzujących się zmiennością⁴⁹⁶.

Próbując precyzyjnie określić przedmiot swoich zainteresowań, badacze, wykorzystujący perspektywę analizy tekstu, często powołują się na klasyfikacje informacji tekstowych, zgodnej z podziałem na dwie kategorie, to jest: fakty i opinie. Fakty dotyczyć mają obiektywnego przekazu na temat podmiotów, wydarzeń i ich właściwości. Z kolei dla opinii charakterystyczny ma być subiektywizm, poprzez który wyrażane są sentymenty, uczucia i emocje. Autorka pragnie jednak zwrócić uwagę, że takie podejście może nadmiernie upraszczać problem. Nie można bowiem z całą pewnością rozróżnić tych dwóch form przekazu. Przytoczone w dalszej części rozdziału badania, pokazują, że profesjonalny przekaz informacyjny może być nacechowany emocjonalnie, tak samo, jak post na forum internetowym może mieć na celu jedynie przekazanie faktów. Dychotomiczny rozdział faktów od opinii wydaje się raczej utopijny. Mimo to warto zauważyć, że analiza sentymentu traktowana jest albo jako część obszaru określanego mianem eksploracji opinii (ang. *opinion mining*), albo wręcz z tym obszarem utożsamiana⁴⁹⁷. Zagadnienie to może być klasyfikowane – podobnie jak eksploracja tekstu w ogóle – jako subdyscyplina eksploracji danych oraz lingwistyki. Rozumie się przez nie „wydobywanie, klasyfikowanie, rozumienie i szacowanie opinii wyrażanych w różnych źródłach internetowych, komentarzach mediów społecznościowych i innych treściach generowanych przez użytkowników”⁴⁹⁸.

Kolejnym kluczowym zagadnieniem jest samo pojęcie sentymentu. Z kilku powodów jednak, jego zdefiniowanie, nie jest zadaniem prostym. Po pierwsze, termin ten stosowany jest przez przedstawicieli nauk ekonomicznych od bardzo dawna. Na przykład w pracy Charlesa Raymonda Whittlesey’a z 1943 roku, *The Banking System and War*

⁴⁹⁶ Por. Goswami, P., Madan, S., *A Survey on Big Data...*, op.cit., s.397.

⁴⁹⁷ Por. Liu, B., *Sentiment Analysis and Subjectivity* [w:] *Handbook of natural language processing*, (red.) Herbich, R., Graepel, T., Boca Raton: Taylor & Francis Group, 2010, s. 627–628.

⁴⁹⁸ Kim, Y., Jeong, R.S., Ghani, I., *Text Opinion Mining to Analyze News for Stock Market Prediction*, „International Journal of Advances in Soft Computing and Its Applications”, 2014, nr 6/1, s. 2.

Finance, użyty został w kontekście nastrojów rynkowych, związanych z działaniami Systemu Rezerwy Federalnej podczas I Wojny Światowej⁴⁹⁹. Wydaje się, że doprowadziło to do sytuacji, w której większość autorów używa go intuicyjnie, licząc, że odbiorcy, posługując się tym samym tezaurem, zrozumieją znaczenie, bez konieczności jego definiowania. Po drugie, patrząc na rynek jako na dynamiczny proces, przebiegający od zdarzenia rynkowego do decyzji inwestycyjnej, trudno jednoznacznie wskazać miejsce powstawania owego sentymentu. Być może więc klarowne zdefiniowanie tego pojęcia, nawet jeśli zawężymy je wyłącznie do sfery ekonomicznej, zwyczajnie nie jest możliwe, bez jasnego oznaczenia kontekstu, w jakim jest ono używane.

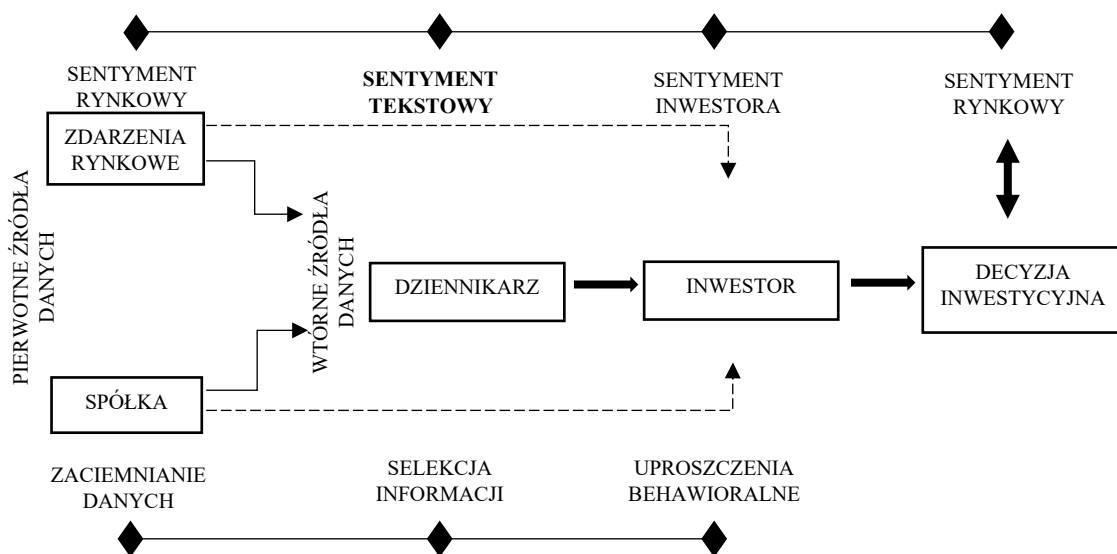
W związku z powyższym, autorka niniejszej dysertacji proponuje dokonanie rozróżnienia w sposobie rozumienia sentymentu, właśnie ze względu na kontekst, w jakim jest używany. Przy czym poprzez kontekst rozumieć należy miejsce jego powstawania w diachronicznym ujęciu dynamiki rynku finansowego, z uwzględnieniem jego uczestników. Podejście to wydaje się mieć sens z kilku powodów. Po pierwsze w zależności od tego, o jakim miejscu powstania sentymentu mowa inne podmioty będą brały udział zarówno w jego wytwarzaniu, jak i odbieraniu – mogą być to emitenci, dziennikarze czy też inwestorzy. Po drugie podmioty te mają właściwe sobie przyczyny kreowania sentymentu. W przypadku emitentów może być to chęć „zaciemnienia danych”, których ujawnienie jest dla nich potencjalnie niekorzystne⁵⁰⁰. W przypadku dziennikarzy mowa jest o szeregu czynników umiejscowionych na zewnątrz lub wewnątrz organizacji, dla której pracują, jak również takich, które zostały przez nich zinternalizowane. Czynniki te przejawiają się w procesach selekcji informacji oraz kadrowania. Oba przypadki zostały szczegółowo opisane w trzecim rozdziale niniejszej rozprawy. Sami inwestorzy, zgodnie z założeniami paradygmatu behawioralnego, narażeni są na różnego rodzaju uproszczenia poznawcze i błędy decyzyjne, wynikające z ich ekspozycji na zewnętrzne lub wewnętrzne bodźce. Koncepcja ta zaprezentowana została w drugim rozdziale pracy. Po trzecie w zależności od kontekstu, pojawienie się

⁴⁹⁹ Zob. Whittlesey, Ch. R., *The Banking System and War Finance*, National Bureau of Economic Research, Cambridge, 1943.

⁵⁰⁰ Zob. Goldman, E., Martel, J., Schneemeier, J., A Theory of Financial Media, „Journal of Financial Economics”, 2022, nr 145/1, s. 239–258.

sentymencie będzie miało pośrednie lub bezpośrednie przełożenie na wynik gry rynkowej. O ile, wszakże emitenci i dziennikarze mogą prezentować raporty lub artykuły, zawierające w sobie mniejszy lub większy stopień sentymentu, to ostatecznie decyzje co do ulokowania swojego kapitału podejmują inwestorzy. Wreszcie zauważyć należy, że sentyment, niezależnie od jego kontekstu, jest kreowany, przetwarzany i odbierany w całym procesie mechanizmu rynkowego, począwszy od momentu zdarzenia gospodarczego, aż do końcowej decyzji inwestorów. Idea ta przedstawiona została na rysunku nr 6.

Rysunek 6. Sentyment tekstowy w procesie kształtowania sentymentu rynkowego



Źródło: opracowanie własne

Intuicyjnie stwierdzić można, że najczęściej sformułowanie „sentyment” wykorzystywane jest do opisu sytuacji rynkowej. Chodzi wówczas o tak zwany sentyment rynku (ang. *market sentiment*). W dowolnym analizowanym okresie rynek najczęściej określić można jako w większości pozytywny (wzrostowy) lub negatywny (spadkowy)⁵⁰¹. Potocznie mówi się odpowiednio o tak zwanym rynku byka (ang. *bullish market*) oraz niedźwiedzia (ang. *bearish market*). O ile zwroty te są powszechnie znane, to należy mieć na uwadze, że wciąż nie został ustalony konsensus, co do przyczyn powstawania sentymentu rynkowego. Można powiedzieć, że na tym polu ścierają się przynajmniej dwie koncepcje. Pierwsza zakłada, że idea sentymentu na rynkach

⁵⁰¹ Picasso, A., Merello, S., Ma, Y., Cambria, E., *Technical analysis and sentiment embeddings for market trend prediction*, „Expert Systems with Applications”, 2019, nr 135, s. 64.

finansowych ma swoje korzenie w fundamentalnym założeniu, że decyzje inwestorów wynikają z ich nastrojów⁵⁰². Dynamika rynku jest efektem dwóch czynników: przejściowych nastrojów irracjonalnych inwestorów (podlegających bodźcom egzogenicznym) oraz ograniczonych zdolności handlowych racjonalnych arbitrażystów⁵⁰³. Innymi słowy, zgodnie z tym podejściem, ilekroć rozpatrywany jest sentyment rynkowy, jego przyczyn poszukuje się w dominujących w danym momencie czynnikach emocjonalnych i psychologicznych. Potwierdzając pośrednio tę koncepcję, Pedro Reis i Carlos Pinho twierdzą, że „sentymenty wyjaśniają ruchy na rynkach finansowych, zwłaszcza w okresach irracjonalnej paniki lub nieuzasadnionego optymizmu”⁵⁰⁴. W ten sposób sentyment niejako utożsamiony jest z zachowaniami irracjonalnymi, powodującymi odejście ceny instrumentów od ich wartości fundamentalnej⁵⁰⁵. Druga perspektywa przypisuje sentyment częściowo czynnikom racjonalnym, a częściowo nieracjonalnym⁵⁰⁶. Z tego punktu widzenia założyć można, że nawet jeśli wszyscy inwestorzy zachowywaliby się zgodnie z hipotezą efektywności informacyjnej, to rynek zawsze będzie charakteryzował się jakimś sentymentem.

Drugim kontekstem, w jakim pojawia się omawiany termin, są inwestorzy. Często stosowaną w tym przypadku klasyfikacją jest podział na inwestorów racjonalnych i nieracjonalnych. Przytoczyć można tu koncepcję Stiglitz i Grossmana⁵⁰⁷. Podobną klasyfikację stosują Bradford De Long, Andrei Skleifer, Lawrence Summers i Robert Waldmann⁵⁰⁸. Z tej perspektywy, sentyment rozumiany jest jako emocje, którymi najczęściej kierują się mniej poinformowani inwestorzy. To na nich najmocniej wpływać będą czynniki behawioralne i to dla nich cena będzie stanowiła źródło informacji.

⁵⁰² Zob. De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., *Positive feedback investment strategies and destabilizing rational speculation*, „Journal of Finance”, 1990, nr 45/2, s. 379–395.

⁵⁰³ Mukhopadhyay, S., *Opinion mining in management research: the state of the art and the way forward*, „Opsearch”, 2018, nr 55, s. 240.

⁵⁰⁴ Nogueira Reis, P. M.; Pinho, C., *A new European investor sentiment index (EURsent) and its return and volatility predictability*, „Journal of Behavioral and Experimental Finance”, 2020, nr 27, s. 1.

⁵⁰⁵ Por. Nogueira Reis, P. M.; Pinho, C., *A new European investor...*, op.cit., s. 1; Zob. Też. Giglio, S., Kelly, B., *Excess volatility: Beyond discount rates*, „The Quarterly Journal of Economics”, 2018, nr 133/1, s. 71–127.

⁵⁰⁶ Zob. Verma, R., Soydemir, G., *The Impact of U.S. Individual and Institutional Investor Sentiment on Foreign Stock Markets*, „Journal of Behavioral Finance”, 2006, nr 7/3, s. 128–144.

⁵⁰⁷ Zob. Stiglitz, J. E., Grossman, S. J., *On the Impossibility of Informationally Efficient Markets*, „The American Economic Review”, 1980, nr 70/3, s. 393–408.

⁵⁰⁸ Zob. De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., *Noise Trader Risk in Financial Markets*, „Journal of Political Economy”, 1990, nr 98/4, s. 703–738.

Chociaż, co zostało scharakteryzowane w rozdziale drugim, obecnie przyjmuje się, że inwestorzy instytucjonalni – którzy co do zasady powinni stanowić grupę dobrze poinformowanych traderów – również nie są wolni od niektórych uprzedzeń poznawczych i motywacyjnych.

Trzecim kontekstem, w jakim pojawia się omawiany termin, jest przestrzeń, którą nazwać można pierwotnymi oraz wtórnymi źródłami informacji. Są to: sprawozdania spółek, media społecznościowe (ang. *social media sentiment*) oraz wiadomości (ang. *news sentiment* lub *media sentiment*). Powodami do takiego umiejscowienia sentymentu są właściwości tekstów, publikowanych zarówno przez spółki, internautów, jak i dziennikarzy. Sanjiv Ranjan Das wymienia następujące właściwości tekstu.

1. Zawiera wartość emocjonalną, która może być użyteczna w ocenie nastrojów rynkowych.
2. Zawiera opinie i związki, które mogą być zbierane i oceniane pod kątem reguł handlowych i analizy rynkowej.
3. Może wyrażać fakty, których nie da się przedstawić za pomocą liczb.
4. Zmniejszają tendencję do nadmiernego agregowania informacji.

Z jednej strony Das zwraca uwagę na potencjalne zalety, płynące z analizy tekstu, a z drugiej wskazuje na ograniczenia, które wiążą się ze standardowymi analizami ilościowymi. Jego zdaniem usilne przekształcanie danych w wartości kwantyfikowane wiąże się z utratą jasności przekazu lub nawet merytorycznej zawartości informacji. Ponieważ za pomocą prezentacji liczbowych dokonuje się – zdaniem tego autora – przesadnego streszczenia informacji, to nie można za ich pomocą uchwycić niuansów, które mogłyby być wyrażone poprzez tekst⁵⁰⁹.

W tym miejscu autorka pragnie wyjaśnić przyczyny stosowania w niniejszej pracy anglicyzmu „sentyment”, zamiast – wspominanych już określeń – „nastrój”, „ton”, czy „nacechowanie emocjonalne”. Terminy te niosą za sobą bowiem istotne konotacje, które mogą mieć przełożenie na precyzyjne określenie przedmiotu badań, prezentowanych w rozdziale empirycznym.

⁵⁰⁹ Das. S. R., *Text and Context: Language Analytics in Finance...*, op.cit., s. 4–5.

Tabela 3. Przykładowe klasyfikacje emocji na gruncie literatury światowej w kontekście wykorzystania metod eksploracji tekstu w obszarze finansów i ekonomii

Rok	Autorzy	Tytuł	Nazwa klasyfikacji emocji	Typy emocji
2022	Arianna Lazzini, Simone Lazzini, Federica Balluchi, Marco Mazza	Emotions, moods and hyperreality: social media and the stock market during the first phase of COVID-19 pandemic	Amotional analysis	podziw, radość, smutek, strach, zaskoczenie, złość, dezaprobata, uśmiech i nuda
2021	Yue Li, John W. Goodell, Dehua Shen	Does happiness forecast implied volatility? Evidence from nonparametric wave-based Granger causality testing	Daily happiness sentiment data (DHS)	szczęście
2020	Wenhao Chen, Kin Keung Lai, Yi Cai	Exploring public mood toward commodity markets: a comparative study of user behavior on Sina Weibo and Twitter	Ekman emotion model	szczęście, dobro, złość, smutek, strach, wstręt, zaskoczenie
2020	Samant Saurabh, Kushankur Dey	Unraveling the relationship between social moods and the stock market: Evidence from the United Kingdom	Psychometric words	miłość, szczęście, złość, smutek, strach i stres
2019	Ruwei Zhao	Quantifying the correlation and prediction of daily happiness sentiment and stock return: The Case of Singapore	Daily Twitter happiness index	szczęście
2018	Tanya Araújo, Samuel Eleutério, Francisco Louçã	Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood	Thomson Reuters MarketPsych indexes (TRMI)	zaufanie, stres, przemoc, konflikt, niepewność, złość, przygnębienie
2018	Zhenkun Zhou, Ke Xu, Jichang Zhao	Tales of emotion and stock in China: volatility, causality and prediction	Online emotions	złość, wstręt, radość, smutek, strach
2017	Jiancheng Shen, Mohammad Najand and Feng Dong, Wu He	News and social media emotions in the commodity market	Commodity-specific emotions	optymizm, strach i radość
2015	Michael Nofer, Oliver Hinz	Using Twitter to Predict the Stock Market: Where is the Mood Effect?	Depressive mood	smutek, beznadzieja, zmęczenie, złość, pozytywny nastrój
2011	Johan Bollen, Huina Mao, Xiaojun Zeng	Twitter mood predicts the stock market	Google-Profile of Mood States (GPOMS)	spokój, czujność, pewność, witalność, bycie miłym, szczęście

Źródło: opracowanie własne

Przeprowadzenie przez autorkę przeglądu literatury pozwoliło zauważyć pewną prawidłowość, w kontekście sposobów klasyfikacji i definiowania wydźwięku emocjonalnego tekstów, podlegających eksploracji. Przykładowo Johan Bollen, Huina Mao oraz Xiaojun Zeng w 2011 roku podjęli próbę predykcji zachowań rynkowych, sprawdzając na platformie Twitter obecność takich emocji jak: spokój, pewność czy czujność⁵¹⁰. Samant Saurabh i Kushankur Dey, zajmując się rynkiem Wielkiej Brytanii, analizowali media społecznościowe pod kątem zwrotów, które wskazywałyby na przeżywanie: miłości, smutku, radości, złości, strachu czy stresu, co określili jako „słowa psychometryczne” (ang. *psychometric words*)⁵¹¹. Wenhao Chen, Kin Keung Lai i Yi Cai badali chiński rynek surowców, wykorzystując do tego – znaną z psychologii – teorię emocji podstawowych Paula Ekmana, którymi są: strach, złość, smutek, pogarda, wstręt, zaskoczenie i radość⁵¹². Z kolei Arianna Lazzini, Simone Lazzini, Federica Balluchi oraz Marco Mazza analizowali zachowanie rynku finansowego podczas pandemii COVID-19, za pomocą takich emocji jak: podziw, radość, smutek, strach, zaskoczenie, złość, dezaprobata, uśmiech oraz znudzenie⁵¹³. Badania te zaprezentowane zostały w tabeli nr 3.

Uwagę zwraca fakt, że autorzy ci, charakteryzując swoje badania, używają takich sformułowań jak „nastrój” czy „emocje”, ale nie „sentyment”. Można więc dojść do wniosku, że część badaczy poszukuje alternatywnych klasyfikacji wydźwięku analizowanych tekstów, którą można by stosować w obszarach związanych z rynkami finansowymi. Przeprowadzenie przeglądu literatury, w zakresie wykorzystania metod eksploracji tekstu w obszarze finansów i ekonomii, pozwala jednak stwierdzić, że – co do zasady – określenie „sentyment” używane jest dla scharakteryzowania kontinuum między wydźwiękiem pozytywnym i negatywnym.

⁵¹⁰ Zob. Bollen, J., Mao, H., Zeng, X., *Twitter mood predicts the stock market*, “Journal of Computational Science”, 2011, nr 2, s. 1–8.

⁵¹¹ Zob. Saurabh S., Dey K., *Unraveling the relationship between social moods and the stock market: Evidence from the United Kingdom*, “Journal of Behavioral and Experimental Finance”, 2020, nr 26, s. 1–9.

⁵¹² Zob. Chen W., Lai, K. K., Cai, Y., *Exploring public mood toward commodity markets: a comparative study of user behavior on Sina Weibo and Twitter*, “Internet Research”, 2020, nr 31/3, s. 1102–1119.

⁵¹³ Zob. Lazzini A., Lazzini S., Balluchi F., Mazza M., *Emotions, moods and hyperreality: social media and the stock market during the first phase of COVID-19 pandemic*, “Accounting, Auditing and Accountability Journal”, 2022, nr 35/1, s. 199–215.

Tabela 4. Zdefiniowanie sentymentu na gruncie literatury światowej w kontekście wykorzystania metod eksploracji tekstu w obszarze finansów i ekonomii, jako stopnia pozytywności i negatywności treści

Rok	Autorzy	Tytuł	Czasopismo
2022	Jing-Rung Yu i inni	Dynamic rebalancing portfolio models with analyses of investor sentiment	International Review of Economics and Finance
2022	Elena Fedorova i inni	Impact of news sentiment and topics on IPO underpricing: US evidence	International Journal of Accounting and Information Management
2022	Ahmed Bouteska i inni	Predictive power of investor sentiment for Bitcoin returns: Evidence from COVID-19 pandemic	Technological Forecasting and Social Change
2021	Laura Rocca i inni	Environmental disclosure and sentiment analysis: state of the art and opportunities for public-sector organisations	Meditari Accountancy Research
2021	Samuel P. Fraiberger i inni	Media sentiment and international asset prices	Journal of International Economics
2021	Ernest N. Biktimirov i inni	Sentiment and hype of business media topics and stock market returns during the COVID-19 pandemic	Journal of Behavioral and Experimental Finance
2020	Haider Maqsood i inni.	A local and global event sentiment based efficient stock exchange forecasting using deep learning	International Journal of Information Management
2020	Saeed Rouhani, Ehsan Abedin	Crypto-currencies narrated on tweets: a sentiment analysis approach	International Journal of Ethics and Systems
2020	Chen Gu, Alexander Kurov	Informational role of social media: Evidence from Twitter sentiment	Journal of Banking and Finance
2020	Alan J. Hanna i inni	News media and investor sentiment during bull and bear markets	European Journal of Finance
2020	Francesco Audrino i inni	The impact of sentiment and attention measures on stock market volatility	International Journal of Forecasting
2020	Yelin Li i inni	The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning	International Journal of Forecasting
2019	Minchae Song, Kyung-shik Shin	Forecasting economic indicators using a consumer sentiment index: Survey-based versus text-based data	Journal of Forecasting

Rok	Autorzy	Tytuł	Czasopismo
2019	Frank Z. Xing i inni	Sentiment-aware volatility forecasting	Knowledge-Based Systems
2018	Stephen Kelly, Khurshid Ahmad	Estimating the impact of domain-specific news sentiment on financial assets	Knowledge-Based Systems
2018	Yasutomo Tsukioka i inni	Investor sentiment extracted from internet stock message boards and IPO puzzles	International Review of Economics and Finance
2018	Mark Johnman i inni	Predicting FTSE 100 returns and volatility using sentiment analysis	Accounting and Finance
2018	Juan C. Reboredo, Andrea Ugolini	The impact of Twitter sentiment on renewable energy stocks	Energy Economics
2018	Shuyuan Deng i inni	The interaction between microblog sentiment and stock returns: An empirical examination	MIS Quarterly: Management Information Systems
2016	Nuno Oliveira i inni	Stock market sentiment lexicon acquisition using microblogging data and statistical measures	Decision Support Systems
2015	Svetlana Maslyuk- Escobedo i inni	News sentiment and jumps in energy spot and futures markets	Pacific Basin Finance Journal
2013	Yang Yu i inni	The impact of social and conventional media on firm equity value: A sentiment analysis approach	Decision Support Systems
2013	Liang-Chih Yu i inni	Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news	Knowledge-Based Systems
2012	Robert P. Schumaker i inni	Evaluating sentiment in financial news articles	Decision Support Systems
2007	Paul C. Tetlock	Giving content to investor sentiment: The role of media in the stock market	The Journal of Finance

Źródło: opracowanie własne.

Robert P. Schumaker i inni w artykule “*Evaluating sentiment in financial news articles*”, stwierdzają:

Ogólnie rzecz biorąc, analiza sentymentu dotyczy analizy tekstu opartego na kierunku i próbuje określić, czy tekst jest obiektywny, czy subiektywny, a także czy subiektywne części zawierają pozytywne lub negatywne uczucia. Ta klasyfikacja na uczucia pozytywne i negatywne jest powszechnie problemem dwuklasowym [...] ⁵¹⁴.

Przykładowo Bing Li i inni, opisując wykorzystane w badaniach dane tekstowe, kategoryzują je w skali od *pozytywny*+ do *negatywny*–, przypisując każdej kategorii dane numeryczne, które nazywają „stopniami sentymentu” (ang. *degree of sentiment*) ⁵¹⁵. Podobnie Svetlana Borovkova i Alan Lammiman, badając reakcję rynków kontraktów terminowych, klasyfikowali sentyment w podziale na pozytywny oraz negatywny ⁵¹⁶. Zijia Du i inni, budując słownik do eksploracji tekstu w języku chińskim, stwierdzili: „terminem *sentymentu finansowego* określamy słowa, które niosą ze sobą negatywną lub pozytywną konotację w kontekście finansowym” ⁵¹⁷. Innymi słowy, wydaje się, że wśród autorów zaczyna panować pewien konsensus, dotyczący tego, co kryje się pod pojęciem *sentymentu tekstowego*. Jednocześnie pozostawiana jest przestrzeń dla alternatywnych klasyfikacji, związanych z różnymi emocjami i nastrojami. Wybrane publikacje, w ramach których autorzy przyjęli taki sposób rozumienia sentymentu tekstowego, zaprezentowane zostały w tabeli numer 4.

W ogólnym ujęciu stwierdzić można, że analiza sentymentu to obszar polegający na takim przetwarzaniu języka naturalnego, aby umożliwić badaczom określenie i wydobycie – z wybranych materiałów źródłowych – nastrojów, będących z ich

⁵¹⁴ Schumaker, R.P., Zhang, Y., Huang, Ch-N., Chen, H., *Evaluating sentiment in financial news articles*, „Decision Support Systems”, 2012, nr 53, s. 458–464.

⁵¹⁵ Zob. Li B., Chan, K.C.C., Ou C., Ruifeng S., *Discovering public sentiment in social media for predicting stock movement of publicly listed companies*, „Information Systems”, 2017, nr 69C, 81–92.

⁵¹⁶ Borovkova, S., Lammiman, A., *News Analytics for Energy Futures*, <https://ssrn.com/abstract=1719582>. Dostęp: 27.12.2023.

⁵¹⁷ Du, Z., Huang, A.G., Wermers, R., Wu, W., *Language and Domain Specificity: A Chinese Financial Sentiment Dictionary*, “Review of Finance”, nr 26, s. 674.

perspektywy użytecznymi danymi⁵¹⁸. O ile jednak praktyka badawcza pokazuje, że sentyment tekstowy definiowany jest między krańcami pozytywnego i negatywnego wydźwięku tekstu, to należy zaznaczyć, że jego szczegółowe określenie musi odbyć się w procesie badawczym. To jakich emocji lub uczuć będzie poszukiwał badacz, zależy bowiem od tego, co uzna za użyteczne dane. Dlatego w kolejnym rozdziale przedstawiony zostanie przegląd dotychczasowych badań, związanych z eksploracją tekstu i jej zastosowaniami w analizach dotyczących rynków finansowych.

⁵¹⁸ Por. Im, L.T., San, W. P., Anthony, P., On, K. C., *Improving polarity classification for financial news using semantic similarity techniques*, „International Journal of Intelligent Information Technologies”, 2018, nr 14/4, s. 41.

Rozdział V – Możliwość zastosowania eksploracji tekstu w finansach

Informacje tekstowe odgrywają niebagatelną rolę zarówno w kontekście pracy analityków, jak i inwestorów rynkowych. Przytoczone w dalszej części rozdziału badania stanowią przesłankę by sądzić, że podejmując decyzje, rejestrują oni nie tylko dane ilościowe, ale również tekstowe⁵¹⁹. Tematyczne artykuły prasowe mogą mieć wpływ na oczekiwania uczestników rynku co do kierunku zmian cen walorów notowanych na giełdzie. Traktować więc można je jako nośnik wiedzy na temat działalności poszczególnych firm, czy też sytuacji makroekonomicznej krajów. Ponadto, gracze rynkowi stali się w ostatnich latach aktywnymi uczestnikami komunikacji internetowej. Szacuje się, że istnieją fora internetowe dla niemal 8 000 akcji, na których inwestorzy mogą wymieniać się informacjami. Komunikacja ta jest cennym źródłem wiedzy, dotyczącej spostrzeżeń uczestników rynku, ich reakcji na informacje oraz panującego na rynku sentymentu⁵²⁰. Zasadniczym, choć bardzo ogólnie pojętym celem badań, których przedmiotem są powyższe zagadnienia, jest takie wykorzystanie technologii, w tym modelowania komputerowego, aby zidentyfikować możliwe relacje, występujące między odbieraną przez inwestorów informacją a ich praktyką ekonomiczną⁵²¹.

Choć ekstrakcję wiedzy z nieustrukturyzowanych danych – zwłaszcza w ramach tak skwantyfikowanych dziedzin, jak finanse – z punktu widzenia procesu badawczego, można określić jako zadanie trudne, to warto dostrzec zawarty w niej potencjał, przemawiający na korzyść podejmowania tego typu prób. Można bowiem zauważyć, że wielu faktów opisujących rzeczywistość, nie da się wyrazić liczbami. A nawet jeżeli jest to możliwe, to ich prosta konwersja na wartości liczbowe może rzutować na czytelność przekazu lub wręcz na jego zawartość merytoryczną. Istnieją również podstawy by sądzić, że tekst potraktować można jako przekaznik emocji towarzyszących jego autorowi, lub odwzorowujący ogólny sentyment rynkowy, co w dalszej kolejności może mieć przełożenie na decyzje inwestycyjne. Na kanwie tych spostrzeżeń można więc

⁵¹⁹ Loughran, T., McDonald, B., *Textual Analysis in Accounting and Finance: A Survey*, „Journal of Accounting Research”, 2016, nr 54/4, s. 1189.

⁵²⁰ Das, S. R., Chen. M. Y., *Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web*, „Management Science”, 2007, nr 53/9, s. 1375.

⁵²¹ Nassirtoussi. A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D.C.L., *Text mining for market prediction: A systematic review*, „Expert Systems with Applications”, 2014, nr 41, s. 7654.

wskazać argumenty przemawiające za tym, że badanie tekstów związanych z szeroko pojętą przestrzenią biznesową i ekonomiczną, może wnieść dodatkową wartość w analizach dotyczących oceny nastrojów rynkowych⁵²².

Należy jednak zastanowić się, co w praktyce – w obszarze zagadnień finansowych – stanowiłyby o wartości tego typu badań. W ogólnym ujęciu cel eksploracji danych jest dwojaki. Z jednej strony jest to klasyfikacja, której istota leży w takim grupowaniu danych, aby najlepiej odwzorowywały rzeczywistość. Z drugiej jest to predykcja, której celem jest prognozowanie przyszłych wartości zmiennych i ich dynamiki⁵²³. Przegląd literatury w zakresie zastosowania analizy tekstu w finansach wskazuje, że badacze z reguły wyznaczają sobie cele oscylujące wokół pytania o możliwość wykorzystania danych tekstowych do celów predykcyjnych. Dla inwestorów, próba przewidzenia przyszłych zdarzeń jest kwestią fundamentalną, a co za tym idzie stanowi jeden z trzonów zainteresowań badawczych. Zdolność do prognozowania procesów rynkowych i innych zdarzeń o charakterze ekonomicznym jest bowiem równoznaczna nie tylko z umiejętnością generowania bogactwa, ale również unikania strat, co w – omawianym w pierwszym rozdziale – paradygmacie behawioralnym ma niebagatelne znaczenie⁵²⁴. Zmiany w wycenie akcji są jednak związane z wieloma czynnikami, takimi jak wydarzenia polityczne, poziom przedsiębiorczości, ogólna sytuacja ekonomiczna, oczekiwania inwestorów, wybory inwestorów instytucjonalnych itd.⁵²⁵. Z natury rzeczy więc przewidzenie dynamiki rynków finansowych, o ile w ogóle jest możliwe, należy do zadań niebywale trudnych⁵²⁶.

Analitycy rynkowi, podejmując próby oszacowania przyszłych wartości instrumentów giełdowych, dokonują tego opierając się na technikach, które powszechnie definiuje się jako analizę techniczną oraz fundamentalną. Jak wskazuje Wiesław Dębski,

⁵²² Zob. Das. S. R., *Text and Context: Language Analytics in Finance. Foundations and Trends in Finance*, now Publishers Inc, Hanover 2014, s. 4–5.

⁵²³ Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies*, John Wiley & Sons Inc, New Jersey 2013, s. 141–149.

⁵²⁴ Zob. Nassirtoussi. A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D.C.L. *Text mining for market prediction...*, op.cit., s. 7653

⁵²⁵ Wang, J-Z. *Forecasting stock indices with back propagation neural network*, „Expert Systems with Applications”, 2011, nr 38, s. 14346.

⁵²⁶ Nassirtoussi. A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D.C.L., *Text mining for market prediction...*, op.cit., s. 7653.

analiza fundamentalna – co do zasady – uznawana jest za bardziej pracochłonną i kosztowną niż analiza techniczna. Jednej z przyczyn takiego stanu rzeczy upatrywać można w tym, że jej przeprowadzenie wymaga wiedzy zarówno z zakresu makroekonomii, polityki, analizy branży, jak również sytuacji finansowej spółek⁵²⁷. Można również stwierdzić, że dane o charakterze fundamentalnym cechują się różnorodnością. Z jednej strony są to oczywiście dane ustrukturyzowane, liczbowe (np. sprawozdania finansowe). Duża część informacji przydatnych podczas tego rodzaju analizy, ma jednak charakter nieustrukturyzowany, w tym tekstowy (np. prasa branżowa)⁵²⁸. Z uwagi na wymagania kompetencyjne oraz zróżnicowanie danych, automatyzacja analizy fundamentalnej jest zjawiskiem relatywnie rzadko spotykanym⁵²⁹. Ponieważ w ostatnich latach technologia związana z eksploracją danych zaczęła być szeroko stosowana właśnie do celów związanych z analizą finansową, w tym obszarze zaczęto również wykorzystywać analizę tekstu. Przeprowadzone dotychczas na tym polu badania, stanowią przesłankę do stwierdzenia, że informacje tekstowe oraz zawarta w nich warstwa emocjonalna, mogą być definiowane jako zmienne, umożliwiające opisanie dynamiki giełd, zarówno na ogólnym poziomie rynkowym⁵³⁰, jak i na poziomie poszczególnych walorów⁵³¹.

⁵²⁷ Dębski, W., *Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki*, Polskie Wydawnictwo Naukowe, Warszawa 2014, s. 217.

⁵²⁸ Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D. C. L., *Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment*, „Expert Systems with Applications”, 2015, nr 42, s. 308.

⁵²⁹ Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D. C. L., *Text mining for market prediction...*, op.cit., s. 7655.

⁵³⁰ Zob. Tetlock, P. C., *Giving Content to Investor Sentiment: The Role of Media in the Stock Market Published*, „The Journal of Finance”, 2007, nr LXII/3, s. 1139–1168.

Zob. Dougal, C., Engelberg, J., García, D., Parsons, Ch. A., *Journalists and the stock market*, „The Review of Financial Studies”, 2012, nr 25/3, s. 639–679.

Zob. Garcia, D., *Sentiment during recessions*, „The Journal of Finance”, 2013, nr 68/3, s. 1267–1300.

Zob. Dzielinski, M., Hasseltoft, H., *Aggregate News Tone, Stock Returns, and Volatility*. „SSRN Electronic Journal”, 2013, nr , s.1–62.

⁵³¹ Zob. Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., *Which News Moves Stock Prices? A Textual Analysis*. *National Bureau of Economic Research*, „Working Paper”, 2013, nr 18725, s. 1–45.

Zob. Heston, S. L., Sinha, N.R., *News versus Sentiment: Predicting Stock Returns from News Stories*, „Finance and Economics Discussion Series. Washington: Board of Governors of the Federal Reserve System”, 2016, nr 2016–048, s. 1–35.

Zob. Chen, H., De, P., Hu, Y., Hwang, B.-H., *Wisdom of crowds: The value of stock opinions transmitted through social media*, „The Review of Financial Studies”, 2014, nr 27/5, s. 1367–1403.

Zob. Heston, S.L., Nitish, R. S., *News versus Sentiment...*, op.cit., s. 1–35.

5.1. Źródła danych tekstowych wykorzystywanych do badań w obszarze finansów

Przeprowadzone dotychczas badania, związane z zastosowaniem analizy tekstu w obszarze finansów, pozwalają podzielić wykorzystywane do nich dane na trzy kategorie: a) dane upubliczniane przez spółki w oficjalnych dokumentach (ang. *corporate disclosures*), b) artykuły publikowane w prasie (w tym dostępne online), c) media społecznościowe. Pierwszy typ źródeł danych to wszelkiego rodzaju sprawozdania, raporty i oficjalne ogłoszenia, przekazywane przez osoby zarządzające do zewnętrznych interesariuszy. Badanie zawartego w nich sentymentu może być użyteczne, jeśli celem badaczy jest na przykład sprawdzenie roli, jaką odgrywają informacje dotyczące bieżącej lub planowanej sytuacji spółki, na ceny akcji poszczególnych przedsiębiorstw⁵³². Przydatność tego typu źródeł danych wynika przede wszystkim z tego, że menedżerowie – odpowiedzialni za konstruowanie raportów – posiadają z reguły pełniejsze i dokładniejsze informacje na temat spółek, którymi zarządzają, niż jakikolwiek zewnętrzny inwestor lub dziennikarz. Poprzez wykorzystywanie publikowanych przez nich tekstów, upatrywać można sposobu na zmniejszenie asymetrii informacji⁵³³. W tym kontekście interesującym zagadnieniem jest – zaproponowana przez Roberta J. Bloomfielda – hipoteza, mówiąca o tym, że raporty spółek, znajdujących się w lepszej sytuacji finansowej, charakteryzują się wyższym poziomem czytelności (ang. *readability*), niż raporty spółek znajdujących się w sytuacji relatywnie słabszej (ang. *management obfuscation hypothesis*)⁵³⁴. Weryfikacji tej hipotezy podjął się na przykład Feng Li, wykazując, że dochody przedsiębiorstw, których raporty są bardziej złożone (tj. skomplikowane pod względem czytelności), charakteryzują się krótkotrwałością. Analogicznie, straty w takich przedsiębiorstwach utrzymywały się dłużej niż w przedsiębiorstwach, których sprawozdania cechował wyższy poziom czytelności⁵³⁵. Oficjalne sprawozdania i raporty analizowane są również pod względem dużo prostszych miar, na przykład ich długości. Przykładowo, przeprowadzone w 2009

⁵³² Kearney, C., Liu, S., *Textual sentiment in finance: A survey of methods and models*, „International Review of Financial Analysis”, 2014, nr 33, s. 173.

⁵³³ Li, F., *Textual Analysis of Corporate Disclosures: A Survey of the Literature*, „Journal of Accounting Literature”, 2010, nr 29, s. 151–152.

⁵³⁴ Bloomfield, R. J., *The “Incomplete Revelation Hypothesis” and financial reporting*, „Accounting Horizons”, 2002, nr 16/3, s. 233–243.

⁵³⁵ Zob. Li, F., *Annual Report Readability, Current Earnings, and Earnings Persistence*, „Journal of Accounting and Economics”, 2008, nr 45, s. 221–247.

roku badania przez Haifeng You i Xiao-jun Zhang wykazały, że wraz z długością raportów 10-K rośnie niedoszacowanie przez rynek cel akcji spółek⁵³⁶.

Wraz z rozwojem badań, do analiz opierających się na oficjalnych publikacjach spółek giełdowych, zaczęto włączać problematykę sentymentu. Jedno z ciekawszych – zdaniem autorki niniejszej dysertacji – pytań badawczych, jakie zostało postawione w tym kontekście, związane było z badaniem zależności między poziomem pesymizmu a potencjałem oportunistycznych zachowań menedżerów. Badania te zaproponował w 2016 roku George E. Iatridis, poddając analizie spółki notowane na giełdzie londyńskiej. Doszedł on do wniosku, że wśród firm, które używają pesymistycznego języka i wykazują wyższy poziom konserwatyzmu w swoich raportach rocznych, zwykle większą rolę odgrywa ład korporacyjny. Ponadto, uzyskane przez niego wyniki wskazywały na to, że pozytywnie związane z poziomem pesymizmu są takie czynniki, jak: własność instytucjonalna, rozdział funkcji CEO i prezesa, czy też zmiany personalne wśród kadry kierowniczej⁵³⁷.

Należy zaznaczyć, że wykorzystywanie oficjalnych sprawozdań spółek do badań w obszarze finansów obarczone jest również wadami. Po pierwsze, autorami sentymentu zawartego w firmowych przekazach są menedżerowie. O ile można założyć, że posiadają oni wiedzę i związane z nią emocje, które nie są dostępne zewnętrznym obserwatorom, to można mieć również wątpliwości co do tego, że zależy im na tym, żeby osoby czytające raport odebrały sytuację firmy jako niekorzystną. Po drugie, raporty sporządzane są najczęściej w odstępach kwartalnych, półrocznych i rocznych. Ten fakt, jak wskazują Colm Kearney oraz Sha Liu, utrudniać może zbadanie ich regularnego wpływu na wycenę akcji⁵³⁸.

⁵³⁶ Zob. You, H., Zhang, X., *Financial reporting complexity and investor underreaction to 10-K information*, „Review of Accounting Studies”, 2009, nr 14/4, s. 559–586.

⁵³⁷ Zob. Iatridis, G. E., *Financial reporting language in financial statements: Does pessimism restrict the potential for managerial opportunism?*, „International Review of Financial Analysis”, 2016, nr 45, s. 1–17.

⁵³⁸ Kearney, C., Liu, S., *Textual sentiment in finance...*, op.cit., s. 173–174.

Tabela 5. *Wybrane badania, wykorzystujące oficjalne dokumenty spółek jako źródło danych w obszarze zastosowania analizy tekstu w finansach.*

Rok	Autorzy	Publikacja	Źródło danych
2009	Matthew Butler, Vlado Kešelj	Financial Forecasting using Character N-Gram Analysis and Readability Scores of Annual Reports	Raporty roczne
2009	Ronen Feldman, Suresh Govindaraj, Joshua Livnat, Benjamin Segal	Management's tone change, post earnings announcement drift and accruals	Raporty roczne, raporty kwartalne
2010	Ramji Balakrishnan, Xin Ying Qiu, Padmini Srinivasan	On the predictive ability of narrative disclosures in annual reports	Raporty roczne
2011	Sven S. Groth, Jan Muntermann	An intraday market risk management approach based on textual analysis	Oświadczenia korporacyjne
2011	Tim Loughran, Bill McDonald	When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks	Raporty roczne
2013	Michael Hagenau, Michael Liebmann, Dirk Neumann	Automated news reading: Stock price prediction based on financial news using context-capturing features	Oświadczenia korporacyjne
2015	Tomasz Piotr Wisniewski, Liafisu Sina Yekini	Stock market returns and the content of annual report narratives	Raporty roczne
2015	Mark Lang, Lorien Stice-Lawrence	Textual analysis and international financial reporting: Large sample evidence	Raporty roczne
2016	Feng-Tse Tsai, Hsin-Min Lu, Mao-Wei Hung	The impact of news articles and corporate disclosure on credit risk valuation	Raporty roczne, raporty kwartalne
2016	Liafisu Sina Yekini, Tomasz Piotr Wisniewski, Yuval Millo	Market reaction to the positiveness of annual report narratives	Raporty roczne
2017	Mathias Kraus, Stefan Feuerriegel	Decision support from financial disclosures with deep neural networks and transfer learning	Oświadczenia korporacyjne
2018	Stefan Feuerriegel, Julius Gordon	Long-term stock index forecasting based on text mining of regulatory disclosures	Oświadczenia korporacyjne

Źródło: opracowanie własne.

Po trzecie, ponieważ sporządzanie sprawozdań ma swój z góry określony czas, to reakcja rynku w momencie ich publikowania jest raczej sytuacją oczekiwaną. Dlatego ewentualny efekt rynkowy może być spowodowany nie tyle absorpcją informacji, ile jej antycypacją⁵³⁹.

Zagadnienia związane z drugim typem danych, jakimi są artykuły prasowe, zostały zaprezentowane w rozdziale trzecim rozprawy oraz kontynuowane będą w dalszej części niniejszego rozdziału. W tym miejscu warto jednak zaznaczyć, że badacze bazujący na źródłach prasowych najczęściej wybierają do tego celu gazety związane z tematyką finansową (np. „The Wall Street Journal” czy „Forbes”) lub tematyką branży, w której funkcjonują badane spółki. Niewątpliwie ten typ danych nie jest obarczony problemem czasu, w rozumieniu, w jakim ma to miejsce w przypadku sprawozdań finansowych. Większość gazet ma obecnie wydania internetowe, których aktualizacja następuje nawet kilka razy dziennie. Z drugiej jednak strony, jest to przestrzeń, w ramach której – wspomniana w rozdziale czwartym – trudność w rozgraniczeniu opinii od faktów, może mieć swoje najpoważniejsze konsekwencje. Niemniej jednak wskazać można szereg badań, wykorzystujących publikacje prasowe do analiz w domenie finansów, co zaprezentowane zostanie w dalszej części rozdziału.

Trzeci typ źródeł danych – media społecznościowe – postrzegane są jako atrakcyjny przedmiot analiz z uwagi na liczbę osób czytających różnego typu fora internetowe i zamieszczających na nich posty oraz komentarze, w tym również takie, które dotyczą rynków⁵⁴⁰. To źródło danych, zarówno z punktu widzenia predykcyjnego, jak i analitycznego, ma jednak pewne wady. Po pierwsze, nie można przyjmować, że prowadzone na ich podstawie badania, odzwierciedlać będą sentyment i zachowania całego rynku. Główną grupą wykazującą aktywność w mediach społecznościowych są inwestorzy indywidualni⁵⁴¹.

⁵³⁹ Nassirtoussi, A.K., Aghabozorgi, S., Wah, T.Y., Ngo, D. C. L., *Text mining for market prediction...*, op.cit., s. 7657.

⁵⁴⁰ Kearney, C., Liu, S., *Textual sentiment in finance...*, op.cit., s. 174.

⁵⁴¹ Zob. Bukovina, J., *Social media big data and capital markets – An overview*, „Journal of Behavioral and Experimental Finance”, 2016, nr 11, s. 20.

Tabela 6. *Wybrane badania, wykorzystujące publikacje prasowe jako źródło danych w obszarze zastosowania analizy tekstu w finansach.*

rok	autorzy	publikacja	źródło danych
1998	Beat Wuthrich, Vincent Cho, Steven Leung, D. Permuntilleke, Jian Zhang	Daily stock market forecast from textual web data	Wall Street Journal, Financial Times, Reuters, Dow Jones, Bloomberg
2007	Gil Rachlin, Dima Alberg, Mark Last	ADM RAL: A data mining based financial trading system	Forbes.com, today.reuters.com
2008	Paul C. Tetlock, Maytal Saar-Tsechansky, Sofus Macskassy	More Than Words: Quantifying Language to Measure Firms' Fundamentals	Wall Street Journal, Dow Jones News Service
2012	Robert P. Schumaker, Yulei Zhang, Chun-Neng Huang, Hsinchun Chen	Evaluating Sentiment in Financial News Articles	Associated Press, Financial Times, PRNewsWire
2013	Jan Kleinnijenhuis, Friederike Schultz, Dirk Oegema, Wouter van Atteveldt	Financial news and market panics in the age of high-frequency sentiment trading algorithms	NRC Handelsblad, de Volkskrant, Algemeen Dagblad, Trouw, and the special-interest financial newspaper Het Financieele Dagblad
2014	Manuel Ammann, Roman Frey, Michael Verhofen	Do Newspaper Articles Predict Aggregate Stock Returns?	Handelsblatt
2015	Marjan Van de Kauter, Diane Breesch, Véronique Hoste	Fine-grained analysis of explicit and implicit sentiment in financial news articles	De Tijd
2015	Fabrizio Lillo, Salvatore Micciché, Michele Tumminello, Jyrki Piilo, Rosario N. Mantegna	How news affects the trading behaviour of different categories of investors in a financial market	archiwum NewsScope
2016	Sheung Yin Kevin Mo, Anqi Liu, Steve Y. Yang	News sentiment to market impact and its feedback effect	Wall Street Journal, New York Times
2016	Nadine Strauß, Rens Vliegenthart, Piet Verhoeven	Lagging behind? Emotions in newspaper articles and stock market prices in the Netherlands	Algemeen Dagblad, De Telegraaf, de Volkskrant, Het Financieele Dagblad, NRC Handelsblad, Trouw
2017	Qiang Song, Anqi Liu, Steve Y. Yang	Stock portfolio selection using learning-to-rank algorithms with news sentiment	Bloomberg, Thomson Reuters News Analytics

rok	autorzy	publikacja	źródło danych
2018	Kun Chen, Peng Luo, Libo Liu, Weiguo Zhang	News, search and stock co-movement: Investigating information diffusion in the financial market	Ponad 3000 źródeł online
2018	Stephen Kelly, Khurshid Ahmad	Estimating the impact of domain-specific news sentiment on financial assets	Wall Street Journal, Financial Times
2018	Adam Atkins, Mahesan Niranjan, Enrico Gerding	Financial news predicts stock market volatility better than close price	archiwum wiadomości Reuters US
2019	Charles W. Calomiris, Harry Mamaysky	How news and its context drive risk and returns around the	archiwum Thomson Reuters Machine Readable News
2019	David E. Allena, Michael McAleerd, Abhay K. Singh	Daily market news sentiment and stock prices	Thomson Reuters News Analytics
2020	Xiaodong Li, Pangjing Wu, Wenpeng Wang	Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong	Finiet News
2020	Alan J. Hanna, John D. Turner, Clive B. Walker	News media and investor sentiment during bull and bear markets	Financial Times Historical Archive
2021	Sabri Boubaker, Zhenya Liu, Ling Zhai	Big data, news diversity and financial market crash	Financial Times
2022	Jianping Li, Guowen Li, Mingxi Liu, Xiaoqian Zhu, Lu Wei	A novel text-based framework for forecasting agricultural futures using massive online news headlines	Hexun.com
2022	Yu He, Linshan Qu, Ran Wei, Xuankai Zhao	Media-based investor sentiment and stock returns: a textual analysis based on newspapers	China Securities Journal, Shanghai Securities News, Securities Times and Securities Daily
2022	Saeede Anbaee Farimani, Majid Vafaei Jahan, Amin Milani Fard, Seyed Reza Kamel Tabbakh	Investigating the informativeness of technical indicators and news sentiment in financial market price prediction	FXStreet, NewsBTC, Cointelegraph, Investing.com

Źródło: opracowanie własne.

Nie istnieją również sposoby weryfikacji tego, czy post zamieszczony został w Internecie przez osobę dobrze poinformowaną oraz czy jej autorem nie jest ktoś, kogo intencją jest zagłuszanie rzeczywistej oceny sytuacji (ang. *noise traders*)⁵⁴². Dodatkowo, właściwością tego typu komunikacji jest to, że przepełniona jest ona kolokwializmami, sarkazmem, znakami graficznymi, czy też skrótami myślowymi. Z punktu widzenia fazy przedprocesowej, to jest na przykład oczyszczania danych tekstowych, jak również późniejszej analizy, posty internetowe stanowią największe wyzwanie⁵⁴³. Dotychczas prowadzone w tym zakresie badania opierały się przede wszystkim o trzy źródła: Twitter, Facebook oraz fora internetowe, służące do komunikacji między inwestorami (ang. *stock boards*). Do tej kategorii autorka zalicza również wyniki wyszukiwarek internetowych, jako że są one pochodną działań podobnych w swoim charakterze do aktywności w mediach społecznościowych.

Na podstawie powyższych rozważań, zauważyć można, że każda z kategorii źródeł danych tekstowych, wykorzystywanych w dziedzinie finansów, niesie za sobą pewne trudności, które uwydatniają się podczas samego procesu badawczego albo mogą mieć potencjalny wpływ na efektywność wyników. Zdaniem autorki największe nadzieje można jednak wiązać z analizą tekstów publikowanych w prasie. Z jednej strony bowiem, obecnie możliwe jest zgromadzenie tego typu danych na temat właściwie każdej giełdy. Z drugiej strony, możliwe jest również pobieranie tekstów publikowanych przez dziennikarzy znajdujących się poza obszarem geograficznym spółek, będących podmiotami badania. Tworzy to przestrzeń do uchwycenia – pod względem narracyjnym oraz emocjonalnym – możliwie szerokiej perspektywy. Ponadto przyjąć można, że artykuły prasowe są wynikiem pracy osób obserwujących rynek, a więc dobrze poinformowanych. Pozwala to uniknąć obaw o to, że klasyfikacja sentymentu zaburzona zostanie przez osoby, chcące – świadomie lub nie – wpłynąć na zmianę emocji graczy rynkowych. Z badawczego punktu widzenia, co warto dodać, stwierdzić można, że liczba publikowanych obecnie artykułów prasowych stanowi pewną gwarancję możliwości skonstruowania próby badawczej. W związku z powyższym, w części empirycznej wykorzystane zostaną właśnie teksty prasowe.

⁵⁴² Kearney, C., Liu, S., *Textual sentiment in finance...*, op.cit., s. 174.

⁵⁴³ Loughran, T., McDonald, B., *Textual Analysis in Accounting...*, op. cit., s. 43.

5.2. Wykrywanie nadużyć

Zastosowanie tekstu do analizy i prognozowania zdarzeń o charakterze finansowym rozumiane może być bardzo szeroko. Dla technik tych zastosowanie znajduje się w tematyce biznesowej i finansowej, wykraczającej poza problem dynamiki cen na rynku kapitałowym. Eksploracja tekstu wykorzystywana jest również na przykład w bankowości, w związku z szacowaniem ryzyka kredytowego lub do wykrywania defraudacji i oszustw finansowych⁵⁴⁴. Można powiedzieć, że ta problematyka od zawsze związana była z eksploracją danych, w tym danych tekstowych. Działo się tak jeszcze zanim termin Big Data przedostał się do powszechnej świadomości. Barierą, przed którą stawali badacze, były jednak ograniczenia technologiczne, które uniemożliwiały wykorzystywanie narzędzi, pozwalających na wykrywanie nadużyć na dużych zbiorach danych⁵⁴⁵.

Wydaje się, że rozwój programowania przyczynił się do zwiększenia poziomu wiedzy właśnie na temat wykrywania nadużyć, powodowanych zarówno przez managerów spółek, jak i pracowników. Przeprowadzony przegląd literatury wskazuje na to, że w tym pierwszym przypadku, najsilniej eksploatowanym źródłem wiedzy są sprawozdania finansowe. W ogólnym ujęciu, oszustwa związane ze sprawozdawczością finansową polegają na takim prezentowaniu sytuacji spółki, aby bezpodstawnie wydawała się bardziej atrakcyjna niż jest w rzeczywistości⁵⁴⁶. Często takie działanie wynika z chęci ukrycia faktycznego stanu finansowego przedsiębiorstwa lub osiągnięcia przez nie korzyści kosztem akcjonariuszy. Z perspektywy przedstawianych w części empirycznej badań, szczególnie istotna jest ta druga motywacja. Można bowiem przyjąć, że oficjalnie publikowane przez korporacje dokumenty, postrzegane są jako dane o wysokim stopniu wiarygodności. Na ich podstawie inwestorzy mogą więc sformułować

⁵⁴⁴ Zob. Ghailan, O., Mokhtar, H., Hegazy, O., *Improving Credit Scorecard Modeling Through Applying Text Analysis*, „International Journal of Advanced Computer Science and Applications”, 2016, nr 7/4, s. 516.

Zob. Thomson Reuters, *Starmine Quantitative Models*, http://professional.starmine.com/static_content/StarMine_Text_Mining_Credit_Risk_Model_overview.pdf. Dostęp: 20 sie. 2017.

⁵⁴⁵ Sinayobye, J. O., Kiwanuka, F., Kaawaase Kyanda S., *A state-of-the-art review of machine learning techniques for fraud detection research*, ACM/IEEE Symposium on Software Engineering in Africa 2018 Gothenburg, Sweden, s.11.

⁵⁴⁶ West, J., Bhattacharya, M., *Intelligent financial fraud detection: A comprehensive review*, „Computers and Security”, 2016, nr 57, s. 51.

pozytywny obraz spółki, a przez to zawyżyć swoje oczekiwania co do jej przyszłej wartości⁵⁴⁷.

W 2005 roku Camilla Magnusson, Antti Arppe, Tomas Eklund i Barbro Back zwrócili uwagę na to, że dane tekstowe pochodzące ze sprawozdań finansowych, mogą być związane, lub nawet poprzedzać, niekorzystne zmiany wyników finansowych. Ich zdaniem, gdy otoczenie firmy oraz jej wewnętrzny potencjał wydajności ulegają zmianie, znajduje to odzwierciedlenie w sposobie myślenia kierownictwa firmy, co z kolei przekłada się na treść raportów, a z pewnym opóźnieniem również na wyniki spółki⁵⁴⁸. Jednocześnie, Sunita Goel, Jagdish Gangolly, Sue R. Faerman oraz Ozlem Uzuner zauważają, że wiele wczesnych badań na temat wykrywania oszustw pomijało dane jakościowe, będące przecież nieodzownym elementem raportów przedsiębiorstw. Badacze ci stwierdzili, że o ile tego typu źródła nie mogą zawierać bezpośrednich znaków fałszerstw, to eksploracja danych może ujawnić wskazujące na te zmienne, które maskowane są za pomocą semantyki i składni, a więc zjawisk właściwych językowi naturalnemu⁵⁴⁹.

Wspomniani autorzy przeprowadzili w 2010 roku jedno z pierwszych badań związanych z tą tematyką. Ich celem było zbudowanie modelu predykcyjnego, który pozwalałby na wykrywanie oszustw. Do badania wykorzystali roczne raporty amerykańskich spółek publicznych, zebrane z 14 lat (1993–2006). Trafność zaproponowanego przez nich modelu przekraczała 87%. Co jednak ciekawsze, wskazali oni systematyczne cechy jakościowe, występujące między dwiema klasami sprawozdań – spółek, u których występowało zjawisko defraudacji, oraz spółek, które nie były obarczone takim problemem. Cechy te związane były ze sposobem komunikacji, czy też stylem pisania. Okazało się, że raporty spółek dopuszczających się oszustw, charakteryzowały się częstszym stosowaniem strony biernej oraz zwrotów przypuszczających, jak również wyższym wskaźnikiem różnorodności leksykalnej oraz

⁵⁴⁷ Ravisankar, P., Ravi, V., Raghava Rao, G., Bose, I., *Detection of financial statement fraud and feature selection using data mining techniques*, „Decision Support Systems Journal”, 2011, nr 50, s. 491–492.

⁵⁴⁸ Zob. Magnusson, C., Arppe, A., Eklund, T., Back, B., Vanharanta, H., Visa, A., *The language of quarterly reports as an indicator of change in the company's financial status*, „Information & Management”, 2005, nr 42, s. 561–574.

⁵⁴⁹ Goel, S., Gangolly, J., Faerman, S.R., Uzuner, O., *Can Linguistic Predictors Detect Fraudulent Financial Filings?*, „Journal of Emerging Technologies in Accounting”, 2010, nr 7, s. 26.

niższym poziomem indeksu czytelności⁵⁵⁰. W tym samym roku Mark Cecchini, Haldun Aytug, Gary J. Koehler oraz Praveen Pathak zaprezentowali model wczesnego wykrywania dwóch typów zdarzeń: bankructwa oraz defraudacji. Zastosowanie samych danych tekstowych pozwoliło im osiągnąć trafność na poziomie 75% w przypadku defraudacji oraz 80% w przypadku bankructw. Dalsza eksploracja, polegająca na połączeniu w modelu danych finansowych oraz tekstowych, pozwoliła w obu przypadkach osiągnąć trafność przekraczającą 80%⁵⁵¹. Rok później, Tim Loughran i Bill McDonald przeprowadzili analizy raportów 10-K w celu sprawdzenia czy przedsiębiorstwa, które zostały oskarżone o nieprawidłowości księgowe lub zgłosiły uchybienia podczas kontroli wewnętrznych, używają innego języka niż firmy, które nie miały takich doświadczeń. Stworzona przez nich lista słów okazała się pomocna w identyfikacji organizacji o niestandardowych zachowaniach⁵⁵². Również w 2011 roku Fletcher H. Glancy i Surya B. Yadav, opierając się na raportach SEC⁵⁵³, zaproponowali model CFDM (*Computational Fraud Detection Model*). Wyniki zastosowanego przez nich klastrowania pozwoliły na zidentyfikowanie około 90% sprawozdań z klasy „defraudacja” oraz 80% sprawozdań z klasy „nie-defraudacja”. Na podstawie swoich analiz, badacze stwierdzili, że zdolność ich modelu do wykrywania nadużyć oznaczać może, że autorzy raportów byli pod wpływem stresu, który przekładał się na używany przez nich język. Może to oznaczać, że byli oni świadomi lub przynajmniej mieli pewien pogląd na temat tego, że opisywana przez nich sytuacja spółki nie znajduje potwierdzenia w rzeczywistości⁵⁵⁴. Podobnie w 2015 roku, Lynnette Purda i David Skillicorn zaproponowali metodę, której podstawą jest analiza języka, a celem predykcja defraudacji na podstawie danych, pochodzących z raportów 10-K. Do badań wykorzystali 200 wyrażen, które charakteryzowały się największą korelacją z działaniami malwersacyjnymi, a kolejno posegregowali raporty na dwie grupy: sfalszowane oraz

⁵⁵⁰ Zob. Goel, S., Gangolly, J., Faerman, S.R., Uzuner, O., *Can Linguistic Predictors Detect...*, op.cit., s. 25–46.

⁵⁵¹ Zob. Cecchini, M., Aytug, H., Koehler, G. J., Pathak, P., *Making words work: Using financial text as a predictor of financial events*, „Decision Support Systems”, 2010, nr 50, s. 164–175.

⁵⁵² Loughran, T., McDonald, B., *Barron's Red Flags: Do They Actually Work?*, „Journal of Behavioral Finance”, 2011, nr 12/2, s. 90–97.

⁵⁵³ Komisja Papierów Wartościowych i Giełd Stanów Zjednoczonych (ang. *United States Securities and Exchange Commission*)

⁵⁵⁴ Zob. Glancy, F. H., Yadav, S. B., *A computational model for financial reporting fraud detection*, „Decision Support Systems”, 2011, nr 50, s. 595–601.

prawdziwe. Używając zbudowanego modelu, osiągnęli trafność klasyfikacyjną na poziomie 82%⁵⁵⁵. Z kolei Wei Dong, Shaoyi Liao oraz Liang Liang zaproponowali metodę FSF (*Financial Statement Fraud Detection Using Text Mining*). W wyniku jej zastosowania osiągnięta została średnia zdolność wykrywania nadużyć przekraczająca 82%⁵⁵⁶.

Należy jednak zaznaczyć, że nie wszystkie badania potwierdzają możliwość wykorzystania zmiennych tekstowych do wykrywania defraudacji. Przykładowo Petr Hajek i Roberto Henriques, posilając się swoimi wynikami badań, stwierdzili, że zastosowanie danych tekstowych nie pozwoliło na zbudowanie satysfakcjonującego modelu predykcyjnego. Co więcej, wyłączenie z całego zbioru zmiennych, tych o charakterze lingwistycznym, nie zmniejszyło w znaczącym stopniu zdolności ich modelu do prawidłowej klasyfikacji spółek⁵⁵⁷. Podobnie, Chandra S. Throckmorton, William J. Mayew, Mohan Venkatachalam i Leslie M. Collins, w ramach swoich badań wykazali, że większość zmiennych lingwistycznych nie dawała możliwości wiarygodnego wskazania, które z badanych przez nich spółek dopuściło się oszustw⁵⁵⁸. Niejednoznaczność w wynikach przytaczanych badań, wskazuje na konieczność dalszej eksploracji zagadnienia. Zaproponowanie systemu klasyfikacji sprawozdań z wykorzystaniem informacji tekstowych mogłoby bowiem znaleźć szerokie zastosowanie zarówno w kontekście inwestycyjnym, jak i nadzorczym⁵⁵⁹.

Powyższe przykłady związane były z możliwością wykorzystania analizy tekstu do predykcji oszustw na poziomie menedżerskim. Podobne próby podejmowane są w kontekście wykrywania nadużyć popełnianych przez pracowników. W 2009 roku Carolyn Holton przeprowadziła badania, których celem było zidentyfikowanie wiadomości mailowych, zawierających w sobie treści wskazujące na niezadowolenie

⁵⁵⁵ Purda, L., Skillicorn, D., *Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection*, „Contemporary Accounting Research”, 2014, nr 32/3, s. 1193–1223.

⁵⁵⁶ Dong, W., Liao, S., Liang, L., *Financial Statement Fraud Detection using Text Mining: a systemic functional linguistics theory perspective*, PACIS 2016, Proceedings 2016, Chiayi, Taiwan, s. 188.

⁵⁵⁷ Zob. Hajek, P., Henriques, R., *Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods*, „Knowledge-Based Systems”, 2017, nr 128, s. 139–152.

⁵⁵⁸ Zob. Throckmorton, Ch. S., Mayew, W. J., Venkatachalam, M., Collins, L. M., *Financial fraud detection using vocal, linguistic and financial cues*, „Decision Support Systems Journal”, 2015, nr 74, s. 78–87.

⁵⁵⁹ Cecchini, M., Aytug, H., Koehler, G.J., Pathak, P., *Making words work...*, op.cit., s. 174.

pracowników z ich sytuacji zawodowej. Autorka, bazując na psychologicznych determinantach skłonności do dokonywania oszustw⁵⁶⁰, założyła, że sposób komunikacji tego typu pracowników, charakteryzować będzie się osobliwymi zmiennymi o charakterze jakościowym. Wyniki jej badania wskazywały na to, że możliwe jest algorytmiczne klasyfikowanie maili z podziałem na dwie klasy: zadowoleni i niezadowoleni. Badaczka postulowała, żeby taki sposób analizy włączyć w standardową procedurę audytu, stwierdzając, że mogłoby to potencjalnie przyczynić się do zmniejszenia liczby nadużyć lub ich wcześniejszego wykrywania⁵⁶¹. Niewątpliwie tę koncepcję wykorzystało przedsiębiorstwo Ernst&Young, które we współpracy z FBI zaproponowało metodę wykrywania oszustw pracowniczych za pomocą wyszukiwania fraz w treści komunikacji mailowej⁵⁶². Pracownicy Ernst&Young, rozwijając metodykę analizy tekstu oraz zawartego w nim sentymentu, skonstruowali również model FDA (*Forensic Data Analytics*). Jak sami autorzy twierdzą:

eksploracja tekstu, analiza sentymentu i mapowanie pojęć mogą być użyteczne w sprawach związanych z oszustwami, korupcją, marnotrawstwem i nadużyciami w celu identyfikowania transakcji, e-maili i danych tekstowych, które posiadają niepokojące sformułowania i frazy lub wykrywania emocji związanych z danymi⁵⁶³.

Z uwagi na to, że obecnie firmy zwykle konsumują i produkują dane o znaczących rozmiarach, konieczne wydaje się wdrażanie procesów ciągłego monitorowania, aby móc w odpowiednim momencie zidentyfikować anomalie w strumieniu danych lub wzorce zachowań, które niosą ze sobą potencjał oszustwa⁵⁶⁴. Z badawczego punktu widzenia, tego typu analizy obarczone są jednak podstawowym ograniczeniem, związanym z faktem, że do ich przeprowadzenia niezbędne są wewnętrzne dane spółek, często

⁵⁶⁰ ang. *Fraud triangle employee disgruntlement drivers*.

⁵⁶¹ Zob. Holton, C., *Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem*, „Decision Support Systems Journal”, 2009, nr 46, s. 853–864.

⁵⁶² Na przykład: *cover up, write off, failed investment, off the books, nobody will find out, grey area*. Donohue, B., *How to Fail at Corporate Fraud*, <https://threatpost.com/how-fail-corporate-fraud-010813/77378/>. Dostęp: 1 kw. 2023.

⁵⁶³ Ernst&Young, *Forensic data analytics – Globally integrated compliance review, litigation support and investigative services*, [http://www.ey.com/Publication/vwLUAssets/EY_-_Forensic_Data_Analytics_\(FDA\)/\\$FILE/EY-forensic-data-dnalytics.pdf](http://www.ey.com/Publication/vwLUAssets/EY_-_Forensic_Data_Analytics_(FDA)/$FILE/EY-forensic-data-dnalytics.pdf). Dostęp: 20 sie. 2017, s. 7.

⁵⁶⁴ Banarescu, A., *Detecting and Preventing Fraud with Data Analytics*, „Procedia Economics and Finance.”, 2015, nr 32, s. 1827.

definiowanych jako dane wysokiej wrażliwości. W efekcie ich pozyskanie – przez zewnętrznego badacza – jest niezwykle trudne⁵⁶⁵.

5.3. Analiza ryzyka

Eksploracja tekstu zyskała również popularność w ramach zagadnień związanych z analizą ryzyka. Wydaje się, że w ramach tej problematyki, jest ona stosowana w dwóch kluczowych obszarach. Pierwszy związany jest z bankowością, drugi natomiast z rynkiem finansowym. Jako przykład pierwszego zastosowania podać można badania, które przeprowadzili Omar Ghailan, Hoda M. O. Mokhtar, oraz Osman Hegazy, wykorzystując formularze aplikacyjne wniosków o pożyczki instytucji finansowych w Jemenie. Doszli oni do wniosku, że włączenie danych tekstowych do analizy scoringowej poprawia trafność modelu kredytowego, podnosząc liczbę prawidłowych klasyfikacji. Dzięki zastosowaniu ich metody, ogólny wynik zagregowanego wskaźnika zwiększył się z 76% do 93%⁵⁶⁶. Innym przykładem jest model zaproponowany przez pracowników Thomson Reuters o nazwie TMCR (*The StarMine Text Mining Credit Risk Model*). Jest to komercyjny model analizy ryzyka kredytowego, bazujący na specjalistycznych słownikach związanych z analizą kondycji finansowej podmiotów gospodarczych. Za pomocą TMCR, ryzyko w spółkach publicznych oceniane jest poprzez ewaluację języka, używanego w Reuters News, transkrypcjach połączeń konferencyjnych, raportach spółek oraz wyszukiwaniach brokerów. TMCR służyć może zarówno do wykrywania nieprawidłowości, jak i predykcji bankructwa⁵⁶⁷.

Przykładem analizy ryzyka w ramach zagadnień związanych z rynkiem finansowym są badania, które przeprowadzili Qing Li, TieJun Wang, Ping Li, Ling Liu, Qixu Gong oraz Yuanzhu Chen. Badając *sentiment ryzyka*, stworzyli wskaźnik, którego relacja okazała się odwrotnie proporcjonalna do ceny akcji. Oznaczało to, że przedsiębiorstwa obciążone wysokim poziomem tego wskaźnika, wykazywały niskie wyniki stóp zwrotu z akcji. Natomiast przedsiębiorstwa, dla których wskaźnik

⁵⁶⁵ Sinayobye, J. O., Kiwanuka, F., Kaawaase Kyanda S., *A state-of-the-art review...*, op.cit., s. 12.

⁵⁶⁶ Ghailan, O., Mokhtar, H., Hegazy, O., *Improving Credit Scorecard Modeling...*, op.cit., s. 516.

⁵⁶⁷ Thomson Reuters, *StarMine Quantitative Models*,

http://professional.starmine.com/static_content/StarMine_Text_Mining_Credit_Risk_Model_overview.pdf.
Dostęp: 20 sie. 2017.

sentymetu ryzyka nie był wysoki – odwrotnie⁵⁶⁸. Z kolei Zhi Da, Joseph Engelberg oraz Pengjie Gao, badając relację między zapytaniami w wyszukiwarkach internetowych a cenami akcji, zbudowali dość popularny indeks FEARS (*Financial and Economic Attitudes Revealed by Search Index*). W tym celu zagregowali miliony wyszukiwań takich słów, jak: „recesja”, „bankructwo” czy „bezrobocie”. Jak się okazało, zastosowanie tego indeksu umożliwiło predykcję dziennych zwrotów na analizowanych rynkach. Podobnie jak w przypadku badań Qing Li i zespołu, wysoki poziom indeksu FEARS skorelowany był z niskimi zwrotami z danego dnia, a jednocześnie prognozował wysokie zwroty w kolejnych kilku dniach. Wysoki poziom indeksu wiązał się również z tymczasowym wzrostem wahań rynkowych oraz odpływem funduszy z rynku akcji do rynku obligacji⁵⁶⁹. Podobna logika przyświecała twórcom University of Michigan Social Media Job Loss Index, których celem było przewidzenie poziomu roszczeń z tytułu zasiłku dla bezrobotnych (ang. *unemployment insurance*). Co ciekawe, indeks ten opiera się na frazach wpisywanych przez użytkowników Twittera⁵⁷⁰. Próbę zastosowania indeksów opierających się na mediach społecznościowych w celach predykcyjnych podjęli także Antonios Siganos, Evangelos Vagenas-Nanos oraz Patrick Verwijmeren. Do swoich badań wykorzystywali oni wskaźnik FGNHI (*Facebook's Gross National Happiness Index*)⁵⁷¹, który reprezentował codzienne natężenie emocji użytkowników platformy Facebook. Badanie polegało na analizie relacji między sentymentem a poziomem stóp zwrotu z akcji. Autorzy wykazali, że w ramach krótkoterminowych zwrotów, relacja ta ma charakter pozytywny, a amplituda sentymentu skorelowana jest czasowo z poziomami indeksów giełdowych⁵⁷².

W powyższych przykładach, metody eksploracji tekstu służą do badania niedostrzegalnych na pierwszy rzut oka informacji dotyczących szeroko rozumianej

⁵⁶⁸ Li, Q., Wang, T., Li, P., Liu, P., Gong, Q., Chen, Y., *The effect of news and public mood on stock movements*, „Information Sciences”, 2014, nr 278, s. 827.

⁵⁶⁹ Zob. Da, Z., Engelberg, J., Gao, P., *The sum of all FEARS: Investor Sentiment and Asset Prices*, „Review of Financial Studies”, 2015, nr 28/1, s. 1–32.

⁵⁷⁰ American Association for Public Opinion Research, *AAPOR Report on Big Data*, AAPOR Big Data Task Force, 2015, s. 7.

⁵⁷¹ Indeks nie jest już dostępny.

⁵⁷² Siganos, A., Vagenas-Nanos, E., Verwijmeren, P., *Facebook's daily sentiment and international stock markets*, „Journal of Economic Behavior & Organization”, 2014, nr 107/B, s. 731.

sytuacji finansowej, zarówno przedsiębiorstw, jak i całych rynków⁵⁷³. Należy jednak zaznaczyć, że jest to wciąż wschodzący obszar finansów i efekty opartych na nim technik wciąż mogą okazywać się niejasne lub nieprecyzyjne⁵⁷⁴. Pomimo tego, na obecnym etapie wiedzy istnieją przesłanki by sądzić, że za pomocą analizy tekstu można uzyskać statystycznie istotne wyniki oceny poziomu ryzyka w domenie finansowej.

5.4. Zastosowanie analizy sentymentu, pochodzącego z publikacji prasowych, do badań na rynku finansowym

Próby odnalezienia związków między danymi tekstowymi a dynamiką rynku podejmowane są również w kontekście indeksów giełdowych. Głównym powodem, dla którego są one podejmowane, jest istnienie instrumentów finansowych opartych na wartościach owych indeksów, takich jak instrumenty pochodne czy ETF-y. Koncepcję wykorzystania narzędzi związanych z eksploracją tekstu do analiz, umiejscowionych w tym właśnie kontekście, zastosowano zarówno na rynku amerykańskim, jak i europejskim, w tym polskim.

Wydaje się, że jedno z pierwszych badań w tym zakresie przeprowadzone zostało przez Beata Wuthricha z zespołem jeszcze w 1998 roku. Autorzy zaprezentowali model klasyfikacji danych tekstowych i predykcji ruchów pięciu indeksów: Dow Jones, FT-SE 100, Nikkei 225, Hang Seng oraz Singapore Straits. Celem modelu było zarekomendowanie działania, które podjąć powinien potencjalny inwestor, to jest: „kup”, „sprzedaj”, „nie rób nic”. Badacze weryfikowali trafność modelu w interwałach dziennych. Jednak po zakończeniu badań stwierdzono trafność decyzyjną modelu na poziomach od 40% (Singapore Straits) do 47% (FT-SE 100)⁵⁷⁵. Biorąc jednak pod uwagę, że autorzy podzielili możliwe trendy na trzy kategorie (wzrost, spadek, brak ruchu), osiągnięty przez model wynik jest wyższy niż potencjalne 33%. Chociaż uzyskane wyniki trudno uznać za satysfakcjonujące, to należy mieć na uwadze, że próba

⁵⁷³ Por. Wang, B., Huang, H., Wang, X., *A novel text mining approach to financial time series forecasting*, „Neurocomputing”, 2012, nr 83, s. 136–145.

⁵⁷⁴ Loughran, T., McDonald, B. *Textual Analysis in Accounting...*, op.cit., s. 1189.

⁵⁷⁵ Zob. Wuthrich, B., Cho, V., Leung, S., Permunetilleke, D., Sankaran, K., Zhang, J., Lam, W., *Daily stock market forecast from textual web data*, „Conference Paper”, San Diego CA, USA 1998.

ta podjęta została pod koniec lat 90., a sam okres badawczy określić można jako relatywnie krótki (12.1997–03.1998)⁵⁷⁶.

Badaczem, który niewątpliwie zaznaczył swój wkład w kontekście analizy związków między sentymentem tekstowym a zmiennymi finansowymi, jest Paul Tetlock. W 2007 roku opublikował on artykuł *Giving Content to Investor Sentiment: The Role of Media in the Stock Market Published*, w którym eksplorował relację, jaka zachodzi między negatywnym sentymentem generowanym z jednej z kolumn „The Wall Street Journal”, *Abreast of the Market*, a wynikami giełdowymi. Jednym z celów jego badania było sprawdzenie, czy pesymizm medialny może prognozować dzienne stopy zwrotu indeksu Dow Jones Industrial Average. Okazało się, że średnio, zmiana pesymizmu medialnego o jedno odchylenie standardowe, wpływała na zmianę stóp zwrotu z indeksu z następnego dnia o 8,1 punktów bazowych (czyli więcej niż bezwarunkowa stopa zwrotu, która wynosiła 5,4 punktu bazowego). Udało mu się jednocześnie zaobserwować trzy zjawiska. Po pierwsze, odrzucił hipotezę mówiącą o tym, że treści pochodzące z mediów, zawierają nowe informacje na temat fundamentalnych wartości aktywów, oraz że nie mają one związku z rynkiem akcji, a są jedynie jego efektem ubocznym. Po drugie, stwierdził, że niestandardowo wysokie lub niskie poziomy pesymizmu mogą być postrzegane jako prognoza wysokiego wolumenu obrotów. Po trzecie, jego zdaniem pesymizm medialny wyprzedza spadki cen, po których następuje rewersja do ceny fundamentalnej⁵⁷⁷.

Inne przeprowadzone w tym zakresie badanie, zaprezentowane zostało przez Sahila Zubaira i Krzysztofa J. Ciosa w 2015 roku. Przedmiotem analiz był indeks S&P 500. Badanie obejmowało okres siedmiu lat (2007–2013). Źródła tekstowe pochodziły z archiwum Reutersa. Celem badania było zweryfikowanie hipotezy, mówiącej o występowaniu korelacji między wynikami indeksu a sentymentem tekstowym. Sentyment podzielony został na dwie kategorie: pozytywny i negatywny, a do jego klasyfikacji wykorzystano słownik Harvard General Inquirer⁵⁷⁸. Uzyskane wyniki wskazywały na występowanie korelacji między negatywnym sentymentem a dynamiką

⁵⁷⁶ Mittermayer, M.-A., Knolmayer, G. F., *Text mining systems for market response to news: a survey* Institute of Information Systems, University of Bern, „Working Paper”, 2006, nr 184, s. 4.

⁵⁷⁷ Zob. Tetlock, P. C., *Giving Content to Investor Sentiment...*, op.cit., s. 1139–1168.

⁵⁷⁸ Zob. General Inquirer, <https://inquirer.sites.fas.harvard.edu/>. Dostęp: 1 kw. 2023.

indeksu (dotyczyło to pięciu z siedmiu badanych lat). Natomiast w przypadku sentymentu pozytywnego nie zauważono takiej właściwości⁵⁷⁹. Inne badanie przeprowadzone na rynku amerykańskim przedstawione zostało przez Branka Kavšeka w roku 2017. Istotą badania nie było jednak wykazanie korelacji między sentymentem a wynikami indeksu (w tym przypadku Dow Jones Industrial Average), lecz wskazanie słów, które mają związek z ruchami rynku. W efekcie badania udało się stworzyć listę zwrotów, charakteryzujących się wysokim poziomem korelacji ze spadkami oraz wzrostami indeksu⁵⁸⁰.

Analogiczne badania prowadzone były również na rynku europejskim. W 2018 roku Stefan Feuerriegel i Julius Gordon podjęli próbę zaprognozowania zmian w poziomach indeksów: DAX, CDAX oraz STOXX. W odróżnieniu od wcześniejszych badaczy, Feuerriegel i Gordon zainteresowani byli możliwościami wykorzystania koncepcji eksploracji tekstu w analizach o charakterze długoterminowym. Uwzględnione przez nich teksty pochodziły ze sprawozdań finansowych z lat 1997–2015. Efektywność prognoz różniła się w zależności od takich czynników jak: horyzont prognozy, model prognostyczny, czy też dane wejściowe. Chociaż wydajność prognoz opartych na danych tekstowych nie we wszystkich eksperymentach była satysfakcjonująca, to badaczom udało się wskazać przypadki, w których włączenie tego typu danych do modeli, pozwoliło obniżyć błędy prognozy. W szczególności dotyczyło to właśnie długoterminowych prognoz badanych indeksów⁵⁸¹.

Wykorzystanie eksploracji tekstu do analizy ruchów indeksów, testowane było również na rynku polskim. W 2018 roku Katarzyna Rostek wraz z Piotrem Młodzianowskim opublikowali badania dotyczące indeksów: WIG, WIG20, mWIG40 oraz sWIG80. Okres badawczy obejmował 280 dni. Sentyment definiowany był zgodnie z założonymi przez autorów klasami słów kluczowych. Do sentymentu pozytywnego przypisano takie słowa, jak: „hossa”, „rośnie”, „byk”, „zielony”, „zysk”, „ożywienie”,

⁵⁷⁹ Zob. Zubair, S., Cios, K. J., *Extracting News Sentiment and Establishing its Relationship with the S&P500 Index*, 48th Hawaii International Conference on System Sciences Extracting, 2015 Kauai, HI, USA.

⁵⁸⁰ Zob. Kavšek, B., *Using Words from Daily News Headlines to Predict the Movement of Stock Market Indices*, „Managing Global Transitions”, 2017. nr 15/2, s. 109–121.

⁵⁸¹ Zob. Feuerriegel, S., Gordon, J., *Long-term stock index forecasting based on text mining of regulatory disclosures*, „Decision Support Systems”, 2018, nr 112, s. 88–97.

natomiast dla sentymentu negatywnego były to: „bessa”, „spada”, „niedźwiedź”, „czerwony”, „strata”, „kryzys”. Zastosowanie zaproponowanych słów pozwoliło na trafne sklasyfikowanie ruchów rynku (wzrost / spadek) na poziomie od 61% do 68% w zależności od indeksu. Próba zbudowania na tej podstawie modelu prognostycznego nie pozwoliła jednak na osiągnięcie satysfakcjonujących rezultatów. Współczynnik trafności prognozy wahał się w tym przypadku między wynikiem 48% a 59%⁵⁸².

Ameryka Północna

Jednym z pierwszych modeli zbudowanych w oparciu o eksplorację tekstu – wykorzystujących artykuły prasowe – których przeznaczeniem była predykcja cen akcji, był model NewsCATS (*News Categorization and Trading System*). Jego celem było sklasyfikowanie trendu cenowego, który może pojawić się zaraz po ujawnieniu informacji. Model składał się z trzech komponentów. Pierwszy pobierał i wstępnie przetwarzał odpowiednie komunikaty prasowe. Drugi sortował komunikaty zgodnie z wcześniej zdefiniowanymi klasami: „dobre wiadomości”, „złe wiadomości”, „wiadomości niegenerujące ruchu” (ang. *no movers*). Trzeci komponent miał za zadanie dobrać adekwatną strategię inwestycyjną. Niestety efektywność algorytmu klasyfikującego nie była wysoka. Większość prawidłowo zaklasyfikowanych informacji dotyczyła kategorii wiadomości, które nie wygenerują ruchu. Paradoksalnie jednak algorytm odpowiedzialny za dobór decyzji inwestycyjnej – na podstawie pojawiania się komunikatów prasowych – pozwolił osiągnąć średni zysk na poziomie 0,11% i był on wyższy niż zysk osiągnięty przez najlepszego losowego inwestora (ang. *random trader*). Na podstawie swoich analiz autor badania wyciągnął wniosek, że prawdopodobieństwa w podążaniu – przez ceny akcji – ścieżką zgodną z modelem błędzenia losowego nie są takie same zaraz po pojawieniu się na rynku informacji prasowej oraz że właściwość ta wyjaśniona może być wyłącznie poprzez pojawienie się komunikatu prasowego⁵⁸³.

W 2009 roku Robert Schumaker i Hsinchun Chen opublikowali wyniki badania przeprowadzonego na spółkach notowanych w ramach indeksu S&P500. Autorzy

⁵⁸² Zob. Rostek, K., Młodzianowski, P., *Współzależność informacji sieciowych oraz zmian indeksów zachodzących na Gieldzie Papierów Wartościowych w Warszawie*, „Zeszyty Naukowe Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach”, 2017, nr 115, s. 249–263.

⁵⁸³ Zob. Mittermayer, M-A., *Forecasting Intraday Stock Price Trends with Text Mining Techniques*, 37th Hawaii International Conference on System Sciences, Big Island, HI, USA 2004.

zwrócili uwagę, że większość prowadzonych w tym zakresie badań koncertowała się jedynie na ustaleniu kierunku zmian cen. W związku z tym, celem ich badania była estymacja bezpośredniej ceny akcji w krótkim okresie, to jest dwadzieścia minut po opublikowaniu artykułu. Do badania wykorzystano ponad dziewięć tysięcy artykułów oraz ponad dziesięć tysięcy dziennych notowań giełdowych. Uzyskane wyniki pozwoliły stwierdzić, że model wykorzystujący zarówno artykuły prasowe, jak i cenę akcji, w momencie publikacji artykułu osiągał najlepsze rezultaty w kontekście miary bliskości, dokładności kierunkowej oraz symulacji handlu. Zbudowany przez nich model nazwany został AZFinText (*Arizona Financial Text*)⁵⁸⁴. Kilka lat później narzędzie to zostało ponownie zastosowane, przy czym do dwóch wymienionych już badaczy dołączyli Yulei Zhang oraz Chun-Neng Huang. Autorzy zwrócili szczególną uwagę na wyniki związane z negatywnym sentymentem, twierdząc, że dzięki niemu najłatwiej było przewidzieć kierunek zmian cen i możliwe było osiągnięcie średniej stopy zwrotu na poziomie 3,3%⁵⁸⁵. Na wagę analizy negatywnego sentymentu w przekazach medialnych uwagę zwrócił już wcześniej Paul Tetlock, który zauważył, że wysoki poziom pesymizmu prognozuje spadkową presję na ceny rynkowe, a w kolejnej fazie ich rewersję do podstaw. Dodatkowo – jego zdaniem – nadzwyczaj wysokie lub niskie wartości pesymizmu związane są ze znaczącym wolumenem obrotów⁵⁸⁶.

Kolejnym indeksem, którego spółki podlegały analizie w ramach problematyki eksploracji tekstu, jest indeks NASDAQ. Badania w tym kontekście przeprowadzone zostały przez Ayman E. Khedr, S. E. Salama i Nagwa Yaseen. Ich celem było zbudowanie modelu, umożliwiającego przewidzenie trendów rynkowych. W modelu uwzględniono analizę nastrojów wiadomości finansowych oraz historyczne ceny notowanych w indeksie akcji trzech spółek. W pierwszym etapie badania wykorzystane zostały wyłącznie nastroje prasowe. Pozwoliło to uzyskać wyniki dokładności prognostycznej na poziomie od 73% do 86%. Włączenie – w następnym etapie – do modelu cen historycznych, co poprawiło dokładność niemal do 90%. Autorzy określili

⁵⁸⁴ Zob. Schumaker, R. P., Chen, H., *Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFinText System*, „ACM Transactions on Information Systems”, 2009, nr 27/2, s. 1–19.

⁵⁸⁵ Zob. Schumaker, R. P., Zhang, Y., Huang, Ch-N., Chen, H., *Evaluating sentiment in financial news articles*, „Decision Support Systems”, 2012, nr 53, s. 458–464.

⁵⁸⁶ Zob. Tetlock, P. C., *Giving Content to Investor Sentiment...*, op.cit., s. 1139–1168.

więc dane tekstowe jako zmienne o istotnym charakterze dla analizy dynamiki spółek, znajdujących się w indeksie⁵⁸⁷. Do podobnych wniosków doszli badacze, którzy testowali możliwość zastosowania tego typu badań, na spółkach notowanych w ramach indeksu Dow Jones. Wyniki badań na ten temat przedstawili David E. Allen, Michael McAleer i Abhay K. Singh. Warto jednak zauważyć, że skonstruowana przez nich próba uwzględniała wszystkie notowane w indeksie spółki. Autorzy wskazali, że nawet po uwzględnieniu czynników rynkowych, wyniki sentymentu tekstowego mają wciąż znaczący wpływ na stopy, mieszczących się w indeksie, spółek. Ponadto zauważono, że informacje nie zawsze znajdują natychmiastowe odzwierciedlenie w cenie walorów, a nawet reakcja z pięciodniowym opóźnieniem, wciąż – w niektórych przypadkach – okazywała się istotna statystycznie⁵⁸⁸.

Oczywiście nie wszystkie badania potwierdzają, że wykorzystywanie danych tekstowych, w celu opisanie dynamiki rynku, niesie za sobą dodatkową wartość. Analizę danych, charakteryzujących się krótkim interwałem, przeprowadzili Adam Atkins, Mahesan Niranjan i Enrico Gerding. Badaniom poddano wyniki dwóch indeksów: NASDAQ i Dow Jones, jak również dwóch spółek: Goldman Sachs oraz J. P. Morgan. W przypadku tych badań próby predykcji zamykały się w oknie 60 minut. Autorzy dostrzegli, że wskaźnik klasyfikacji F1 osiągał wyższe rezultaty w przypadku analizy zmienności wolumenu obrotów niż w przypadku cen zamknięcia. Badacze wykazali, w przeciwieństwie do wyników większości przeprowadzanych na ten temat badań, niezdolność modelu opartego na sentymencie do przewidywania cen walorów na poziomie wyższym niż 50%⁵⁸⁹.

Ameryka Południowa

W 2012 roku P.S.M. Nizer i J.C. Nievola zaprezentowali wyniki badań dotyczące giełdy brazylijskiej. Celem badań było stworzenie metody, umożliwiającej analizowanie wiadomości tekstowych, pisanych w języku portugalskim oraz sprawdzenie ich wpływu

⁵⁸⁷ Zob. Khedr, A. E., Salama, S. E., Yaseen, N., *Predicting stock market behavior using data mining technique and news sentiment analysis*, „International Journal of Intelligent Systems and Applications”, 2017, nr 7, s. 22–30.

⁵⁸⁸ Zob. Allen, D. E., McAleer, M., Singh, A. K., *Daily market news sentiment and stock prices*, „Applied Economics”, 2019, nr 51/30, s. 3212–3235.

⁵⁸⁹ Zob. Atkins, A., Niranjan, M., Gerding, E., *Financial news predicts stock market volatility better than close price*, „The Journal of Finance and Data Science”, 2018, nr 4, s. 120–137.

na ceny akcji. Próbkę stanowiły 33 spółki, notowane w indeksie BOVESPA w roku 2009, przy czym dane cenowe pobierane były w pięciominutowych interwałach. Warto zwrócić uwagę na sposób klasyfikowania danych tekstowych. Dzielone były one na dwie kategorie: interesujące i nieinteresujące. Obie grupy weryfikowane były na podstawie ponadnormatywnych stóp zwrotu. Jeśli wykryto reakcję rynku na dany artykuł, klasyfikowany był on jako „interesujący”. W ten sposób zbudowana została próba treningowa, na podstawie której uczony był algorytm klasyfikacyjny. Wyniki automatycznej klasyfikacji pozwoliły na osiągnięcie precyzji na poziomie 80%, jednak w kontekście czułości, wynik ten wynosił jedynie 20%. Mimo to autorzy uznali, że taki sposób segregacji danych tekstowych pozwaliby na zmniejszenie liczby artykułów, która niezbędna jest analitykowi do bieżącej pracy⁵⁹⁰. Indeks BOVESPA analizowany był również kilka lat później przez Tanya’a Araújo, Samuela Eleutério oraz Francisco Louçã. Próbkę badawczą zbudowano na podstawie 35 spółek notowanych w indeksie w latach 2007–2015, dla których obliczone zostały dzienne stopy zwrotu. Celem badania było sprawdzenie, czy nastroje wyrażone w wyspecjalizowanych mediach, mogą zostać uznane za zmienną objaśniającą dynamikę rynku. Autorzy udowodnili jednak, że to zmiany cen walorów rynkowych mają większy wpływ na sentyment przekazywany w mediach niż odwrotnie⁵⁹¹.

Azja

Azjatyccy badacze zwrócili uwagę na to, że inwestorzy często zainteresowani są nie tyle ceną pojedynczych akcji, ile wewnętrzną relacją, jaka zachodzi między trendami cenowymi. Znajomość takich zależności potencjalnie pozwoliłaby na przewidywanie ruchów posiadanych walorów poprzez monitorowanie wydarzeń dotyczących instrumentów powiązanych. Kierując się tym spostrzeżeniem Gabriel P.C. Fung i Jeffrey Xu Yu oraz Wai Lam przeprowadzili badanie, w którym podjęli próbę eksploracji i predykcji cen akcji, używając do tego wyłącznie danych tekstowych. Założenie poczynione na potrzeby tego badania brzmiało następująco: jeśli mamy do czynienia z akcjami X i Y, a zależność między nimi jest taka, że ilekroć rośnie cena akcji X,

⁵⁹⁰ Zob. Nizer, P. S. M., Nievola, J. C., *Predicting published news effect in the Brazilian stock market*, „Expert Systems with Applications”, 2012, nr 39, s. 10674–10680.

⁵⁹¹ Zob. Araújo, T., Eleutério, S., Louçã, F., *Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood*, „Physica A.”, 2018, nr 505, s. 1139–1149.

z pewnym opóźnieniem rośnie również cena akcji Y, to artykuły prasowe (zdarzenia), które wywoływałyby trend wzrostowy akcji X, będą przydatne dla przewidywania ruchów cenowych akcji Y. Przedmiotem badania były 33 spółki wyselekcjonowane z indeksu Hang Seng, do których przyporządkowano około 600 000 artykułów prasowych z okresu 01.10.2002–30.04.2003. Autorzy przeprowadzili trzy eksperymenty. W pierwszym monitorowano pojedyncze szeregi czasowe, w drugim zaimplementowano autorski system oparty na monitorowaniu dwóch szeregów czasowych, w trzecim natomiast posłużono się strategią „kup i trzymaj”. Efektywność każdego eksperymentu mierzona była skumulowanymi stopami zwrotu. Wyniki testów wykazały, że zbudowany przez badaczy model pozwolił na uzyskanie niemal dwukrotnie wyższych stóp zwrotu niż w pozostałych dwóch eksperymentach⁵⁹².

W 2018 roku Kun Chen, Peng Luo, Libo Liu i Weiguo Zhang przeprowadzili badania na 300 spółkach notowanych w indeksie Shanghai-Shenzhen. Zauważyli oni, że w finansowych artykułach prasowych występuje zjawisko, które nazwali „współwspominaniem” (ang. *co-mention*). Chodzi tu o sytuacje, w których autorzy artykułów pisząc o jednej spółce, nawiązują do innej. Autorzy założyli, że między takimi spółkami musi istnieć jakiś specyficzny związek, a im większa częstotliwość występowania tego zjawiska, tym silniejszy ten związek jest. Dla zbadania problemu zastosowali metodę porównawczej analizy sentymentu w celu zidentyfikowania dwóch typów tego zjawiska: „współspominania wspierającego” (ang. *supporting co-mention*) i „współspominania przeciwstawnego” (ang. *opposing co-mentions*). Ich badanie koncentrowało się na sprawdzeniu, czy zjawisko to jest czynnikiem, mogącym mieć wpływ na relację stóp zwrotu między akcjami spółek. Uzyskane wyniki wskazały na pozytywną i statystycznie istotną korelację między współwspominaniem o spółkach, a ich stopami zwrotu. Efekt ten dotyczył obu typów zjawiska. Jednakże w przypadku współspominania przeciwstawnego, współczynnik wpływu na ruchy akcji był niemal dwukrotnie większy niż w przypadku współspominania wspierającego. Zdaniem badaczy taka właściwość

⁵⁹² Zob. Fung, G. P. Ch., Yu, J. X., Lam, W., *Stock prediction: Integrating text mining approach using real-time news*, IEEE International Conference on Computational Intelligence for Financial Engineering, Hong Kong, China 2003.

wyjaśniona może być teorią perspektywy, zgodnie z którą ludzie z reguły przykuwają większą uwagę to złych wiadomości niż dobrych⁵⁹³.

Europa

Belgia

W ramach rozwoju badań, związanych z wykorzystaniem eksploracji tekstu w kontekście rynków finansowych, zauważyć można, że koncepcja ta ewoluuje w dwóch kierunkach. Pierwszym jest stosowanie coraz bardziej zaawansowanych, często autorskich narzędzi, których podstawą są algorytmy uczenia maszynowego lub sztucznej inteligencji. Drugim kierunkiem jest próba odnalezienia w analizowanych tekstach głębszych znaczeń i informacji. Przykładem obu nurtów są badania przeprowadzone na rynku belgijskim, których autorami są Marjan Van de Kauter, Diane Breesch oraz Véronique Hoste. Celem tych badań nie była jednak predykcja cen walorów, a zaproponowanie nowego podejścia do ekstrakcji sentymentu. Badacze ci zauważyli, że większość prowadzonych w tym obszarze analiz koncentruje się na wykrywaniu nastrojów, które określili jako jawne. Tekst, jak sami twierdzą, zawiera również znaczenia ukryte, które również odpowiadają za odbiór publikacji jako pozytywnej lub negatywnej. Aby przetestować zaproponowany przez nich algorytm, skorzystali z pomocy adnotatorów (to jest osób, które manualnie definiowały poszczególne teksty jako pozytywne, negatywne lub neutralne). Dzięki temu stworzony został *złoty standard*, z którym porównywane były wyniki innych podejść, wśród których zastosowano: podejście leksykalne, nadzorowane uczenie maszynowe oraz model autorski. Uzyskane wyniki pozwoliły stwierdzić niemal pełną zgodność rezultatów ostatniego podejścia, z – wcześniej określonym – złotym standardem⁵⁹⁴.

Finlandia

Pytanie o relację, jaka zachodzi między dynamiką walorów a wiadomościami tekstowymi na rynku finlandzkim, zbadali Fabrizio Lillo, Salvatore Micciché, Michele Tumminello, Jyrki Piilo i Rosario N. Mantegna. Problem, który sobie postawili, dotyczył jednak nie tyle cen instrumentów, ile zachowań graczy rynkowych. Zbadali oni decyzje handlowe

⁵⁹³ Zob. Chen, K., Luo, P., Liu, L., Zhang, W., *News, search and stock co-movement: Investigating information diffusion in the financial market*, „Electronic Commerce Research and Applications”, 2018, nr 28, s. 159–171.

⁵⁹⁴ Zob. Van de Kauter, M., Breesch, D., Hoste, V., *Fine-grained analysis of explicit and implicit sentiment in financial news articles*, „Expert Systems with Applications”, 2015, nr 42, s. 4999–5010.

dużej grupy inwestorów, którzy w latach 2003–2008 nabywali lub sprzedawali akcje firmy Nokia. W badaniu uwzględniono dwie grupy czynników: endogeniczne (stopy zwrotu i zmienność cen) oraz egzogeniczne (dzienna liczba artykułów i zawarte w nich nastroje). Przeprowadzone analizy wykazały, że występowanie korelacji między konkretnym czynnikiem a zachowaniami, zależy od tego, do jakiej grupy należy dany inwestor. Podmioty rządowe oraz organizacje trzeciego sektora okazały się mniej wrażliwe zarówno na wiadomości, jak i stopy zwrotu oraz zmienność cen, chociaż wykazywały one tendencję do reagowania na czynniki egzogeniczne. Z kolei gospodarstwa domowe i przedsiębiorstwa okazały się bardziej wrażliwe na oba typy czynników, ale reakcja zdefiniowana została jako silniejsza w przypadku czynników endogenicznych. Analizie poddano również instytucje finansowe oraz podmioty zagraniczne, ale w ich przypadku wyniki nie były jednoznaczne⁵⁹⁵.

Holandia

Interesujące badanie, zwracające uwagę na konsekwencje przedstawiania lub transmitowania przez media negatywnego sentymentu, przeprowadzone zostało w Holandii. Autorami badania są Jan Kleinnijenhuis, Friederike Schultz, Dirk Oegema oraz Wouter van Atteveldt. Badanie było próbą odpowiedzi na pytanie o rolę i znaczenie mediów w okresie kryzysu finansowego 2007. Badacze ci stwierdzili, że treści wiadomości nie tylko były nośnikiem informacji o kryzysie, ale również wpływały na jego pogłębianie, co miało swoje odwzorowanie – między innymi – w zmianach, zachodzących w indeksie AEX⁵⁹⁶. W konkluzji swojego wywodu, na pytanie o to, czy wiadomości finansowe mogą wywołać panikę na rynku, odpowiedzieli twierdząco⁵⁹⁷. Mimo to, analizując spółki notowane w ramach tego samego indeksu, badacze Nadine Strauß, Rens Vliegthart oraz Piet Verhoeven, doszli do zgoła odmiennych wniosków. Podobnie jak w przypadku poprzednich badań, eksploracji podlegały emocje przekazywane przez artykuły prasowe gazet holenderskich, ale próba badawcza składała się z 21 spółek, a okres badawczy obejmował dwanaście lat (2002–2013). Wyniki nie

⁵⁹⁵ Zob. Lillo, F., Micciché, S., Tumminello, M., Piilo, J., Mantegna, R. N., *How news affects the trading behaviour of different categories of investors in a financial market*, „Quantitative Finance”, R. 2015, nr 15/2, s. 213–229.

⁵⁹⁶ *Amsterdam EXchange Market*.

⁵⁹⁷ Zob. Kleinnijenhuis, J., Schultz, F., Oegema, D., van Atteveldt, W., *Financial news and market panics in the age of high-frequency sentiment trading algorithms*, „Journalism”, 2013, nr 14/2, s. 271–291.

potwierdziły jednak wpływu sentymentu na ceny walorów z kolejnego dnia po publikacji ani w przypadku sentymentu pozytywnego, ani negatywnego. Przeprowadzone testy wykazały, że media raczej odzwierciedlają ruchy, mające miejsce na rynku, a nie odwrotnie⁵⁹⁸. Można zauważyć, że wnioski z tego badania są analogiczne do tych wysnutych przez Tanya'a Araújo, Samuela Eleutério oraz Francisco Louçã, którzy analizowali rynek brazylijski.

Norwegia

W 2017 roku Vegard H. Larsen oraz Leif Anders Thorsrud zaprezentowali wyniki badań przeprowadzonych na rynku norweskim. Próba badawcza składała się z 233 spółek notowanych na giełdzie w Oslo w latach 1988–2014. Autorzy doszli do kilku wniosków. Po pierwsze stwierdzili oni, że proste strategie inwestycyjne – oparte na wiadomościach – przynoszą znaczące roczne stopy zwrotu, sięgające nawet 20 procent. Po drugie, w sytuacji, w której na rynek nie napływały żadne wiadomości prasowe (co spowodowane było strajkiem), wyniki spółek, które wykazały wrażliwość na działanie informacji prasowych, odnotowywały spadki. Po trzecie, badacze zwrócili uwagę na brak reakcji rynku w dniu opublikowania informacji oraz wyraźny, istotny statystycznie wynik skumulowanych stóp zwrotu w dniach kolejnych (aż do piętnastego dnia)⁵⁹⁹.

Włochy

W 2022 roku Francesco Colasanto, Luca Grilli, Domenico Santoro i Giovanni Villani przeprowadzili badania na czterech włoskich spółkach: Unicredit, EssilorLuxottica, Intesa SanPaolo, UnipolSai. Do zbudowania wskaźników sentymentu wykorzystano publikacje gazety „Milano Finanza” oraz jeden z popularnych modeli służących do przetwarzania języka naturalnego, BERT⁶⁰⁰. Ich celem było wykorzystanie wyników polaryzacji nastrojów do poprawy prognoz cen akcji. Testom podlegały możliwości predykcyjne w dwóch wariantach. Pierwszy dotyczył reakcji rynku w interwałach godzinowych, drugi natomiast – dziennych. Autorzy wykazali krótkoterminowy wpływ

⁵⁹⁸ Strauß, N., Vliegenthart, R., Verhoeven, P., *Lagging behind? Emotions in newspaper articles and stock market prices in the Netherlands*, „Public Relations Review”, 2016, nr 42, s. 548–555.

⁵⁹⁹ Zob. Larsen, V. H., Thorsrud, L. A., *Asset returns, news topics, and media effects*, „CAMP Working Paper Series”, 2017, nr 5, s. 1–48.

⁶⁰⁰ *Bidirectional Encoder Representations from Transformers*.

wiadomości na rynek akcji, a tym samym stwierdzili, że najlepszym zastosowaniem, zaproponowanej przez nich metodologii, jest predykcja w interwałach godzinowych⁶⁰¹.

Wielka Brytania

W 2021 roku wyniki badań, przeprowadzonych na giełdzie londyńskiej, zaprezentowali Mohammad Alomari, Abdel Razzaq Al Rababa, Ghaith El-Nader, Ahmad Alkhataybeh oraz Mobeen Ur Rehman. Okres badawczy obejmował lata 1998–2017. Celem badania była analiza wpływu sentymentu przekazywanego przez prasę oraz media społecznościowe na wyniki akcji i obligacji, oraz sprawdzenie zachodzącej między nimi korelacji. Autorzy stwierdzili większy wpływ nastrojów, przekazywanych przez prasę na dynamikę rynku akcji, w porównaniu z emocjami przekazywanymi przez media społecznościowe. Rozszerzenie modelu predykcyjnego o nastroje zawarte w wiadomościach pozwoliło również na uzyskanie wyższych wyników w kontekście stóp zwrotu, w porównaniu do modelu rozszerzonego o nastroje przekazywane w mediach społecznościowych. Autorzy doszli do wniosku, że mimo rosnącej roli mediów społecznościowych, informacje przekazywane przez tradycyjne media wciąż stanowią siłę napędową rynku akcji⁶⁰².

W przypadku części z przytoczonych badań, zastosowanie metod eksploracji tekstu pozwoliło na uzyskanie ponadnormatywnych stóp zwrotu lub umożliwiło polepszenie wyników modeli predykcyjnych. Należy jednak zauważyć, że istnieją również takie badania, które wskazują, że tego typu dane nie mogą być definiowane jako zmienne objaśniające dynamikę rynków finansowych. Można więc stwierdzić, że dotychczasowe badania nie pozwalają na uzyskanie w tym kontekście jednoznacznej odpowiedzi. Warto jednak zaznaczyć, że jest to wschodzący obszar finansów i efekty opartych na nim technik wciąż mogą pozostawać niejasne lub nieprecyzyjne⁶⁰³. Pomimo tego, na obecnym etapie badań istnieją przesłanki, by sądzić, że za pomocą analizy tekstu

⁶⁰¹ Zob. Colasanto, F., Grilli, L., Santoro, D., Villani, G., *ALBERTino for stock price prediction: a Gibbs sampling approach*, „Information Sciences”, 2022, nr 597, s. 341–357.

⁶⁰² Zob. Alomari, M., Al Rababa’a, A. R., El-Nader, G., Alkhataybeh, A., Rehman, U. M., *Examining the effects of news and media sentiments on volatility and correlation: Evidence from the UK*, „The Quarterly Review of Economics and Finance”, 2021, nr 82, s. 280–297.

⁶⁰³ Loughran, T., McDonald, B., *Textual Analysis in Accounting...*, op.cit., s. 2.

można uzyskać statystycznie istotne wyniki predykcji oraz klasyfikacji cen instrumentów finansowych, ich wahań lub wolumenów. Stanowi to jeden z argumentów przemawiających za kontynuacją badań w tym obszarze i weryfikacji związanych z nim hipotez na kolejnych rynkach.

Rozdział VI – Sentyment informacji niefinansowych jako zmienna objaśniająca reakcję rynku

Prezentowany w tej części pracy problem badawczy ma charakter interdyscyplinarny i dotyczy zagadnienia roli, jaką informacja odgrywa na rynku finansowym oraz jej potencjalnego wpływu na notowania spółek giełdowych. Odpowiedzi na to pytanie upatrywać można w toczącej się od lat dyskusji między przedstawicielami neoklasycznej teorii finansów a finansami behawioralnymi. Zwolennicy pierwszej szkoły dowodzą trafności hipotezy efektywności rynku w rozumieniu Eugene’a Famy, przyjmując przy tym założenie o racjonalności podmiotów rynkowych⁶⁰⁴. W ramach hipotezy efektywności informacyjnej argumentuje się, że inwestorzy są w stanie przetworzyć pojawiające się informacje natychmiast, co powinno doprowadzić do sytuacji, w której rynek w pełni odzwierciedla dostępną wiedzę⁶⁰⁵. Z kolei przedstawiciele podejścia behawioralnego wskazują na istnienie różnego typu anomalii rynkowych oraz na obciążenie – zdolności percepcyjnych inwestorów – błędami poznawczymi⁶⁰⁶. Jednocześnie od lat 70. obserwować możemy spór o założenie racjonalności graczy rynkowych. Finansiści behawioralni przyjmują stanowisko, że ludzki mózg często przetwarza informacje za pomocą skrótów i filtrów o charakterze emocjonalnym⁶⁰⁷. Można więc przyjąć, że oba nurty wskazują na odmienny sposób dyskontowania informacji przez uczestników rynku kapitałowego. W tym kontekście odpowiedź na pytanie o rolę informacji na rynkach finansowych wydaje się wciąż nierozstrzygnięta, co uzasadnia podejmowanie dalszych analiz w tym obszarze.

Jednocześnie prezentowany problem mieści się w nurcie badawczym określanym mianem *Natural Language Processing* (NLP), przez co rozumieć należy każdą komputerową operację, która dokonywana jest na języku naturalnym⁶⁰⁸. Zagadnieniem wpisany w tę tematykę, jest eksploracja tekstu, to jest zautomatyzowane przetwarzanie

⁶⁰⁴ Stiglitz, J., E., *Informacja i zmiana paradygmatu w ekonomii*, „Gospodarka Narodowa”, 3/2004, s. 89.

⁶⁰⁵ Zob. Fama, E. F., *Efficient Capital Markets: A Review of Theory and Empirical Work*, „The Journal of Finance”, 1970, nr 25/2, s. 383–417.

⁶⁰⁶ Zob. Kahneman, D., Tversky, A., *Prospect Theory: An Analysis of Decision under Risk*, „The Econometric Society”, 1979, nr 47/2, s. 263–292.

⁶⁰⁷ Szyszka, A.: *Behavioral Finance and Capital Markets: How Psychology Influences Investors and Corporations*. Palgrave Macmillan. New York, 2013, s. 82-85.

⁶⁰⁸ Por. Bird. S., Klein. E., Loper. E., *Natural Language Processing with Python*, O’Reilly Media, Sebastopol 2009, s. 9.

języka naturalnego, czego celem jest eksploracja danych, umożliwiająca pozyskanie przez badacza użytecznych informacji o charakterze ilościowym bądź jakościowym⁶⁰⁹. Jedną z gałęzi tego podejścia jest analiza sentymentu, która w literaturze przedmiotu definiowana jest jako „stopień pozytywności lub negatywności tekstu”⁶¹⁰. Jego zastosowanie w kontekście finansów wymaga włączenia do procesu badawczego wiedzy pochodzącej przynajmniej z trzech obszarów: lingwistyki, programowania oraz finansów behawioralnych⁶¹¹. Połączenie wiedzy z zakresu języka i programowania pozwala na automatyczną klasyfikację danych i odkrywanie wzorców, zawartych w dokumentach tekstowych⁶¹². Służy to realizacji nadrzędnego celu, jakim jest wyodrębnienie sentymentu i odpowiedź na pytanie o sposób jego dyskontowania przez inwestorów giełdowych. Z uwagi na dynamiczny rozwój wykorzystywania analizy tekstu w finansach, w najnowszej literaturze anglojęzycznej, ten nurt badawczy, zaczęto określać terminem *Natural Language Based Financial Forecasting* (NLFF)⁶¹³.

Istotną motywacją do zgłębiania omawianej kwestii jest również konstatacja związana z pojawieniem się w ostatnich latach nowego typu danych, określanych mianem Big Data. Cechy charakterystyczne tego rodzaju danych, takie jak wielkość, zmienność, wielowymiarowość czy brak struktury, skłaniają badaczy i analityków do poszukiwania nowych narzędzi, umożliwiających wydobywanie z nich wartości o charakterze informacyjnym⁶¹⁴. Przyczynkiem do podjęcia badań we wskazywanym obszarze jest również fakt dynamicznych zmian w zakresie ilości i rodzaju dostępnych danych. Przeprowadzony przez autorkę przegląd literatury pozwala stwierdzić, że dynamika kreowania danych ma silną tendencję wzrostową⁶¹⁵, a 80% z nich cechuje

⁶⁰⁹ Das. S. R., *Text and Context: Language Analytics in Finance. Foundations and Trends in Finance*, now Publishers Inc, Hanover 2014., s. 6.

⁶¹⁰ Kearney, C.; Liu, S. *Textual sentiment in finance: A survey of methods and models*. International Review of Financial Analysis. R.2014. nr 33. s. 172.

⁶¹¹ Nassirtoussi. A. K.; Aghabozorgi, S.; Wah, T.Y.; Ngo, D. C. L. *Text mining for market prediction: A systematic review*. Expert Systems with Applications. R.2014. nr 41, s. 7654.

⁶¹² Khan, A.; Baharudin, B.; Lee, L. H.; Khan, K. A., *Review of Machine Learning Algorithms for Text-Documents Classification*. Journal of Advances in Information Technology R.2010. nr 1/1. s. 4.

⁶¹³ Zob. Xing, F. Z.; Cambria, E.; Welsch, R. E. *Natural language based financial forecasting: a survey*. Artificial Intelligence Review R.2018. nr 50. s. 49–73.

⁶¹⁴ Zikopoulos P.C.; Eaton, Ch.; deRoos, D.; Deutsch, T.; Lapis, G.: *Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill. New York, 2011, s. 3

⁶¹⁵ Zob. PwC: *Where have you been all my life? How the financial services industry can unlock the value in Big Data*, 2013.

brak struktury lub jedynie fragmentaryczne ustrukturyzowanie⁶¹⁶. Ponadto dominująca część z nich ma charakter tekstowy⁶¹⁷. Równocześnie obserwujemy znaczący wzrost danych związanych z biznesem⁶¹⁸. Można pokusić się o wskazanie dwóch przyczyn takiego stanu rzeczy. Po pierwsze, upatrywać można go w obecnym trendzie kładzenia coraz większego nacisku na sprawozdawczość niefinansową (np. raporty CSR, czy raportowanie realizacji najlepszych praktyk dla spółek giełdowych). Po drugie, z uwagi na rewolucję internetową, nastąpił gwałtowny wzrost liczby źródeł informacji na temat rynków finansowych (strony internetowe, czasopisma wydawane w formie elektronicznej, blogi). Z uwagi na powyższe, istotna wydaje się weryfikacja metod i narzędzi umożliwiających analizę tego typu danych.

Przeprowadzony przez autorkę, przegląd literatury uprawnia również do konkluzji, że rozwój badań z zakresu analizy sentymentu na rynkach wschodzących – w porównaniu do rynków rozwiniętych – określić można jako relatywnie słaby. W efekcie brak jest punktów odniesienia w postaci wyników badań, które opisywałyby sposób dyskontowania informacji na rynkach Europy Środkowo-Wschodniej z perspektywy ich nacechowania emocjonalnego. Powyższe argumenty stanowiły przesłankę do stwierdzenia, że w podejmowanej tematyce istnieje luka badawcza.

6.1. Metodyka i procedura badawcza

Pierwsza część prezentowanego rozdziału jest szczegółowym omówieniem metodyki i procedury badawczej. Zostały w niej przedstawione hipotezy badawcze oraz narzędzia zastosowane w celu ich weryfikacji. Objąsnione zostały też kryteria doboru źródeł danych oraz sposób ich gromadzenia, co w dalszej kolejności umożliwiło scharakteryzowanie próby badawczej. Szczególna uwaga poświęcona została również etapom przeprowadzonych badań. Ponadto przedstawiono ograniczenia badawcze, które mogły wpłynąć na interpretację wyników i ogólną wiarygodność uzyskanych rezultatów. Autorka wyraża nadzieję, że będzie to stanowiło fundament dla dyskusji nad

⁶¹⁶ Zikopoulos P.C.; Eaton, Ch.; deRoos, D.; Deutsch, T.; Lapis, G.: *Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill. New York, 2011, s. 8.

⁶¹⁷ Das. S.R. *Text and Context: Language Analytics in Finance. Foundations and Trends in Finance*. now Publishers Inc. Hanover, 2014, s. 3.

⁶¹⁸ Zob. McKinsey Global Institute: *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, 2011, s. 156.

kierunkami przyszłych badań, związanych z zastosowaniem eksploracji sentymentu w obszarze finansów.

Celem badania jest zbadanie i ocena sposobu dyskontowania informacji przez inwestorów giełdowych, z punktu widzenia dodatkowej perspektywy, jaką jest sentyment artykułów prasowych, na temat spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie.

Na podstawie treści zaprezentowanych w pierwszej części pracy oraz przeglądu badań, przeprowadzonych dotychczas we wskazywanym obszarze, sformułowana została następująca hipoteza główna:

H0: Rodzaj sentymentu, zawarty w danych tekstowych, jest odzwierciedlony w reakcji rynku.

Hipoteza główna wsparta została poniższymi hipotezami pomocniczymi:

H1: Anglojęzyczne artykuły prasowe, dotyczące spółek notowanych w indeksie WIG20 w okresie od 01.01.2013 do 31.12.2022, charakteryzują się przewagą sentymentu pozytywnego względem sentymentu negatywnego.

H2: Występuje asymetria w sposobie reakcji rynku na pozytywny i negatywny sentyment, zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20.

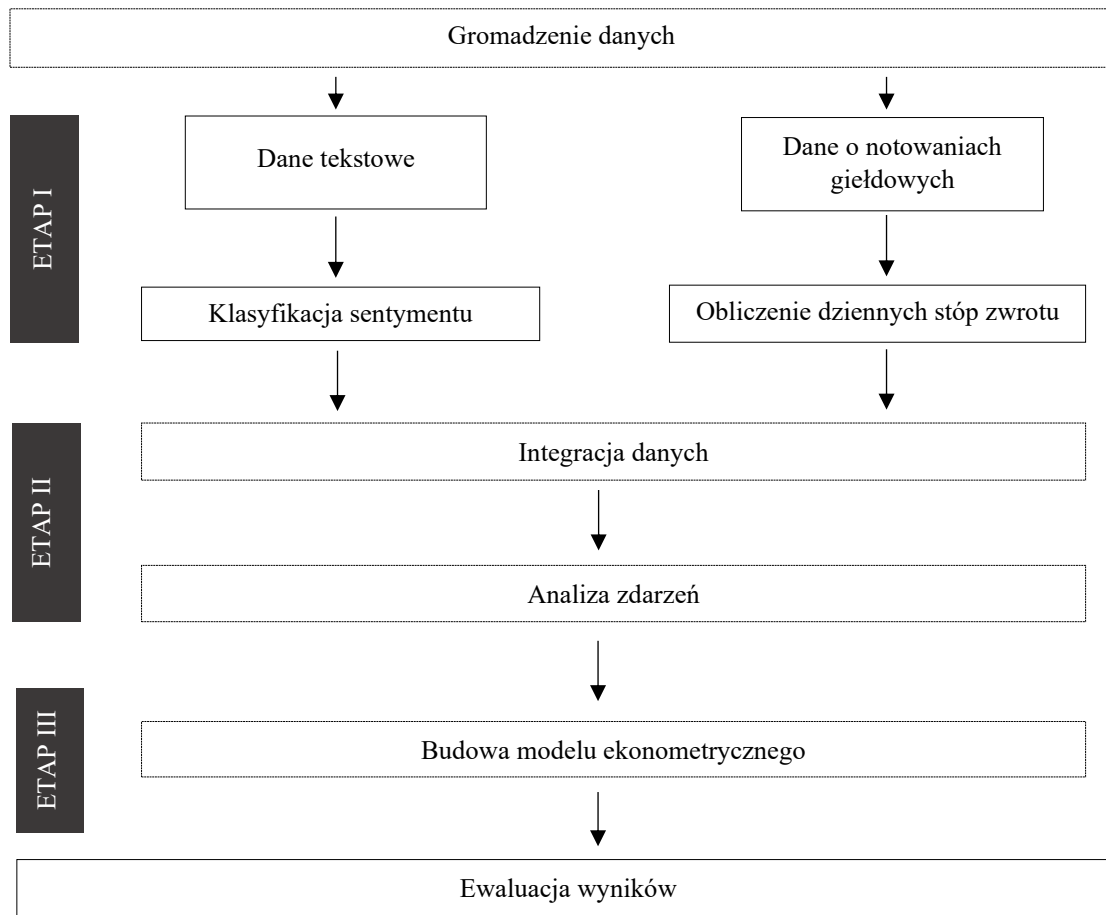
H3: Sentyment zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20, stanowi determinantę inwestorów giełdowych.

6.1.2. Opis procedury badawczej

Procedura badawcza rozpoczęła się od zgromadzenia niezbędnych danych tekstowych – w postaci artykułów prasowych – oraz finansowych, – w postaci dziennych notowań cen akcji spółek, wchodzących w skład próby badawczej. Samo badanie składało się z trzech etapów. Pierwszy związany był z ekstrakcją oraz analizą sentymentu artykułów prasowych, co służyło weryfikacji pierwszej hipotezy pomocniczej. W etapie drugim przeprowadzona została analiza zdarzeń, pozwalająca oszacować ponadnormatywne stopy zwrotu, związane z pojawianiem się artykułów prasowych o określonym sentymencie. Zadaniem tego etapu była weryfikacja drugiej hipotezy pomocniczej.

Ostatnia część badania polegała na estymacji modelu klasyfikacyjnego w postaci regresji logitowej, co umożliwić miało weryfikację trzeciej hipotezy pomocniczej. Ogólny schemat całej procedury zaprezentowany jest na rysunku nr 7.

Rysunek 7. Procedura badawcza



Źródło: opracowanie własne.

Gromadzenie danych

Przed rozpoczęciem właściwego badania zgromadzone zostały zarówno dane tekstowe, jak i finansowe, związane ze spółkami wchodzącymi w skład indeksu WIG20 w okresie od stycznia 2013 roku do końca grudnia 2022 roku. Dane tekstowe, w postaci artykułów prasowych w języku angielskim, pobrane zostały z serwisu *Emerging Markets Information Service (EMIS)*. Dobór tego źródła danych podyktowany był kilkoma przesłankami. Po pierwsze, aktywność twórców bazy *EMIS* skoncentrowana jest na obszarach rynków wschodzących. Spółka gromadzi informacje o ponad dziesięciu milionach przedsiębiorstw, funkcjonujących w 197 krajach Azji, Ameryki Łacińskiej,

Afryki, Bliskiego Wschodu oraz Europy Wschodniej. Stanowiło to przesłankę do stwierdzenia, że znajdują się tam dane na temat polskich podmiotów gospodarczych. W toku dalszej weryfikacji tego źródła danych okazało się, że pozwala ono na dostęp do informacji na temat ponad czterystu tysięcy przedsiębiorstw, funkcjonujących na terenie Polski⁶¹⁹. Wybór tej bazy danych gwarantował więc szeroki dostęp do informacji, podawanych w środowisku międzynarodowym, na temat spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie.

Po drugie, w bazie agregowane są nie tylko dane finansowe (takie jak: sprawozdania finansowe, oceny kredytowe, czy też analizy sektorowe), ale również tekstowe – w tym artykuły prasowe. Ponadto spółka *EMIS*, gromadzi swoje archiwum w ramach współpracy zarówno z globalnymi, jak i lokalnymi dostawcami danych. W kontekście prowadzonego badania niewątpliwym atutem bazy było więc przechowywanie w niej informacji bezpośrednio ze źródeł pierwotnych. W przypadku polskich przedsiębiorstw *EMIS* umożliwiało pobranie artykułów publikowanych zarówno przez instytucje publiczne i agencje informacyjne (np. *Polska Agencja Prasowa*), jak i magazyny komercyjne o tematyce biznesowej (np. *Parkiet*⁶²⁰, *Bankier*⁶²¹, *The Polish News Bulletin*⁶²², *Poland A.M. Daily*⁶²³) oraz ogólnospołecznej (np. *Rzeczpospolita*, *Wprost*). Dzięki skorzystaniu z tej bazy uzyskano artykuły prasowe w możliwie nieprzetworzonej formie. Pozwoliło to przyjąć założenie, że będą one oddawały, zawarty w nich sentyment, zgodnie z intencjami autorów publikacji.

Po trzecie, *EMIS* cieszy się popularnością wśród szerokiego grona odbiorców. Użytkownikami bazy są podmioty rządowe (np. *Polska Agencja Rozwoju Przedsiębiorczości*⁶²⁴), bankowe, firmy konsultingowe, jak również uczelnie wyższe (np. Uniwersytet Harvarda⁶²⁵, Uniwersytet Columbia⁶²⁶, Uniwersytet Warszawski⁶²⁷).

⁶¹⁹ Zob. Emerging Markets Information Service, <https://www.emis.com/pl>, Dostęp: 10 sty. 2024.

⁶²⁰ Patrz: *Parkiet*, <https://www.parkiet.com/>, Dostęp: 10 sty. 2024.

⁶²¹ Patrz: *Bankier*, <https://www.bankier.pl/>, Dostęp: 10 sty. 2024.

⁶²² Patrz: *Polish News Bulletin*, <https://pnb.pl/>, Dostęp: 10 sty. 2024.

⁶²³ Patrz: *Poland A.M. Daily*, <https://polandam.pl/>, Dostęp: 10 sty. 2024.

⁶²⁴ Patrz: Polska Agencja Rozwoju Przedsiębiorczości, *Public Intelligence. The Use of Administrative Data for Monitoring and Evaluation of Public Policies*, 2022, Warszawa, s. 73.

⁶²⁵ Patrz: Harvard Library, *A guide to library resources in support of research in urban planning and design*, <https://guides.library.harvard.edu/gsd/urban>, Dostęp: 10 sty. 2024.

Powszechność wykorzystywania bazy wśród renomowanych jednostek, zainteresowanych tematyką biznesową, świadczy pozytywnie o jej rzetelności i wiarygodności. Powyższe argumenty pozwoliły uznać bazę EMIS za adekwatne źródło danych, odpowiadające potrzebom prowadzonego badania.

W badaniu wykorzystano 56 953 anglojęzycznych artykułów prasowych. Wyjaśnienia wymaga jednak powód wykorzystania artykułów napisanych w języku angielskim. Po pierwsze, zauważyć można, że w – wyniku procesów globalizacyjnych – język ten zyskał status języka międzynarodowego. Dorothea C. Lazaro i Erlinda M. Medalla w swojej publikacji „English as the Language of Trade, Finance and Technology in APEC: An East Asia Perspective” zwróciły uwagę na wzrost jego znaczenia w latach 90. XX wieku, kiedy to powiązany został z handlem, finansami oraz technologią⁶²⁸. W tych kontekstach w XXI stał się wręcz kluczowy. Nie chodzi tu jednak wyłącznie o znajomość języka, ale o wykorzystanie go do skutecznej komunikacji. Obecnie jest wręcz niezbędny dla zacieśniania relacji w handlu międzynarodowym, rozwoju gospodarczym, czy też finansowym⁶²⁹. Jak twierdzi Belinda Crawford, język angielski stał się dla inwestorów swoistym *lingua franca*⁶³⁰.

Podobny proces – w ostatnich 20 latach – zaobserwować można na lokalnym rynku giełdowym. O ile w roku 2002 udział inwestorów krajowych w obrocie akcjami spółek notowanych na Rynku Głównym Giełdy Papierów Wartościowych w Warszawie był dominujący (stanowiąc 65%), to dziesięć lat później odsetek obrotów generowanych przez inwestorów krajowych i zagranicznych zaczął osiągać zbliżone wartości (odpowiednio 52% oraz 48%). Obecnie struktura ta uległa odwróceniu, wobec wyników odnotowywanych na początku wieku. W roku 2022 udział inwestorów zagranicznych

⁶²⁶ Patrz: Columbia University Libraries, Research Guides. A-Z Databases, <https://guides.library.columbia.edu/az.php?a=all>, Dostęp: 10 sty. 2024.

⁶²⁷ Patrz: Biblioteka Uniwersytetu Warszawskiego, Bazy online, <https://www.buw.uw.edu.pl/zasoby-online/bazy-online/>, Dostęp: 10 sty. 2024.

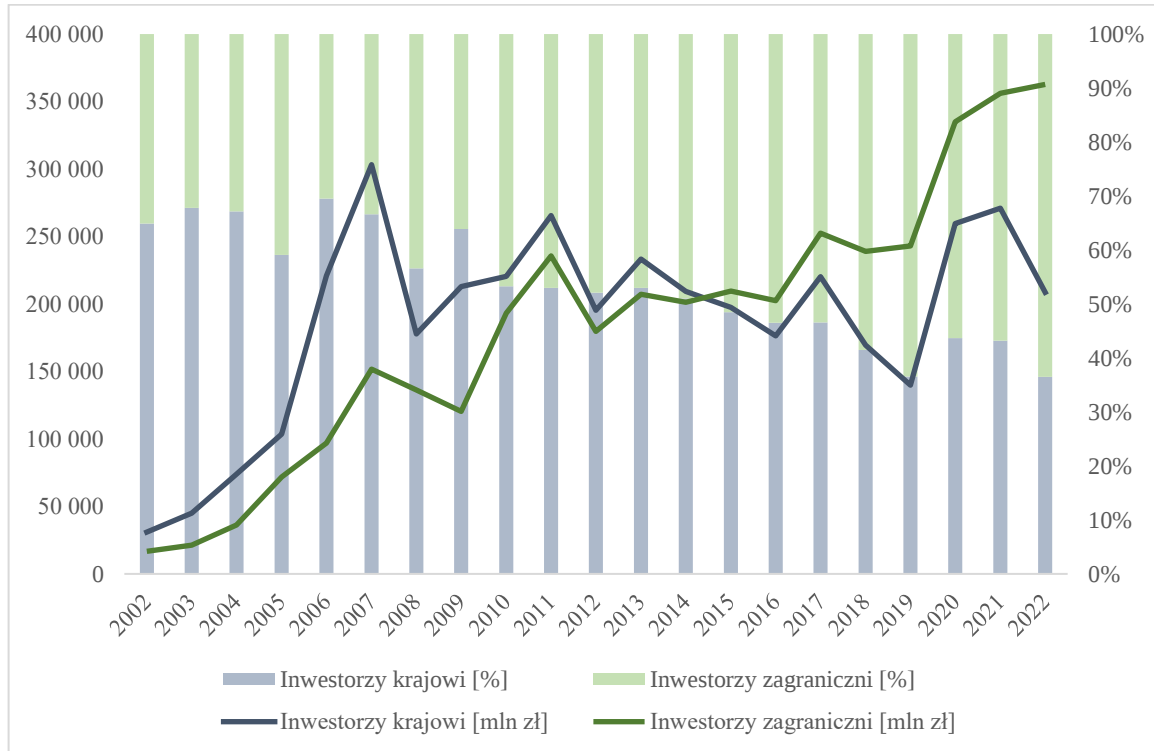
⁶²⁸ Zob. Dorothea C. Lazaro & Erlinda M. Medalla. (2004). English as the Language of Trade, Finance and Technology in APEC: An East Asia Perspective. Philippine Journal of Development, Number 58, Volume XXXI, No. 2.

⁶²⁹ Zob. Shaharuddin, N., Abdul Wahab, M. H., Safian, S. S., & Osman, A. A. (2022). The Role of English Proficiency and Financial Development in Stimulating Economic Growth. Journal of Entrepreneurship, Business and Economics, 10(2S2), 18–37.

⁶³⁰ Zob. Using English as a lingua franca to engage with investors: An analysis of Italian and Japanese companies' investor relations communication policies.

w obrotach akcjami wynosił już bowiem 63%, natomiast w przypadku inwestorów krajowych było to jedynie 37%⁶³¹. Tendencja ta zobrazowana została na rysunku nr 8.

Rysunek 8. Udział inwestorów krajowych i zagranicznych w obrocie Rynku Głównego GPW w Warszawie



Źródło: opracowanie własne na podstawie www.gpw.pl/analizy (dostęp: 5 sty. 2024)

Jednocześnie, z uwagi na lukę badawczą, występującą w analizowanym obszarze, zasadne wydawało się wykorzystanie narzędzi, które podlegały wcześniejszej weryfikacji. Jednym z takich narzędzi jest anglojęzyczny słownik sentymentu *Loughran-McDonald Master Dictionary*, który skonstruowany został z myślą o zastosowaniu w obszarach, związanych z tematyką finansową i ekonomiczną⁶³². Przykłady badań, w których skorzystano z owego słownika, zaprezentowane zostały w tabeli numer 7. Wziąwszy pod uwagę przytoczone argumenty, autorka pracy podjęła decyzję o tym, że do badania zaaplikowane zostaną artykuły anglojęzyczne.

⁶³¹ Zob. https://www.gpw.pl/pub/GPW/analizy/2015_P2_Udzialy_20160222.pdf oraz https://www.gpw.pl/pub/GPW/analizy/Udzial_inwestorow_1H2023.pdf

⁶³² Zob. Loughran, T.; McDonald, B.: *Loughran-McDonald Master Dictionary*. <https://sraf.nd.edu/loughranmcdonald-master-dictionary/>, Dostęp: 1 kw. 2023.

Informacje o notowaniach rynkowych z kolei pobierane były z serwisów *stooq.pl* oraz *Thomson Reuters*, które udostępniają dzienne ceny zamknięcia poszczególnych walorów, notowanych na Giełdzie Papierów Wartościowych w Warszawie. Umożliwiło to obliczenie ponad 50 tysięcy dziennych stóp zwrotu dla trzydziestu trzech badanych spółek.

Weryfikacja pierwszej hipotezy pomocniczej

Pierwszy etap badania związany był przede wszystkim z pracą z danymi tekstowymi. W poniższej części rozdziału przedstawiono, zastosowaną w badaniu, technikę oczyszczania danych oraz metodę ekstrakcji sentymentu. Objąsnione zostało również to, w jaki sposób szacowany był poziom sentymentu, przypisywany do każdego artykułu prasowego. Na koniec opisane zostały warunki, umożliwiające pozytywną lub negatywną weryfikację pierwszej hipotezy pomocniczej.

Oczyszczanie danych

Praca z danymi tekstowymi rozpoczęła się od przeprowadzenia procesu ich oczyszczania. Działania te powodowane były koniecznością redukcji wielowymiarowości danych nienumerycznych (ang. *dimensionality reduction*). Właściwością danych tekstowych jest bowiem to, że wyposażone są w dużą liczbę, niekontrolowanych przez badacza, parametrów. Ich redukcja służy uniknięciu zjawiska określanego mianem przeuczenia modelu (ang. *overfitting*), to jest sytuacji, w której na skutek trenowania modelu na zbiorze wyposażonym w zbyt dużą liczbę cech, zmniejsza się jego jakość i efektywność⁶³³.

W tym celu na wszystkich – zaklasyfikowanych do badania – artykułach przeprowadzona została *tokenizacja* (ang. *tokenization*). Token zdefiniować można jako minimalną jednostkę, która może być “zrozumiana” i przetworzona przez komputer. Proces ten pozwala na podzielenie ciągu słów języka naturalnego, które dla modeli i algorytmów nie niosą ze sobą żadnego znaczenia, na – posiadające znaczenie i możliwe do przetworzenia – tokeny⁶³⁴. Umożliwiło to również wyeliminowanie ze wszystkich

⁶³³ Padhi, D. K., Padhy, N., Bhoi, A. K., Shafi, J., Ijaz, M. F., *A fusion framework for forecasting financial market direction using enhanced ensemble models and technical indicators*, “Mathematics”, 2021, nr 9 s. 12.

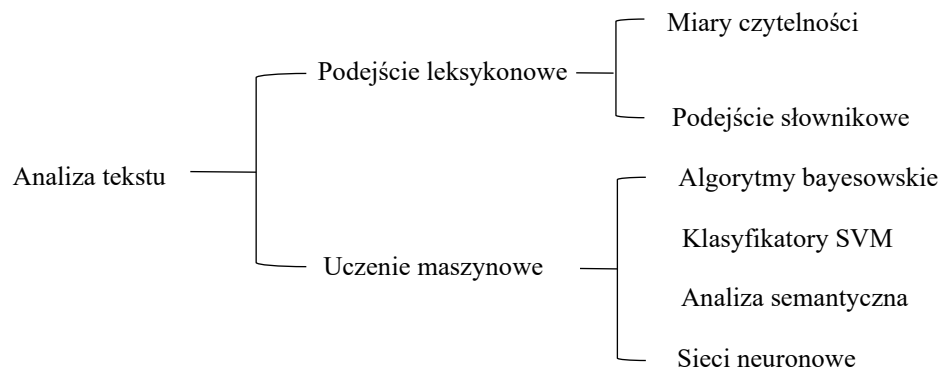
⁶³⁴ Zob. Hardeniya, N., *NLTK Essentials*, Birmingham, 2015, s. 23 – 24.

artykułów cyfr oraz tak zwanych *stop words*, to jest słów lub liter, które nie posiadają warstwy merytorycznej oraz emocjonalnej, ale niezbędne są do porozumiewania się językiem naturalnym. W języku angielskim za tego typu słowa uznaje się na przykład “and”, “him” lub “such”.

Ekstrakcja sentymentu

Oczyszczenie danych umożliwiło rozpoczęcie procedury ekstrakcji sentymentu. Dokonany przez autorkę przegląd badań wykazał, że – w analizowanym kontekście – nie można mówić o jednym modelu, dającym pogląd na cały proces, wykorzystania eksploracji tekstu do celów analizy finansowej. Dzieje się tak, chociażby z tego powodu, że sam pomysł stosowania tego typu danych w ramach finansów jest stosunkowo nowy, zwłaszcza jeśli uwzględnimy włączenie do niego algorytmów uczenia maszynowego. W związku z tym – na obecnym etapie rozwoju wiedzy – różne propozycje i podejścia badaczy wciąż się ze sobą ścierają, a nawet wzajemnie wykluczają⁶³⁵.

Rysunek 9. *Ogólny podział metod analizy tekstu*



Źródło: Guo, Li., Shi, F., Tu, J. (2016). *Textual analysis and machine learning: Crack unstructured data in finance and accounting*, „The Journal of Finance and Data Science”, 2, 153-170, s.155.

W tym nurcie istnieją dwa podejścia, rozwiązujące problem ekstrakcji sentymentu z tekstów, napisanych w języku naturalnym: (1) bazujące na leksykonach, (2) bazujące na algorytmach maszyn uczących się. W literaturze przedmiotu spotkać można również trzeci sposób, jakim jest analiza lingwistyczna. Jest to jednak podejście bardzo trudne do zaaplikowania w ramach technik komputerowych, dlatego w praktyce w zasadzie się go

⁶³⁵ Zob. Nassirtoussi. A. K., Aghabozorgi, S., Wah, T. Y., Ngo, D. C. L., *Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment*, „Expert Systems with Applications”, 2015, nr 42.

nie stosuje⁶³⁶. Zaletą podejścia, wykorzystującego maszyny uczące się, jest między innymi to, że mogą uczyć się z przedstawionych im danych oraz dobrać atrybuty – służące na przykład klasyfikacji lub predykcji – w trakcie tego procesu. Może to skutkować lepszym dopasowaniem zmiennych uczących do analizowanego korpusu tekstów. Wadą, wykorzystania maszyn uczących się, jest natomiast to, że dobór owych atrybutów często pozostaje poza kontrolą badacza (tak dzieje się na przykład w przypadku sieci neuronowych). Ponadto ich stosowanie wymaga wcześniejszego oznaczenia danych pod względem docelowych klas, co określić należy jako działanie zarówno czasochłonne, jak i kosztowne. Modele leksykalne można z kolei uznać za prostsze do zastosowania. Wymagają jednak posiadania gotowego słownika. Ponieważ konstruowanie słowników odbywa się w oparciu o konkretne korpusy (tj. zbiory tekstów z danej dziedziny tematycznej), to istnieje ryzyko, że gotowe słowniki mogą okazać się niedostosowane do – analizowanych przez badacza – dokumentów⁶³⁷.

W odpowiedzi na powyższe cechy obu koncepcji, w badaniu zastosowane zostało podejście leksykalne z wykorzystaniem słownika *Loughran-McDonald Master Dictionary*. Słownik ten zaproponowany został przez profesorów w dziedzinie finansów Tima Loughrana oraz Billa McDonald. Stanowi rozszerzenie wcześniejszego słownika *2of12inf*, poprzez uwzględnienie tokenów występujących w raportach 10-K oraz *earning calls*. Dzięki temu słownik ten dostosowany został do specyfiki domeny finansów, przez co stał się jednym z częściej stosowanych w tym obszarze tematycznym. Wybrane przykłady badań, w ramach których zastosowano wspomniany słownik, zaprezentowane zostały w tabeli nr 7.

Szacowanie poziomu sentymentu

Zaaplikowanie słownika pozwoliło na wskazanie – pojawiających się w analizowanych tekstach – słów, które nie posiadały żadnego nacechowania emocjonalnego (neutralne) oraz tych, które takim nacechowaniem się charakteryzowały. Równocześnie ta druga grupa podzielona została na słowa pozytywne oraz negatywne.

⁶³⁶ Smailovic, J., Grcar, M., Lavrac, N., Znidarsic, M., *Stream-based active learning for sentiment analysis in the financial domain*, „Information Sciences”, R. 2014, nr 285, s.185.

⁶³⁷ Fernandez, R., Guizar, B. P., Rho, C. *A sentiment-based risk indicator for the Mexican financial sector*, Latin American Journal of Central Banking, 2021, nr 2, s. 6.

Tabela 7. Wybrane przykłady badań, wykorzystujących słownik Loughran-McDonald Master Dictionary

rok	autorzy	tytuł	czasopismo
2022	Elena Fedorova, Pavel Drogovoz, Alexandr Nevredinov, Polina Kazinina, Cai Qitan	Impact of MD&A sentiment on corporate investment in developing economies: Chinese evidence	Asian Review of Accounting
2021	Samuel P. Fraiberger, Do Lee, Damien Puy, Romain Ranciere	Media sentiment and international asset prices	Journal of International Economics
2020	Xiaodong Li, Pangjing Wu, Wenpeng Wang	Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong	Information Processing and Management
2019	Ritu Yadava, A. Vinay Kumarb, Ashwani Kumar	News-based supervised sentiment analysis for prediction of futures buying behaviour	IIMB Management Review
2016	Sunita Goel, Ozlem Uzuner	Do sentiments matter in fraud detection? Estimating semantic orientation of annual reports	Intelligent Systems in Accounting, Finance and Management
2015	Kenneth R. Ahern, Denis Sosyura	Rumor Has It: Sensationalism in Financial Media	The Review of Financial Studies
2014	Xiaodong Li, Haoran Xie, Li Chen, Jianping Wang, Xiaotie Deng	News impact on stock price return via sentiment analysis	Knowledge-Based Systems
2014	Allen H. Huang, Amy Y. Zang, Rong Zheng	Evidence on the Information Content of Text in Analyst Reports	The Accounting Review
2013	Diego Garcia	Sentiment during recessions	The Journal of Finance

Źródło: opracowanie własne.

Po przeprowadzeniu klasyfikacji słów każdy z tekstów zdefiniowany został jako negatywny, pozytywny lub neutralny, na podstawie poniższego wzoru⁶³⁸.

$$sentyment = \frac{N_{pos} - N_{neg}}{N_{pos} + N_{neg}}$$

N_{pos} - liczba słów o nacechowaniu pozytywnym

N_{neg} - liczba słów o nacechowaniu negatywnym

Uzyskane wyniki siły sentymentu mieściły się w przedziałach od -1 do +1, odpowiednio dla tekstów negatywnych i pozytywnych. Jako neutralny oznaczany był tekst, który osiągnął wartość sentymentu równą 0.

Weryfikacja hipotezy badawczej

Weryfikacja hipotezy polegała na przeanalizowaniu każdej ze spółek z próby badawczej, pod względem udziału artykułów o pozytywnym i negatywnym sentymencie. Za brak podstaw do odrzucenia hipotezy uznana została sytuacja, w której dla więcej niż połowy podmiotów, artykuły prasowe o sentymencie pozytywnym będą stanowiły ponad 50% ogólnej liczby publikacji na temat każdego z nich. W odwrotnym przypadku hipoteza zostanie zweryfikowana negatywnie.

Weryfikacja drugiej hipotezy pomocniczej

Zintegrowanie przetworzonych danych tekstowych oraz danych o notowaniach rynkowych, umożliwiło przeprowadzenie drugiego etapu badań, podczas którego zastosowana została metoda analizy zdarzeń. Metoda jest powszechnie wykorzystywana w badaniach, związanych z rynkiem finansowym. Dotąd stosowana była zarówno w badaniach nad Giełdą Papierów Wartościowych w Warszawie⁶³⁹, jak i badaniach

⁶³⁸ Li, X., Xie, H., Chen, L., Wang, J., Deng, X., News impact on stock price return via sentiment analysis, „Knowledge-Based Systems”, R. 2014, nr 69, s. 19.

⁶³⁹ Zob. Czekaj, J., Woś, M., Żarnowski, J., *Efektywność giełdowego rynku akcji w Polsce. Z perspektywy dziesięciolecia*, Wydawnictwo Naukowe PWN, Warszawa 2001.; Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna a Wolters Kluwer business, Warszawa 2012.

Tabela 8. Zastosowanie analizy zdarzeń w ramach badań związanych z analizą tekstu w kontekście rynków finansowych

rok	autorzy	tytuł	czasopismo
2023	Svetlana Borovkova, Alan Lammiman	The impact of news sentiment on energy futures returns	Working Paper
2022	Muhammad Umar i inni	Crypto swings and the performance of carbon-intensive equity funds in China	Resources Policy
2022	Vivien E. Jancenelle, Dominic Buccieri	Signaling downsizing intentions after a major crisis: does managerial authenticity matter?	International Journal of Organizational Analysis
2021	Yi Xuan Lim, Consilz Tan	Do negative events really have deteriorating effects on stock performance? A comparative study on Tesla (US) and Nio (China)	Journal of Asian Business and Economic Studies
2021	Maciej Wujec	Analysis of the Financial Information Contained in the Texts of Current Reports: A Deep Learning Approach	Journal of Risk and Financial Management
2021	Marco Caiffa i inni	CEO Duality: Newspapers and Stock Market Reactions	Journal of Risk and Financial Management
2020	Jesus Carmona i inni	Social media sentiment, tariffs, and international equity pricing	International Journal of Electronic Finance
2020	Vivien E. Jancenelle	Moral foundations and financial performance: the effect of moral cues during times of earnings uncertainty	International Journal of Organizational Analysis
2019	Ahmed Bouteska	The effect of investor sentiment on market reactions to financial earnings restatements: Lessons from the United States	Journal of Behavioral and Experimental Finance
2019	Vivien E. Jancenelle i inni	Making investors feel good during earnings conference calls: The effect of warm-glow rhetoric	Journal of General Management
2019	Nadine Strauss, Christopher Holmes Smith	Buying on rumors: how financial news flows affect the share price of Tesla	Corporate Communications: An International Journal
2018	Tahir M. Nisar, Man Yeung Southampton	Twitter as a tool for forecasting stock market movements: A short-window event study	The Journal of Finance and Data Science
2018	Vivien E. Jancenelle	Organizational Psychological Capital During Earnings Conference Calls: Mitigating Shareholders' Sell-Off in the Face of Earnings Surprises?	Journal of Leadership & Organizational Studies
2010	Michael D. Pfarrer i inni	The effects of firm reputation and celebrity on earnings surprises and investors' reactions	Academy of Management Journal
2008	Elaine Henry	Are investors influenced by how earnings press releases are written?	Journal of Business Communication

Źródło: opracowanie własne.

związanych z eksploracją tekstu, prowadzonych w obszarze tematyki finansowej⁶⁴⁰. Wybrane przykłady tego typu publikacji zaprezentowane zostały w tabeli numer 8.

Zdarzenie zdefiniowane zostało jako dzień publikacji artykułu, cechującego się skrajnym sentymentem pozytywnym lub negatywnym. Do oszacowania wartości normatywnych stóp zwrotu zaaplikowany został model rynkowy przy oknie estymacyjnym, wynoszącym 30 dni. Reakcja rynku badana była natomiast poprzez oszacowanie skumulowanych średnich ponadnormatywnych stóp zwrotu (CAAR), w oknie zdarzenia od dnia poprzedzającego zdarzenie do pięciu dni po zdarzeniu. Celem tej części analizy była weryfikacja drugiej hipotezy pomocniczej. Poniżej zamieszczony został szczegółowy opis każdego z podjętych kroków.

Definicja zdarzenia

Pierwszym krokiem w tym etapie badań było zdefiniowanie zdarzenia. Z uwagi na to, że powinno być możliwie jednoznaczne, w analizie wykorzystano artykuły prasowe, do których przypisane zostały jedynie skrajne wyniki sentymentu. Ponieważ miara sentymentu, zastosowana w pierwszym etapie badania, pozwoliła na przypisanie do każdego artykułu wyników mieszczących się w przedziale od -1 do 1, to jako skrajne zdefiniowane zostały wyniki większe niż 0,9 oraz mniejsze niż -0,9, odpowiednio dla sentymentu pozytywnego i negatywnego. Za dzień zdarzenia przyjęto dzień publikacji artykułu prasowego. W ten sposób do badania zakwalifikowanych zostało łącznie 5 861 artykułów, przy czym 3 185 z nich stanowiły artykuły o sentymencie pozytywnym, a pozostałe 2 676 stanowiły artykuły o sentymencie negatywnym. Szczegółowe dane, dotyczące tekstów, przypisanych każdej ze spółek, zaprezentowane są na wykresie nr 5.

Definicja okna zdarzenia

Można stwierdzić, że w przypadku badań opierających się na analizie tekstu wybór długości okna zdarzenia – przynajmniej częściowo – jest wynikiem wcześniejszego doboru źródeł danych tekstowych. Jak zauważa Ahmed Bouteska istotą problemu jest wybór okna zdarzenia, które byłoby wystarczająco długie, aby uniknąć pominięcia istotnych reakcji rynkowych, a równocześnie na tyle krótkie, aby uniknąć wpływu

⁶⁴⁰ Zob. Iatridis, G. E., *Financial reporting language in financial statements: Does pessimism restrict the potential for managerial opportunism?*, „International Review of Financial Analysis”, 2016, nr 45.

niepowiązanych zdarzeń zakłócających ceny akcji⁶⁴¹. Badacze koncentrujący się na analizowaniu relacji między stopami zwrotu a artykułami prasowymi lub danymi, pochodzącymi z mediów społecznościowych, stają bowiem przed innymi problemami niż badacze, zajmujący się oficjalnymi raportami i sprawozdaniem spółek giełdowych. W tym drugim przypadku mamy do czynienia z usankcjonowaną prawnie regularnością publikacyjną. Jasny jest więc moment upubliczniania raportów oraz częstotliwość, z jaką będą pojawiały się kolejne. W przypadku wyboru źródeł danych w postaci artykułów prasowych lub mediów społecznościowych, częstotliwość pojawiania się treści jest zależna od woli dziennikarza, czy też użytkownika owych mediów. Ponadto pobrane dane mogą znacząco różnić się pod względem liczebności. Przykładowo Tahir M. Nisar i Man Yeung, analizując możliwość wykorzystania treści publikowanych na platformie *Twitter* do prognozowania stóp zwrotu z rynku akcji, zagregowali niemal 61 tysięcy wpisów, mimo że okres badawczy wynosił jedynie pięć dni, a wyszukiwanie wpisów opierało się jedynie o dziewięć *hashtagów*⁶⁴².

Henryk Gurgul zauważa, że w początkowych pracach, wykorzystujących metodę analizy zdarzeń, badacze skłaniali się raczej ku konstruowaniu relatywnie szerokich okien zdarzeń, często przekraczających 20 dni. Obecny trend wskazuje jednak na skłonność do zawężania okien, co wynika z kilku przesłanek. Po pierwsze, zastosowanie szerokiego okna zdarzenia zwiększa ryzyko, że wystąpią w nim zdarzenia zakłócające, które mogą zniekształcić wynik analizy. Po drugie, ustanowienie szerokiego okna zdarzenia, może doprowadzić do tego, że mimo wystąpienia znaczącej reakcji rynku na badane zdarzenie (np. w dniu zdarzenia), skumulowane nadwyżkowe stopy zwrotu oszacowane zostaną jako statystycznie nieistotne⁶⁴³. Z uwagi na powyższe przesłanki oraz na sposób definiowania zdarzenia – w przedstawianym badaniu – dla wszystkich spółek zastosowano okno zdarzenia o stałej szerokości, obejmujące dzień przed zdarzeniem oraz pięć dni po zdarzeniu.

⁶⁴¹ Bouteska, A. *The effect of investor sentiment on market reactions to financial earnings restatements: Lessons from the United States*, "Journal of Behavioral and Experimental Finance", 2019, nr 24, s.6.

⁶⁴² Zob. Nisar, T. M., Southampton, M.Y., *Twitter as a tool for forecasting stock market movements: A short-window event study*, "Journal of Finance and Data Science", 2018, nr 1.

⁶⁴³ Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna a Wolters Kluwer business, Warszawa 2012, s. 38- 39.

Tabela 9. Zastosowanie miar uśredniających do prezentacji wyników analizy zdarzeń dla pojedynczych instrumentów

rok	autorzy	tytuł	czasopismo
2024	Eyden Samunderu, Anita Yordanova	The Effect of Public Contract Winning Announcements on Share Prices: An Event-Based Study on the Pharmaceutical Industry	Journal of Economic Analysis
2023	Sunkung Choi	Measuring economic diplomacy using event study method: the case of EU-China summit talks and Airbus stock price changes	Asia Europe Journal
2022	Hafsa Rasheed, Habib Ahmad, Attiya Yasmin Javid	Gold vs. PSX sectors during political uncertainties: an event study analysis	Business Review
2021	Matthijs Jan Kallen, Bert Scholtens	Movers and Shakers: Stock Market Response to Induced Seismicity in Oil and Gas Business	Energies
2021	Erdem Oncu	The impact of COVID-19 on health sector stock returns	SSRN Electronic Journal
2019	Masaaki Yoshimori Former	Shadow Exchange Rates – Changing the Winds with Headwinds and Tailwinds	SocioEconomic Challenges
2013	Timo Korkeamäki, Tuomas Takalo	Valuation of Innovation and Intellectual Property: The Case of iPhone	European Management Review

Źródło: opracowanie własne.

Szacowanie ponadnormatywnych stóp zwrotu

W następnym korku oszacowane zostały dzienne stopy zwrotu dla wszystkich badanych spółek, zgodnie z poniższym wzorem⁶⁴⁴:

$$r_{it} = \frac{P_{it_2}}{P_{it_1}} - 1$$

P_{i1} - cena zamknięcia w drugim dniu

P_{i2} - cena zamknięcia w drugim dniu

Statystyki opisowe dla dziennych stóp zwrotu znajdują się w załączniku nr 1.

Dzięki temu możliwe było obliczenie ponadnormatywnych stóp zwrotu AR (ang. *abnormal returns*), zgodnie ze wzorem⁶⁴⁵:

$$AR_{it} = R_{it} - E(R_{it})$$

AR_{it} - ponadnormatywna stopa zwrotu i-tej akcji w okresie t

R_{it} - rzeczywista stopa zwrotu i-tej akcji w okresie t

$E(R_{it})$ – oczekiwana stopa zwrotu i-tej akcji w okresie t

Należy zaznaczyć, że chociaż badaniu podlegała reakcja inwestorów na informacje, dotyczące pojedynczych spółek, to przedstawienie wyników przy użyciu miary AR, tj. ponadnormatywnych stóp zwrotu, było niemożliwe. Przyczyną takiego stanu rzeczy była liczba artykułów prasowych, opisujących każdą z analizowanych spółek. Ich średnia liczba wynosiła bowiem 178, przy czym dla niektórych przedsiębiorstw wynik ten przekroczył 400 (np. Polskie Górnictwo Naftowe i Gazownictwo, mBank, Powszechna Kasa Oszczędności Bank Polski). Przeprowadzenie w tym zakresie przeglądu literatury, pozwoliło jednak stwierdzić, że z podobnym problemem spotkali się badacze z różnych obszarów finansów, takich jak: surowce, waluty czy akcje. W takiej sytuacji – niezbędne dla zaprezentowania wyników – okazało się zastosowanie miar uśredniających. W związku z tym, w poniższym badaniu podjęto decyzję o wykorzystaniu skumulowanych ponadnormatywnych średnich stóp zwrotu (CAAR). Wybrane przykłady tego typu badań zaprezentowane zostały w tabeli numer 9.

⁶⁴⁴ Por. Jajuga, K., Jajuga, T. *Inwestycje. Instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa*, Wydawnictwo Naukowe PWN, Warszawa, 2008, s. 93.

⁶⁴⁵ Zob. Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Wydawnictwo Nieoczywiste, Warszawa 2019, s. 51.

Na podstawie wyników ponadnormatywnych stóp zwrotu obliczone zostały więc średnie ponadnormatywne stopy zwrotu AAR⁶⁴⁶:

$$AAR_t = \frac{1}{N_t} \sum_{i=1}^{N_t} AR_{it}$$

AAR_t - średnia ponadnormatywna stóp zwrotu w okresie t

N_t - liczba spółek badanej populacji w okresie t

AR_{it} - ponadnormatywna stopa zwrotu i-tej akcji w okresie t

Ostatnim krokiem było oszacowanie, w ramach okna zdarzenia, skumulowanych średnich ponadnormatywnych stóp zwrotu⁶⁴⁷:

$$CAAR_{t_1}^{t_2} = \sum_{t=t_1}^{t_2} AAR_t$$

$CAAR_{t_1}^{t_2}$ - skumulowana średnia ponadnormatywna stopa zwrotu w „oknie obserwacji”

t_1 - początek „okna obserwacji”

t_2 - koniec okna obserwacji

AAR_t - średnia ponadnormatywna stóp zwrotu w okresie t

Istotność statystyczna średnich skumulowanych nadwyżkowych stóp zwrotu wyznaczana była na podstawie parametrycznego testu t-Studenta⁶⁴⁸:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$\bar{X}_1$ - średnia dla pierwszej grupy

s_1^2 - wariancja pierwszej grupy

n_1 - liczebność pierwszej grupy

$\bar{X}_2$ - średnia dla drugiej grupy

s_2^2 - wariancja drugiej grupy

n_2 - liczebność drugiej grupy

⁶⁴⁶ Ibidem.

⁶⁴⁷ Agrawal A., Jaffe J.F., Mandelker G.N., The post-merger performance of acquiring firms: A re-examination of an anomaly, „The Journal of Finance” 1992, vol. XLVII, no. 4, s. 1609.

⁶⁴⁸ Zob. Sobczyk, M. *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2007, s. 136.

Model szacowania oczekiwanej stopy zwrotu

Do oszacowania wartości oczekiwanych dziennych stóp zwrotu wykorzystany został model rynkowy (ang. *market model*). Oczekiwane stopy zwrotu, dla 30-dniowego okna estymacyjnego, wyznaczone zostały na podstawie poniższego wzoru⁶⁴⁹:

$$E(R_{it}) = \alpha_i + \beta_i R_{m,t}$$

α_i - wyraz wolny

β_i - współczynnik ryzyka i-tej akcji

$R_{m,t}$ - stopa zwrotu indeksu rynkowego w okresie t

Z uwagi na to, że analizie podlegały spółki, wchodzące w skład indeksu WIG20, a ich udział w indeksie w poszczególnych okresach przyjmował wartości od 0,42% do 15,00%, indeks ten nie został przyjęty jako podstawa do obliczenia rynkowych stóp zwrotu w modelu. Ponadto, chociaż udział badanych spółek w szerokim indeksie WIG znajdował się na znacznie niższych poziomach, to wysoka korelacja między indeksami WIG i WIG20 (0,98) oraz dodatnia kowariancja (0,0029), uznane zostały za przesłankę do poszukiwania alternatywnego, zewnętrznego indeksu. W związku z tym za rynkowy benchmark przyjęty został *Stooq Poland All Stocks Price Index*. Indeks ten odzwierciedla zachowanie akcji głównego rynku GPW i NewConnect. W przeciwieństwie do indeksu WIG20, w ramach *Stooq Poland*, dla wszystkich uwzględnianych walorów, przyjmuje się taką samą wagę. Wartość indeksu szacowana jest jako iloczyn wartości indeksu z poprzedniego dnia sesyjnego oraz średniej geometrycznej zmian cen w ujęciu procentowym dla wszystkich walorów z bieżącej sesji⁶⁵⁰. Po wyborze modelu oraz indeksu obliczone zostałyienne stopy zwrotu dla wszystkich badanych spółek. Statystyki opisowe znajdują się w załączniku nr 1. Należy zauważyć, że w próbie badawczej znalazły się spółki, których okres notowania w indeksie WIG20 charakteryzował się nieciągłością. Takimi przypadkami są spółki Asseco oraz Jastrzębska Spółka Węglowa. Z uwagi na to ich okresy badawcze podzielone zostały na dwie części i oznaczone odpowiednio jako: ACP1, ACP2 oraz JSW1, JSW2.

⁶⁴⁹ Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Wydawnictwo Nieoczywiste, Warszawa 2019, s. 42.

⁶⁵⁰ Stooq Poland All Stocks Price Index, https://stooq.pl/q/p/?s=_pl, Dostęp: 1 mar 2022.

Weryfikacja hipotezy badawczej

Dla każdej spółki weryfikowana była hipoteza zerowa:

H_0 : wartość oczekiwana $CAAR_t$ równa jest 0.

Pozytywna weryfikacja powyższej hipotezy oznaczała brak reakcji inwestorów na zdarzenie (tj. na pojawienie się artykułu z określonym sentymentem). Jeżeli natomiast uzyskane wyniki dawały podstawy do jej odrzucenia, oznaczało to wystąpienie reakcji, co do której jednocześnie szacowany był kierunek oraz istotność statystyczna.

Weryfikacja trzeciej hipotezy pomocniczej

Celem trzeciego etapu badań było zweryfikowanie hipotezy, zgodnie z którą sentyment, zawarty w artykułach prasowych, może stanowić zmienną, determinującą wynik dziennych stóp zwrotu, spółek uwzględnionych w próbie badawczej. Do przeprowadzenia tego etapu badań, zastosowano regresję logistyczną, której istotą jest zweryfikowanie możliwości przyporządkowania dychotomicznej zmiennej objaśnianej do jednej z – wcześniej zdefiniowanych – klas, za pomocą wybranych zmiennych objaśniających. Wybór tej metody podyktowany był tym, że znajduje częste zastosowanie w badaniach, związanych z eksploracją tekstu, w obszarze analizy sytuacji na rynkach finansowych. Przykłady tego typu badań zawarte zostały w tabeli numer 10.

Budowa modelu klasyfikacyjnego

W kontekście prezentowanego etapu badań za zmienną objaśniającą przyjęto – obliczony w pierwszym etapie – sentyment artykułów prasowych. Podobnie jak w przypadku drugiego etapu badań pod uwagę wzięte zostały jedynie wyniki skrajne, które można uznać za jednoznaczne. Artykuły o sentymencie negatywnym definiowane były jako te, których wynik mieścił się w przedziale od -0,9 do -1. W przypadku artykułów, które określone zostały jako pozytywne, wskaźnik sentymentu przyjmował wartości od 0,9 do 1.

Tabela 10. Zastosowanie modeli klasyfikacyjnych, w tym regresji logitowej, do badań nad reakcją rynków finansowych i ich podmiotów na otrzymanie informacji niefinansowych

Rok	Autorzy	Tytuł	Czasopismo
2022	Fedorova, E., Drogovoz, P., Nevredinov, A., Kazinina, P., Qitan, C.	Impact of MD&A sentiment on corporate investment in developing economies: Chinese evidence	Asian Review of Accounting
2020	Awan, M. J., Rahim, M. S. M., Nobanee, H., Munawar, A., Yasin, A., Zain, A. M.,	Social Media and Stock Market Prediction: A Big Data Approach	Computers, Materials & Continua
2017	Khedr, A. E., Salama, S. E., Yaseen, N.,	Predicting Stock Market Behavior using Data Mining Technique and News Sentiment Analysis	International Journal of Intelligent Systems and Applications
2016	Ghailan O., Mokhtar H.M.O., Hegazy O.	Improving Credit Scorecard Modeling Through Applying Text Analysis	International Journal of Advanced Computer Science and Applications
2016	Li J., Xu Z., Yu L., Tang L.	Forecasting Oil Price Trends with Sentiment of Online News Articles	Procedia Computer Science
2015	Maslyuk-Escobedo, S., Rotaru, K., Dokumentov, A.,	News sentiment and jumps in energy spot and futures markets	Pacific Basin Finance Journal
2014	Feuerriegel, S., Lampe, M. X., Neumann, D.,	News processing during speculative bubbles: Evidence from the oil market	47th Hawaii International Conference on System Science
2014	Kim, Y., Jeong, S. R., Ghani, I.,	Text Opinion Mining to Analyze News for Stock Market Prediction	International Journal of Advances in Soft Computing and Its Applications
2013	Lu Y-Ch., Shen Ch-H., Wei Y-Ch.	Revisiting early warning signals of corporate credit default using linguistic analysis	Pacific-Basin Finance Journal
2011	Ravisankar P., Ravi V., Rao G. R., Bose I.	Detection of financial statement fraud and feature selection using data mining techniques	Decision Support Systems

Źródło: opracowanie własne.

Dla każdej ze spółek zbudowany został model regresji logitowej. Podstawą do oszacowania jego wyników była funkcja logistyczna, którą zapisać można w następujący sposób⁶⁵¹:

$$Y_t = \frac{e^{Z_t}}{1 + e^{Z_t}}$$

Y_t – prawdopodobieństwo wystąpienia danego zdarzenia w chwili t

Z_t – kombinacja liniowa cech wejściowych w chwili t ($Z_t = \beta_0 + \beta_1 x_{1t} + \dots \beta_k x_{kt}$)

e – liczba Eulera

Zmienna objaśniana modelu rozumiana była jako dzienna stopa zwrotu z poszczególnych akcji, przekształcona w ten sposób, że do wyników ujemnych przypisano klasę „0”, natomiast do wyników dodatnich, klasę „1”.

Dla oceny poziomu dokładności oszacowania modelu logitowego zastosowany został współczynnik Pseudo-R² McFaddena, który definiowany jest poniższym wzorem⁶⁵². Do estymacji wszystkich modeli zastosowano Metodę Największej Wiarygodności (ang. *Maximum Likelihood Estimation* (MLE)).

$$R_{McFaddena}^2 = 1 - \frac{\ln L_p}{\ln L_0}$$

$\ln L_p$ – logarytm funkcji wiarygodności dla pełnego modelu

$\ln L_0$ – logarytm funkcji wiarygodności dla modelu zawierającego jedynie wyraz wolny

Współczynniki oceny jakości klasyfikacji

Uzyskane wyniki klasyfikacji zaprezentowane zostały za pomocą czterech miar, pozwalających scharakteryzować efektywność zbudowanego modelu. Są to współczynniki: precyzji (ang. *precision*), czułości (ang. *sensitivity*), swoistości (ang. *specificity*) oraz dokładności (ang. *accuracy*). Powyższe miary oszacowane zostały za pomocą następujących wzorów⁶⁵³:

⁶⁵¹ Gruszczyński, M., Kuszewski, T., Podgórska, M. *Ekonometria i badania operacyjne*, Wydawnictwo Naukowe PWN, Warszawa 2009, s. 142.

⁶⁵² McFadden, D. *Conditional logit analysis of qualitative choice behavior*, [w:] (red.) Zarembka, P. *Frontiers in Econometrics*, Academic Press, Nowy Jork, 1974, s.121.

⁶⁵³ Zob. Fawcett, T. *An introduction to ROC analysis*, "Pattern Recognition Letters", R. 2006, nr 27, s. 861–874. Oraz Powers, D. *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness*,

$$\text{prezycja} = \frac{TP}{TP + FP}$$

$$\text{swoistość} = \frac{TN}{TN + FP}$$

$$\text{czułość} = \frac{TP}{TP + FN}$$

$$\text{dokładność} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP – Prawdziwe Pozytywne;
TN – Prawdziwe Negatywne;

FP – Fałszywe Pozytywne;
FN – Fałszywe Negatywne

Weryfikacja hipotezy badawczej

Przyjęto, że dla pozytywnej weryfikacji trzeciej hipotezy pomocniczej wymagane jest, aby wynik współczynnika pseudo-R² – dla więcej niż połowy badanych spółek – charakteryzował się statystyczną istotnością na poziomie $\alpha \leq 0,1$. Niespełnienie tego warunku oznaczać będzie konieczność odrzucenia tej hipotezy badawczej.

6.1.3. Charakterystyka próby badawczej

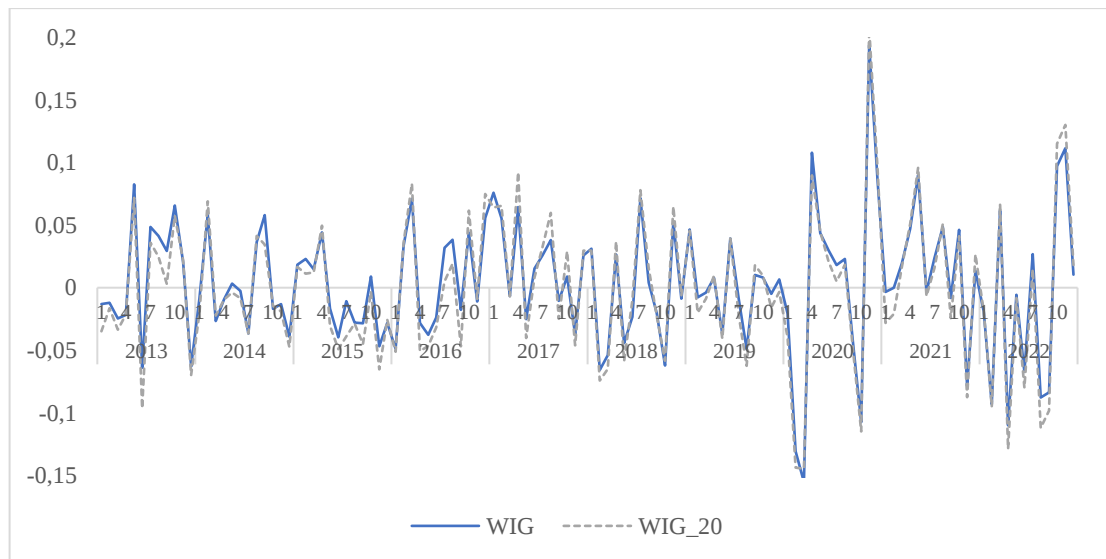
Próba badawcza konstruowana była na podstawie spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie i wchodzących w skład indeksu WIG20 w okresie od 01.01.2013 do 31.12.2022. Indeks WIG20 jest indeksem cenowym, który obejmuje 20 spółek o najwyższych obrotach, szacowanych na podstawie ostatnich 12 miesięcy oraz najwyższej wartości akcji w wolnym obrocie⁶⁵⁴. Wybór tego indeksu był powodowany jego wysoką płynnością oraz kryteriami klasyfikacyjnymi, które wynikają – między innymi – z poziomu kapitalizacji spółek. Z jednej strony zapewniło to dostęp do dziennych cen zamknięcia akcji, co umożliwiło obliczenie stóp zwrotu w krótkim terminie. Z drugiej jednak strony, opierając się na wnioskach wynikających z rozdziału III niniejszej dysertacji, można było przyjąć założenie, że spółki wchodzące w skład tego indeksu, będą cieszyły się zainteresowaniem dziennikarzy, co znajdzie swoje odzwierciedlenie w liczbie publikowanych na ich temat artykułów. Ponadto można stwierdzić, że indeks ten relatywnie dobrze odwzorowuje ogólną sytuację rynkową. Analiza dziennych stóp zwrotu indeksu WIG20 oraz WIG – w badanym okresie – wykazała zachodzącą między nimi korelację na poziomie 0,98. Można też zauważyć,

Markedness & Correlation, “International Journal of Machine Learning Technology”, R. 2011, nr 2/1, s. 37-63.

⁶⁵⁴ Giełda Papierów Wartościowych w Warszawie: WIG20 – Metodologia indeksu. Stan na 2022-12-30. 2022.

że dziesięcioletni okres badawczy pozwolił uchwycić zarówno okresy wzrostów rynku, jak i spadków, co zaprezentowane zostało na wykresie nr 2. Dzięki temu, badania analizy zdarzeń zostały zrealizowane, a zbudowany model przetestowany, w warunkach, które cechowały się różnorodnością.

Wykres 2. *Dynamika miesięcznych stóp zwrotu indeksów WIG20 i WIG w okresie 01.2013 – 12.2022*



Źródło: opracowanie własne.

W badanym okresie w skład indeksu WIG20 wchodziło 39 spółek. Przyjęto dwa założenia, pozwalające zaklasyfikować spółki do dalszych analiz. Pierwsze założenie związane było z długością notowania poszczególnych spółek w ramach indeksu. Indeks WIG20 podlega korektom oraz rewizjom w każdy trzeci piątek marca, czerwca, września oraz grudnia⁶⁵⁵. Oznacza to, że skład indeksu – co każdy kwartał – mógł ulegać zmianie. Na tej podstawie odrzucone zostały spółki, których okres występowania w indeksie był krótszy niż cztery kwartały. Drugie założenie natomiast związane było z liczbą dostępnych artykułów prasowych. W ogólnym ujęciu na temat wszystkich notowanych w okresie badawczym spółek, zgromadzono 285 174 publikacje prasowe w języku angielskim. Biorąc jednak pod uwagę, że cały dziesięcioletni okres badawczy to 2 497 dni sesyjnych, a na drugim etapie badania zastosowana została metoda analizy zdarzeń, zasadne wydawało się ograniczenie liczby artykułów.

⁶⁵⁵ Giełda Papierów Wartościowych w Warszawie: WIG20 – Metodologia indeksu. Stan na 2022-12-30. 2022.

Tabela 11. *Spółki zakwalifikowane do próby badawczej*

spółka			okres badawczy		liczba artykułów	
l.p.	nazwa	symbol	początek	koniec	całość	tytuły
1.	Asseco Poland	ACP	31.12.2012	31.12.2022	6 398	1 259
2.	Allegro.eu	ALE	18.12.2020	31.12.2022	979	202
3.	Alior Bank	ALR	21.03.2014	19.03.2021	9 991	1 506
4.	Bank Handlowy w Warszawie	BHW	21.12.2012	21.03.2014	1 581	168
5.	Bre Bank	BRE	21.12.2012	21.03.2014	1 420	190
6.	Bank Zachodni WBK	BZW	21.06.2013	21.12.2018	3 305	238
7.	CCC	CCC	18.12.2015	31.12.2022	6 738	1 001
8.	CD Projekt	CDR	16.03.2018	31.12.2022	2 496	241
9.	Cyfrowy Polsat	CPS	20.03.2015	31.12.2022	6 581	1 363
10.	Dino Polska	DNP	15.03.2019	31.12.2022	1 482	140
11.	Enea	ENA	20.03.2015	17.03.2017	3 832	717
12.	Energa	ENG	20.03.2015	15.03.2019	6 870	1 039
13.	Eurocash	EUR	15.03.2013	15.03.2019	6 359	874
14.	Globe Trade Centre	GTC	21.12.2012	21.03.2014	1 385	181
15.	Jastrzębska Spółka Węglowa	JSW	31.12.2012	31.12.2022	9 160	1 720
16.	Kernel Holding	KER	21.12.2012	20.03.2015	2 231	176
17.	KGHM Polska Miedź	KGH	31.12.2012	31.12.2022	16 825	3 403
18.	LPP	LPP	21.03.2014	31.12.2022	7 829	1 246
19.	Lotos	LTS	31.12.2012	16.09.2022	11 440	2 366
20.	Lubelski Węgiel Bogdanka	LWB	31.12.2012	18.12.2015	4 137	727
21.	mBank	MBK	21.03.2014	31.12.2022	12 148	1 466
22.	Orange Polska	OPL	21.03.2014	31.12.2022	10 571	3 018
23.	Bank Polska Kasa Opieki	PEO	31.12.2012	31.12.2022	19 613	2 139
24.	PGE Polska Grupa Energetyczna	PGE	31.12.2012	31.12.2022	18 706	4 288
25.	Polskie Górnictwo Naftowe i Gazownictwo	PGN	31.12.2012	16.12.2022	19 870	5 906
26.	Polski Koncern Naftowy ORLEN	PKN	31.12.2012	31.12.2022	19 443	7 641
27.	Powszechna Kasa Oszczędności Bank Polski	PKO	31.12.2012	31.12.2022	24 207	2 565
28.	Play Communications	PLY	15.03.2019	18.12.2020	5 068	435
29.	Powszechny Zakład Ubezpieczeń	PZU	31.12.2012	31.12.2022	18 985	2 461
30.	Synthos	SNS	21.12.2012	18.03.2016	2 660	350
31.	Santander Bank Polska	SPL	21.12.2018	31.12.2022	3 854	429
32.	Tauron Polska Energia	TPE	21.12.2012	18.03.2022	15 661	7 260
33.	Telekomunikacja Polska	TPS	21.12.2012	21.03.2014	1 721	238
suma artykułów					283 546	56 953

Źródło: opracowanie własne.

Zawężenie próby polegało na wyselekcjonowaniu tych artykułów, które w największym stopniu związane były z poszczególnymi spółkami. Przyjęto więc zasadę, że w badaniu wykorzystane zostaną tylko te artykuły, w których tytułach pojawiały się nazwy analizowanych spółek. Pozwoliło to wyselekcjonować 57 196 artykułów prasowych. Z uwagi jednak na to, że do przeprowadzenia analizy zdarzeń, konieczne było wskazanie publikacji, charakteryzujących się skrajnym sentymentem, konieczne było dobranie takich spółek, do których przypisana liczba artykułów, pozwoli na zdefiniowanie więcej niż jednego zdarzenia. Dlatego z próby badawczej wyeliminowane zostały spółki, na temat których opublikowano mniej niż 100 artykułów. Na podstawie obu kryteriów do badania zaklasyfikowano 30 spółek. Należy jednak zaznaczyć, że w przypadku części z nich notowanie w indeksie WIG20 charakteryzowało się nieciągłością. Taka sytuacja dotyczyła spółki Asseco Poland oraz Jastrzębskiej Spółki Węglowej. Ponadto w okresie badawczym trzy spółki, notowane w indeksie WIG20, dokonały *rebrandingu*. Były to: Telekomunikacja Polska (obecnie Orange Polska), Bank Zachodni WBK (obecnie Santander Bank Polska) oraz Bre Bank (obecnie mBank). W każdym z tych przypadków przyjęto założenie, że będą one analizowane jako oddzielne spółki, co rozszerzyło próbę do 35 podmiotów.

Tabela 12. *Spółki niewchodzące w skład próby badawczej*

spółka			rewizje / korekty	liczba artykułów	
l.p.	nazwa	symbol		całość	tytuły
1.	Boryszew SA	BRS	2	480	50
2.	Kruk	KRU	1	16	2
3.	Grupa Kęty	KTY	2	93	6
4.	Mercator Medical	MRC	4	105	15
5.	Pepco Group	PCO	4	322	79
6.	TVN	TVN	1	612	91
suma artykułów				1 628	243

Źródło: opracowanie własne.

Spółki przyjęte do badania oraz odrzucone zaprezentowane zostały odpowiednio w tabelach 11 i 12. Mimo tego, że 6 spółek, notowanych w badanym okresie w indeksie WIG20, nie zostało zakwalifikowanych do dalszych badań, skład próby badawczej w większości okresów rewizyjnych w pełni odzwierciedlał ów indeks. Wyjątek stanowił

rok 2022, gdzie udział badanych spółek w indeksie wahał się od 92 do 95%. Szczegółowa struktura indeksu – we wszystkich badanych kwartałach – zamieszczona została w załączniku nr 2.

6.1.4. Ograniczenia badawcze

Należy jednak zauważyć, że zaprezentowane badania podlegały pewnym ograniczeniom. Wykorzystany w pierwszym etapie badania, słownik nie jest jedynym dostępnym i umożliwiającym ekstrakcję sentymentu. Innym popularnym słownikiem, również w środowisku badaczy z zakresu finansów, jest *General Inquirer Harvard IV-4*⁶⁵⁶. Chociaż słownik ten tworzony był raczej z myślą o zastosowaniu w dziedzinie psychologii, to być może zaaplikowanie go przypisałoby poszczególne artykuły do innych klas, niż stało się to przy użyciu słownika *Loughran-McDonald Master Dictionary*.

W drugim etapie badania – w celu oszacowania oczekiwanych stóp zwrotu w ramach okna estymacyjnego – wykorzystany został model rynkowy. Zastosowanie innych modeli, takich jak: model wyceny aktywów kapitałowych (CAPM), model średniej (MAR), model indeksowy (IM), trójczynnikiowy czy też pięcioczynnikowy model Famy-Frencha, mogłoby potencjalnie przyczynić się do predykcji oczekiwanych stóp zwrotu na innych poziomach.

W trzecim etapie badania zbudowany został model, w ramach którego wykorzystano tylko jedną zmienną objaśniającą. Istnieją jednak przesłanki, by sądzić, że reakcja rynku na napływające informacje, charakteryzujące się pozytywnym lub negatywnym sentymentem, zależna będzie od ogólnej sytuacji rynkowej. Jednym z argumentów, przemawiających za taką możliwością, jest – opisana w rozdziale pierwszym – teoria perspektywy. Jej autorzy podkreślają wagę tak zwanego punktu referencyjnego, w kontekście podejmowania decyzji inwestycyjnych⁶⁵⁷. Być może więc włączenie do modelu zmiennych, opisujących ogólną sytuację rynkową, byłoby działaniem pożądanym. Ponadto model logitowy jest tylko jedną z możliwości weryfikacji hipotezy,

⁶⁵⁶ Zob. General Inquirer, <https://inquirer.sites.fas.harvard.edu/>. Dostęp: 1 kw. 2023.

Zob. Tetlock, P. C., *Giving Content to Investor Sentiment: The Role of Media in the Stock Market Published*, „The Journal of Finance”, 2007, nr LXII/3, s. 1139–1168.

⁶⁵⁷ Zob. Kahneman, D., Tversky, A., *Prospect Theory: An Analysis of Decision under Risk*, „The Econometric Society”, 1979, nr 47/2, s. 263.

mówiącej o możliwości zdefiniowania, zawartego w artykułach prasowych, sentymentu jako zmiennej determinującej reakcję rynku. Model ten, chociaż stosowany do rozwiązania problemów związanych z klasyfikacją, cechuje się ograniczeniami w objaśnianiu zmiennych o charakterze ciągłym.

Mając świadomość powyższych ograniczeń, autorka podjęła się weryfikacji postawionych hipotez badawczych. Do tego celu zastosowana została – opisana powyżej – procedura badawcza, w ramach której przeprowadzono proces ekstrakcji sentymentu ze zgromadzonych artykułów prasowych, dokonano analizy zdarzeń oraz estymacji modelu regresji logitowej. Uzyskane rezultaty zaprezentowane zostały w drugiej części niniejszego rozdziału.

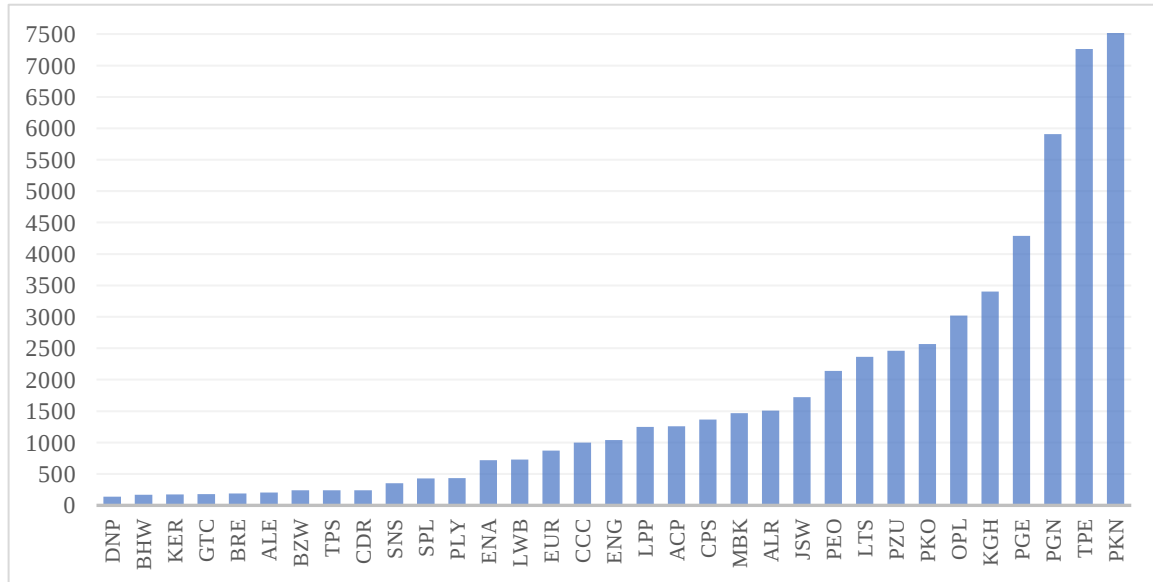
6.2. Badania własne

6.2.1. Gromadzenie i przygotowanie danych

Ogólna liczba zgromadzonych artykułów wynosiła 285 174. Aby jednak zredukować wielowymiarowość danych oraz zminimalizować zjawisko szumów, przyjęto założenie o wykorzystaniu w badaniach jedynie tych artykułów, w których tytułach pojawiały się nazwy analizowanych spółek. Ograniczyło to liczbę artykułów do 57 196 sztuk. Dodatkowo mając świadomość, że do przeprowadzenia drugiego i trzeciego etapu badania, wykorzystane zostaną wyłącznie skrajne wyniki – zawartego w tekstach – sentymentu, konieczny był taki dobór próby, aby możliwe było wskazanie artykułów o tej charakterystyce. W związku z tym jako minimalną liczbę artykułów, pozwalających zakwalifikować spółkę do dalszego badania, przyjęto 100. W efekcie w badaniu wykorzystano próbę składającą się z 56 953 publikacji prasowych w języku angielskim.

Należy zaznaczyć, że spółki - wchodzące w skład próby badawczej - różniły się pod względem liczby, opublikowanych na ich temat, artykułów. Na przykład w przypadku spółki Allegro, która w okresie badawczym w indeksie WIG20 notowana była przez 9 kwartałów, udało się zgromadzić zaledwie 202 artykuły, w tytułach których występowała nazwa spółki. Natomiast na temat spółki PKN Orlen, która znajdowała się w indeksie przez cały okres badawczy (41 kwartałów), zgromadzono 7 203 artykuły.

Wykres 3. Liczba zgromadzonych artykułów dla spółek, znajdujących się w próbie badawczej

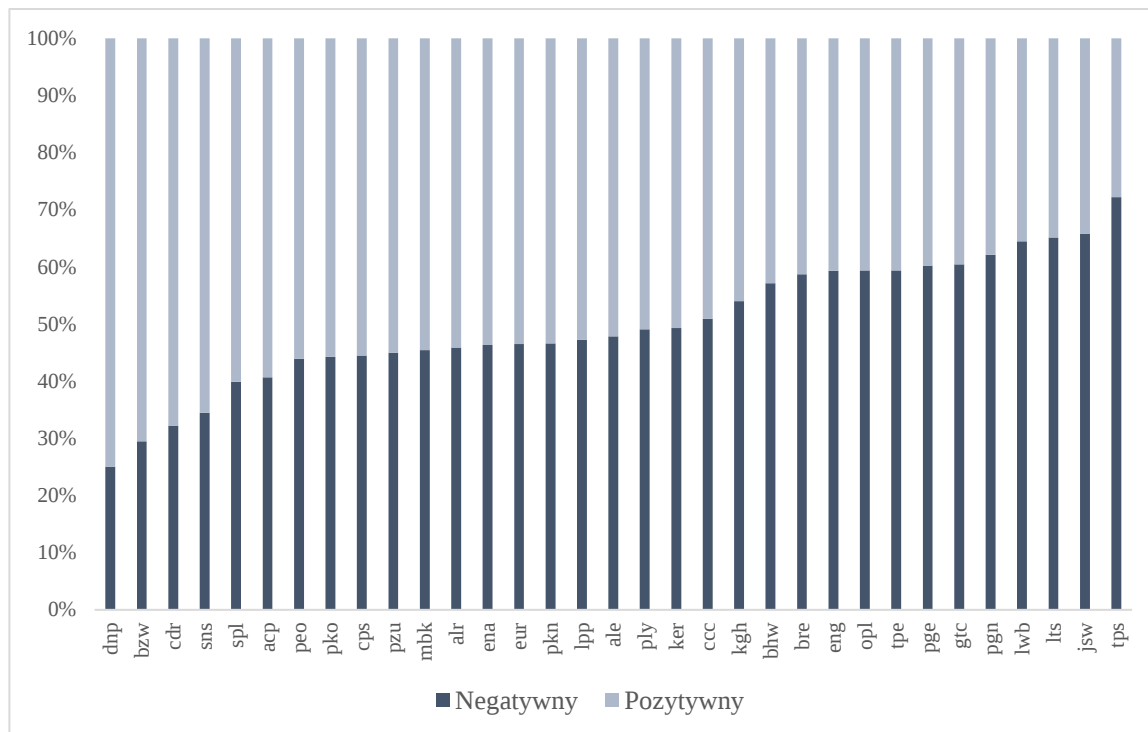


Źródło: opracowanie własne.

6.2.2. Wyniki pierwszego etapu badań

Jak można zauważyć na podstawie tabeli nr 13, w przypadku 19 badanych spółek dominujący udział – wśród wszystkich przeanalizowanych artykułów – miał sentyment neutralny. Dla 9 spółek był to sentyment pozytywny. Sentyment negatywny można określić jako dominujący w przypadku 5 spółek. Z uwagi jednak na cel tego etapu badania, jakim jest weryfikacja pierwszej hipotezy pomocniczej oraz fakt, że sentyment neutralny nie jest brany pod uwagę w etapach II i III, analizie należy poddać relację, jaka zachodzi wyłącznie między sentymentem pozytywnym i negatywnym, co zaprezentowane zostało na wykresie nr 4.

Wykres 4. Struktura artykułów zaklasyfikowanych jako pozytywne lub negatywne



Źródło: opracowanie własne.

W tym wariancie zauważyć można, że w przypadku większości spółek dominował sentyment pozytywny. Taki rezultat zauważalny był wśród 19 spółek, co stanowiło 58% całej próby. Odwrotna sytuacja dotyczyła 14 spółek, czyli 42% próby badawczej. Jako spółki, u których sentyment negatywny wyraźnie przewyższał pozytywny, uznać można przede wszystkim Telekomunikację Polską (72% do 28%), Jastrzębską Spółkę Węglową (66% do 34%), Lotos (65% do 35%), Bogdanę (64% do 36%) oraz Polskie Górnictwo Naftowe i Gazownictwo (62% do 38%). Spółki, które natomiast odznaczyły się najwyższą przewagą sentymentu pozytywnego nad negatywnym, to: Dino (75% do 25%), Bank Zachodni (71% do 29%), CD Projekt (68% do 32%) oraz Synthos (66% do 34%). Należy jednak zwrócić uwagę, że w przypadku spółki CCC różnica między liczbą artykułów o sentymencie negatywnym i pozytywnym wynosiła jedynie 2 p.p., co w kontekście tego typu analiz trudno uznać za wynik jednoznaczny. Szczegółowe wyniki struktury sentymentu zaprezentowane zostały w tabeli nr 13.

Tabela 13. *Liczba artykułów przypisanych do poszczególnych klas sentymentu*

Spółka	Sentyment			Sentyment [%]		
	Neg	Poz	Neu	Neg	Poz	Neu
TPS	122	47	68	29%	51%	20%
JSW	839	436	421	25%	49%	26%
LTS	1 036	554	770	33%	44%	23%
LWB	343	189	191	26%	47%	26%
PGN	2 189	1 334	2 331	40%	37%	23%
GTC	81	53	47	26%	45%	29%
PGE	1 310	866	2 080	49%	31%	20%
TPE	319	218	4 451	89%	6%	4%
OPL	951	650	1 197	43%	34%	23%
ENG	385	264	390	38%	37%	25%
BRE	54	38	96	51%	29%	20%
BHW	60	45	63	38%	36%	27%
KGH	1 093	931	1 365	40%	32%	27%
CCC	279	269	453	45%	28%	27%
KER	70	72	34	19%	40%	41%
PLY	104	108	222	51%	24%	25%
ALE	55	60	87	43%	27%	30%
LPP	341	381	523	42%	27%	31%
PKN	2 097	2 400	2 706	38%	29%	33%
EUR	246	283	345	39%	28%	32%
ENA	214	248	255	36%	30%	35%
ALR	452	533	521	35%	30%	35%
MBK	451	541	474	32%	31%	37%
PZU	184	225	339	45%	25%	30%
CPS	413	515	435	32%	30%	38%
PKO	762	960	836	33%	30%	38%
PEO	235	300	358	40%	26%	34%
ACP	401	585	271	22%	32%	47%
SPL	128	193	108	25%	30%	45%
SNS	72	137	140	40%	21%	39%
CDR	67	141	33	14%	28%	59%
BZW	46	110	88	36%	19%	45%
DNP	17	51	72	51%	12%	36%

Neg – sentyment negatywny; Poz – sentyment pozytywny; Neu – sentyment neutralny.

Źródło: opracowanie własne.

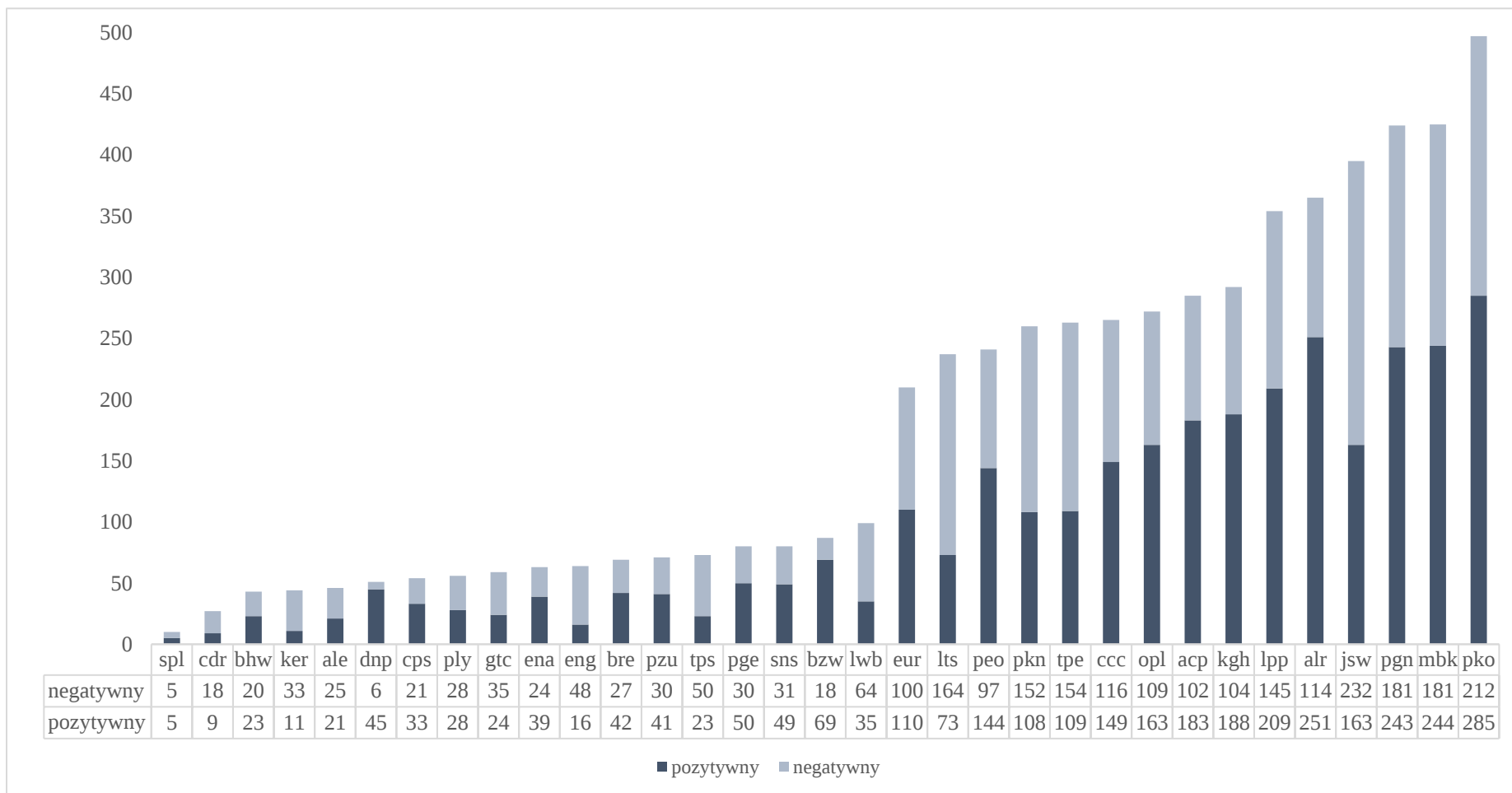
Przeprowadzenie procesu ekstrakcji sentymentu oraz przypisanie wyników jego wskaźnika do każdego z anglojęzycznych artykułów prasowych, pozwoliło stwierdzić, że w okresie od stycznia 2013 do grudnia 2022 większość badanych spółek opisywanych było przez dziennikarzy w sposób pozytywny. Tym samym nie znaleziono podstaw do odrzucenia pierwszej hipotezy pomocniczej.

6.2.3. Wyniki drugiego etapu badań

Jak zostało zaznaczone w pierwszej części rozdziału liczba artykułów, które zakwalifikowano do drugiego etapu badania, została zredukowana. Pod uwagę brane były tylko te artykuły, których wskaźnik sentymentu mieścił się w przedziałach od 0,9 do 1 dla sentymentu pozytywnego oraz od -0,9 do -1 dla sentymentu negatywnego. Pojawienie się w okresie badawczym artykułów o skrajnych wynikach sentymentu, zdefiniowane zostało jako zdarzenia, które podlegały analizie. W rezultacie do analizy włączonych zostało 5 861 artykułów, spośród których 3 185 charakteryzowało się sentymentem pozytywnym, a 2 676 negatywnym. Liczba zdarzeń dla każdej ze spółek – z podziałem na pozytywne i negatywne – zaprezentowany został na wykresie numer 5.

W przypadku wyników uzyskanych podczas analizy artykułów, które sklasyfikowane zostały jako skrajnie negatywne, istotność statystyczną nadwyżkowych stóp zwrotu w ramach okna zdarzenia, wykazało 71% spółek (25). Dla większości spółek istotność statystyczna stóp zwrotu wystąpiła jednak w oknie od dnia zdarzenia do dnia po zdarzeniu. Taki wynik uzyskany został w przypadku 18 spółek. Wraz z rozszerzaniem się okna zdarzenia liczba spółek, których CAAR charakteryzował się istotnością, zmniejszała się odpowiednio do: 12, 10, 10, 9 spółek. Można jednak zauważyć, że wyniki różnią się zarówno pod względem poziomu istotności, jak i szerokości okna, w którym skumulowane średnie ponadnormatywne stopy zwrotu okazywały się istotne. Spółki takie jak: CCC, Bogdanka, Orange, Pekao, PKN Orlen oraz Telekomunikacja Polska charakteryzowały się zarówno wysokim poziomem istotności statycznej, jak i kierunkiem stopy zwrotu, odpowiadającym kierunkowi sentymentu od dnia zdarzenia, aż do ostatniego dnia, uwzględnionego w oknie.

Wykres 5. *Struktura sentymentu artykułów, wykorzystanych w analizie zdarzeń*



Źródło: opracowanie własne.

Zauważyć można również, że wszystkie wymienione spółki (poza CCC) wykazują wzrostową tendencję CAAR wraz z rozszerzaniem się okna. W przypadku 11 spółek reakcję inwestorów na napływające negatywne informacje można określić jako dość gwałtowną i relatywnie silną. Poza wymienionymi wyżej spółkami są to: Bank Handlowy, CD Projekt, Energa, LPP oraz Polskie Górnictwo Naftowe i Gazownictwo. W ich przypadku wysoka istotność stóp zwrotu zauważalna była już w dniu zdarzenia, ale jej poziom malał (choć w różnym tempie) wraz z rozszerzaniem okna. Część spółek wykazała jednak odwrotną tendencję. Taką sytuację zaobserwować można w przypadku Banku Zachodniego oraz Play Communications. Należy również zwrócić uwagę na to, że niektóre spółki, chociaż ich wyniki charakteryzowały się istotnością statystyczną stóp zwrotu, to alfa była równa tylko 0,1. W ich przypadku również sam efekt – w postaci reakcji rynku na informację – był raczej krótkotrwały. W tym kontekście wymienić można spółki: Eurocash, Jastrzębską Spółkę Węglową, KGHM, PKO BP oraz Tauron.

Analiza artykułów, które sklasyfikowano jako skrajnie pozytywne, wskazywała na wyniki, charakteryzujące się istotnością statystyczną stóp zwrotu u 57% badanych spółek (20). Zauważyć można, że większość z nich (16) wykazała istotność statystyczną CAAR w piątym dniu po zdarzeniu. W pierwszym oraz drugim dniu było to 11 spółek, w dniu trzecim 14, a w dniu czwartym i szóstym 13. Podobnie jak w przypadku sentymentu negatywnego w – przypadku części spółek – reakcję inwestorów scharakteryzować można jako raczej krótkotrwałą. Dotyczy to spółek takich jak: Bre Bank, Dino, Enea, Energa, Jastrzębska Spółka Węglowa (w drugim okresie) oraz KGH. W przypadku części spółek, takich jak Eurocash, Bogdanka, mBank, PKN Orlen, zaobserwować można było istotność statystyczną stóp zwrotu w całym oknie zdarzenia (PKO BP, Play Communications, Tauron oraz Telekomunikacja Polska). Należy jednak zauważyć, że aż w przypadku 12 spółek skumulowane średnie ponadnormatywne stopy zwrotu przyjmowały wartości ujemne. W kontekście sentymentu negatywnego taka sytuacja dotyczyła jedynie trzech spółek. Oznacza to, że rynek reagował w sposób odwrotny do charakterystyki sentymentu.

Tabela 14. Wartości CAAR oszacowane w wyniku reakcji na artykuły prasowe, charakteryzujące się negatywnym sentymentem

spółka	okno zdarzenia					
	[-1:0]	[-1:1]	[-1:2]	[-1:3]	[-1:4]	[-1:5]
BHW	-0.016 **	-0.023 ****	-0.025 **	-0.022 **	-0.017	-0.016
BRE	-0.01 *	-0.013 **	-0.008	-0.011	-0.007	-0.006
BZW	0.009	0.011	0.012	0.021 *	0.023 **	0.025 *
CCC	-0.018 ****	-0.015 ****	-0.013 ****	-0.014 **	-0.015 **	-0.013 **
CDR	-0.029 ****	-0.04 ****	-0.018	-0.01	-0.018	-0.021
CPS	-0.01 *	-0.012 *	-0.009	-0.005	-0.008	-0.008
ENA	-0.008	-0.016 **	-0.015 *	-0.008	-0.0	-0.004
ENG	-0.011 **	-0.01 *	-0.003	-0.005	-0.002	-0.002
EUR	-0.004	-0.006 *	-0.006	-0.006	-0.007	-0.007
JSW1	0.009	0.013	0.02 **	0.016	0.007	-0.003
JSW2	-0.005 *	-0.003	-0.002	-0.004	-0.002	-0.002
KER	-0.005	-0.015 *	-0.025 ****	-0.037 ****	-0.033 ****	-0.029 **
KGH	-0.002	0.002	0.002	0.003	0.007	0.01 *
LPP	-0.008 ****	-0.008 ****	-0.004	-0.002	-0.0	0.003
LWB	-0.017 ****	-0.031 ****	-0.041 ****	-0.049 ****	-0.049 ****	-0.054 ****
OPL	-0.007 ****	-0.008 ****	-0.01 ****	-0.012 ****	-0.01 **	-0.011 **
PEO	-0.009 ****	-0.011 ****	-0.013 ****	-0.019 ****	-0.019 ****	-0.022 ****
PGE	-0.01 *	-0.017 **	-0.009	-0.012	-0.015 *	-0.014
PGN	-0.012 ****	-0.015 ****	-0.014 **	-0.012 **	-0.01	-0.01
PKN	-0.008 ****	-0.008 ****	-0.006 *	-0.007 *	-0.007 *	-0.012 **
PKO	-0.002	-0.004 *	-0.002	-0.002	-0.0	-0.0
PLY	-0.013	-0.013	-0.023	-0.023	-0.033 *	-0.039 *
PZU	-0.0	-0.005	-0.011 **	-0.004	-0.003	0.0
TPE	-0.006 *	-0.005	-0.006	-0.007	-0.005	-0.008
TPS	-0.058 ****	-0.064 ****	-0.067 ****	-0.08 ****	-0.084 ****	-0.088 ****

****/**/* – istotność wyników na poziomie $\alpha = 0.01/0.05/0.1$

Źródło: opracowanie własne.

Tabela 15. Wartości CAAR oszacowane w wyniku reakcji na artykuły prasowe, charakteryzujące się pozytywnym sentymentem

spółka	okno zdarzenia					
	[-1:0]	[-1:1]	[-1:2]	[-1:3]	[-1:4]	[-1:5]
ALR	0.011 **	0.008 *	0.013 **	0.013 **	0.018 **	0.022 ***
BRE	-0.007 *	-0.009 *	-0.005	-0.01	-0.01	-0.009
BZW	0.003	0.009 **	0.011 **	0.014 ***	0.019 ***	0.022 ***
CDR	0.015	0.024 *	0.031 *	0.038 **	0.035 *	0.023
CPS	0.003	0.002	-0.002	0.006	0.01	0.005
DNP	0.003	0.016 **	0.014 *	0.01	0.01	0.015
ENA	-0.006	-0.006	-0.01	-0.011	-0.018 *	-0.022 **
ENG	-0.01	-0.014	-0.014	-0.016	-0.026 *	-0.022
EUR	-0.005 *	-0.01 ***	-0.012 ***	-0.014 ***	-0.018 ***	-0.017 ***
JSW1	-0.004	0.006	-0.008	-0.018 *	-0.019 *	-0.025 **
JSW2	-0.001	-0.002	-0.009 *	-0.008	-0.006	-0.005
KGH	-0.004 ***	-0.003	-0.002	0.0	-0.001	-0.002
LTS	-0.002	0.008	0.02 **	0.022 **	0.028 **	0.027 **
LWB	-0.014 ***	-0.025 ***	-0.027 ***	-0.031 ***	-0.034 ***	-0.027 ***
MBK	0.003 *	0.003	0.005 *	0.009 ***	0.008 **	0.005
PEO	0.003	0.003	-0.002	-0.004	-0.007 *	-0.006 *
PKN	-0.006 **	-0.008 **	-0.005	-0.009 **	-0.01 **	-0.01 **
PKO	0.004 ***	0.003 *	0.005 **	0.004 **	0.007 ***	0.007 ***
PLY	0.041 ***	0.064 ***	0.064 ***	0.065 ***	0.068 ***	0.065 ***
TPE	-0.008 **	-0.012 **	-0.013 **	-0.017 ***	-0.023 ***	-0.024 ***
TPS	-0.024 ***	-0.024 ***	-0.034 ***	-0.03 ***	-0.04 ***	-0.038 ***

***/**/* – istotność wyników na poziomie $\alpha = 0.01/0.05/0.1$

Źródło: opracowanie własne.

Ponadto poziom istotności statystycznej stóp zwrotu spółek takich jak Eurocash, Bogdanka, PKN Orlen, Tauron czy Telekomunikacja Polska można określić jako wysoki i widoczny w całym lub przynajmniej w większej części okna zdarzenia. Zauważalne jest, że wyniki wszystkich tych spółek wykazały się istotnymi statystycznie ujemnymi wartościami CAAR podczas analizy sentymentu negatywnego. Stanowi to przesłankę do stwierdzenia, że w ich przypadku, niezależnie od sentymentu – charakteryzującego

otrzymaną informację – inwestorzy reagowali w sposób negatywny. Warto zauważyć, że w przypadku większości spółek, których CAAR osiągał ujemne wartości w reakcji na informacje, charakteryzujące się pozytywnym sentymentem, liczba artykułów zdefiniowanych jako skrajnie negatywna była większa lub zbliżona do liczby artykułów zdefiniowanych jako skrajnie pozytywne. Co do zasady w ramach całej próby widoczna była tendencja odwrotna. Dodatkowo w przypadku przynajmniej części z omawianych spółek średnia sentymentu w całym okresie badawczym osiągała wartości ujemne. Stanowi to przesłankę do stwierdzenia, że odpowiedzi na pytanie o przyczynę takiego stanu rzeczy należy szukać w czynnikach fundamentalnych, które mogły być przyczyną napływania na rynek informacji, wykazujących wysoki poziom negatywnego sentymentu.

Tabela 16. Wartości CAAR dla spółek, które w oknie zdarzenia, nie wykazały istotnych statystycznie stóp zwrotu reakcji na sentyment zawarty w artykułach prasowych

	okno zdarzenia	Spółka					
		ACP1	ACP2	ALE	GTC	SNS	SPL
sentyment pozytywny	[-1:0]	-0.001	-0.004	0.001	-0.003	0.001	-0.016
	[-1:1]	0.000	-0.005	0.004	0.006	-0.000	-0.015
	[-1:2]	-0.001	-0.007	0.013	0.001	0.001	-0.003
	[-1:3]	0.001	-0.005	0.010	0.002	0.001	-0.025
	[-1:4]	0.001	-0.004	0.004	0.005	0.000	-0.018
	[-1:5]	0.001	-0.003	0.008	0.001	-0.000	-0.035
sentyment negatywny	[-1:0]	0.001	-0.005	0.013	0.002	-0.003	-0.016
	[-1:1]	0.000	-0.011	0.007	0.005	-0.002	-0.015
	[-1:2]	0.004	-0.021	0.014	0.010	0.006	-0.003
	[-1:3]	0.004	-0.019	0.016	0.011	0.000	-0.025
	[-1:4]	0.001	-0.020	0.020	0.007	0.006	-0.018
	[-1:5]	0.001	-0.004	0.010	0.007	0.004	-0.035

Źródło: opracowanie własne.

Należy zaznaczyć, że w próbie badawczej znalazły się również takie spółki, które – w reakcji zarówno na sentyment pozytywny, jak i negatywny – nie wykazały istotnych statystycznie stóp zwrotu w całym oknie zdarzenia. Takie wyniki dotyczyły 17% spółek i były to: AssecoPol (ACP1, ACP2), Allegro, GTC, Synthos oraz Santander Bank Polska.

Warto jednak zauważyć, że w przypadku banku Santander do analizy zakwalifikowano jedynie 10 zdarzeń (5 pozytywnych, 5 negatywnych), co w kontekście takich spółek jak JSW lub KGH stanowiło relatywnie niewielką próbę badawczą.

Uzyskane rezultaty wskazują, że reakcja na pojawienie się artykułów prasowych o skrajnie negatywnym sentymencie, znajduje swoje odzwierciedlenie w krótszych oknach zdarzeń, tj. $[-1:0]$ oraz $[-1:1]$, niż ma to miejsce w przypadku sentymentu skrajnie pozytywnego. W tym drugim przypadku reakcja rozciąga się do czwartego, a nawet piątego dnia po zdarzeniu. Wyniki te nie uprawniają więc do odrzucenia drugiej hipotezy pomocniczej, mówiącej o tym, że sposób reakcji inwestorów na sentyment cechuje się asymetrią, w zależności od tego, czy sentyment ten zdefiniowany został jako pozytywny, czy negatywny.

6.2.4. Wyniki trzeciego etapu badań

Dla zobrazowania jakości dopasowania modelu zastosowano miarę *pseudo* – R^2 , której wartości zamykają się w przedziale od 0 do 1. W ujęciu średnim wynik ten znalazł się zaledwie na poziomie 0,02. Maksymalna wartość, którą udało się w tym kontekście osiągnąć, wyniosła 0,31 i dotyczyła spółki Eurocash. W ogólnym ujęciu uzyskane wyniki wskazują na to, że model – w kontekście wszystkich badanych spółek – w niskim stopniu odpowiada rzeczywistym wynikom stóp zwrotu.

Jednocześnie spośród 34 analizowanych spółek, w przypadku czterech (12% próby badawczej) wartość *pseudo* – R^2 określić można jako istotne statystycznie na poziomie $\alpha \leq 0,01$. Są to spółki CCC, LPP, Lubelski Węgiel Bogdanka oraz PKN Orlen. Dla dwóch spółek (6% próby badawczej) *pseudo* – R^2 wskazało na wyniki istotne statystycznie na poziomie $\alpha \leq 0,05$. Były to spółki: Eurocash i PKO BP. W przypadku czterech kolejnych spółek wynik ten okazał się istotny statystycznie na poziomie $\alpha \leq 0,1$. Mowa o spółkach: BreBank, CD Projekt, KGHM oraz Orange Polska. Oznacza to jednak, że w przypadku 71% analizowanych spółek nie stwierdzono istotności statystycznej uzyskanych rezultatów.

Tabela 17. Wyniki estymacji modelu logitowego

spółka	pseudo R-squ	coeff.	p-value	próba
ACP1*	0.001333	-0.0879	0.3658	445
ALE	0.004981	-0.1682	0.4888	73
ALR	0.0002108	0.0345	0.7255	422
BHW	0.01453	0.3233	0.4851	26
BRE	0.03048	0.4278	0.09016	68
BZW	0.009258	0.2773	0.2982	85
CCC	0.01978	0.3361	0.01038	240
CDR	0.02819	0.4611	0.1032	72
CPS	0.0008785	-0.0714	0.4730	423
DNP	0.0006025	0.1003	0.8484	44
ENA	0.006589	-0.1947	0.2070	176
ENG	0.003203	0.1353	0.2675	279
EUR	0.3113	0.01716	0.03482	188
GTC	0.0003946	0.0477	0.8747	46
JSW1	0.006822	0.2315	0.2059	172
JSW2	0.002286	0.1139	0.3485	280
KER	0.01516	0.3233	0.3913	35
KGH	0.002889	-0.1268	0.06684	841
LPP	0.01733	0.3202	0.008722	289
LTS	0.00004059	0.0159	0.9164	196
LWB	0.02727	0.4265	0.005163	208
MBK	0.001295	-0.0857	0.3933	408
OPL	0.001770	0.0995	0.1395	890
PEO	0.002198	0.1133	0.4058	228
PGE	0.000002514	-0.0038	0.9539	959
PGN	0.00000009816	-0.0008	0.9891	1364
PKN	0.002493	-0.1187	0.009506	1949
PKO	0.005175	0.1706	0.02029	752
PLY	0.005536	0.1766	0.3584	110
PZU	0.005342	0.1736	0.2202	203
SNS	0.01325	0.2877	0.2551	71
SPL	0.003531	-0.1473	0.4319	128
TPE	0.004953	-0.1703	0.2106	236
TPS	0.00001878	0.0116	0.9687	60

* Niewielka liczba artykułów w przypadku spółki ACP2 (17) uniemożliwiła przeprowadzenie estymacji.

Źródło: opracowanie własne.

Prezentowane wyniki wskazują również na niską korelację, zachodzącą między skrajnym sentymentem tekstowym a dziennymi stopami zwrotu. W ogólnym ujęciu znajdowała się ona w przedziale od -0,19 (spółka Enea) do 0,46 (spółka CD Projekt). Zauważyć można jednak, że – co do zasady – wyniki charakteryzują się korelacją dodatnią. Taki rezultat zaobserwować można w przypadku 23 podmiotów (68% próby badawczej). Dla pozostałych 11 spółek korelacja między sentymentem artykułów a stopami zwrotu okazała się ujemna, ale jednocześnie nieznacznie słabsza niż przy spółkach z korelacją dodatnią (średnio: -0,11 oraz 0,2). Wyniki estymacji modelu logitowego zaprezentowane zostały w tabeli numer 17.

Charakterystyka rezultatów uzyskanych podczas estymacji modelu logitowego, w kontekście jakości jego dopasowania, istotności statystycznej oraz siły korelacji, znajdują swoje odzwierciedlenie w zdolności modelu do prawidłowego klasyfikowania stóp zwrotu jako ujemne lub dodatnie. Po weryfikacji pod tym kątem całej próby badawczej należy stwierdzić, że przy użyciu sentymentu tekstowego jako zmiennej objaśniającej, umiejętność modelu do poprawnego przyporządkowania zmiennych objaśnianych do jednej z dwóch klas, jest niska. W żadnym z 34 podmiotów model nie przypisał ujemnych stóp zwrotu do prawidłowej kategorii. Wszystkie stopy zwrotu, niezależnie od ich rzeczywistych wartości, zostały zdefiniowane jako dodatnie. Na uwagę zasługuje jednak fakt, że model w 100% poprawnie klasyfikował dodatnie stopy zwrotu jako dodatnie. Nie mniej, w kontekście potencjalnej predykcji, wyniki te czynią go nieprzydatnym. W tabeli numer 18 zaprezentowane zostały sumaryczne wyniki klasyfikacji wszystkich analizowanych spółek.

Tabela 18. *Wyniki klasyfikacji dziennych stóp zwrotu dla spółek w próbie badawczej*

dane rzeczywiste		stopy zwrotu - predykcja	
		ujemne	dodatnie
stopy zwrotu	ujemne	0	6225
	dodatnie	0	5741

Źródło: opracowanie własne.

Z uwagi na niezdolność modelu do przyporządkowania zmiennych objaśnianych do klasy ujemnych stóp zwrotu, współczynnik dokładności w każdym analizowanym

przypadku przyjmował wartości na poziomie 0,0. Z drugiej strony współczynniki czułości oraz swoistości, znajdujące się na poziomie 1,0, wskazują na to, że model prawidłowo identyfikuje stopy zwrotu, które w rzeczywistości są dodatnie jako takie. Podobnie jak w przypadku wyników współczynnika dokładności, wynik ten dotyczy wszystkich badanych spółek. Jedynym współczynnikiem, którego wartości w ramach próby badawczej ulegają zróżnicowaniu, jest wskaźnik precyzji. Za jego pomocą stwierdzić można, jaka część stóp zwrotu, które w wyniku predykcji zaklasyfikowane zostały jako dodatnie, w rzeczywistości takimi są. Warto zwrócić uwagę, że w przypadku kilku spółek wyniki przekroczyły 50%. Mowa tu o: CD Projekt (64%), GTC (57%), Bank Zachodni (55%), Dino Polska (55%), Asseco Poland (ACP1, 55%), Play Communications (52%), Bre Bank (51%), oraz Cyfrowy Polsat (51%). Teoretycznie więc w ich przypadku model wykazywał zdolność do prawidłowej klasyfikacji na poziomie wyższym niż wybór losowy. Po pierwsze jednak sytuacja ta dotyczyła tylko 8 z 34 badanych spółek. Po drugie nie można uznać, że dzięki temu model określić można jako potencjalnie użyteczny dla inwestorów giełdowych. Wyniki współczynników klasyfikacji dla każdej z badanych spółek zaprezentowane zostały w tabeli 19.

Ocena rezultatów – uzyskanych podczas estymacji regresji logitowej – w postaci: stopnia dopasowania modelu, jego statystycznej istotności, korelacji między zmiennymi objaśnianymi i objaśniającymi oraz współczynników klasyfikacji, wskazuje na istnienie podstaw do odrzucenia trzeciej hipotezy pomocniczej. Nie znaleziono bowiem argumentów za tym, że sentyment, zawarty w anglojęzycznych artykułach prasowych, może stanowić zmienną, determinującą wynik dziennych stóp zwrotu spółek, notowanych na Giełdzie Papierów Wartościowych w Warszawie w latach 2013 - 2022.

Tabela 19. Wyniki współczynników klasyfikacji

spółka	współczynniki klasyfikacji			
	precyzja	czułość	swoistość	dokładność
ACP1	0,54	1,00	1,00	0,00
ALE	0,37	1,00	1,00	0,00
ALR	0,49	1,00	1,00	0,00
BHW	0,35	1,00	1,00	0,00
BRE	0,51	1,00	1,00	0,00
BZW	0,55	1,00	1,00	0,00
CCC	0,48	1,00	1,00	0,00
CDR	0,64	1,00	1,00	0,00
CPS	0,51	1,00	1,00	0,00
DNP	0,55	1,00	1,00	0,00
ENA	0,44	1,00	1,00	0,00
ENG	0,45	1,00	1,00	0,00
EUR	0,46	1,00	1,00	0,00
GTC	0,57	1,00	1,00	0,00
JSW1	0,42	1,00	1,00	0,00
JSW2	0,44	1,00	1,00	0,00
KER	0,49	1,00	1,00	0,00
KGH	0,47	1,00	1,00	0,00
LPP	0,44	1,00	1,00	0,00
LTS	0,48	1,00	1,00	0,00
LWB	0,46	1,00	1,00	0,00
MBK	0,46	1,00	1,00	0,00
OPL	0,50	1,00	1,00	0,00
PEO	0,46	1,00	1,00	0,00
PGE	0,48	1,00	1,00	0,00
PGN	0,50	1,00	1,00	0,00
PKN	0,48	1,00	1,00	0,00
PKO	0,48	1,00	1,00	0,00
PLY	0,52	1,00	1,00	0,00
PZU	0,49	1,00	1,00	0,00
SNS	0,45	1,00	1,00	0,00
SPL	0,43	1,00	1,00	0,00
TPE	0,39	1,00	1,00	0,00
TPS	0,43	1,00	1,00	0,00

Źródło: opracowanie własne.

6.3. Wnioski z badań

6.3.1. Podsumowanie uzyskanych wyników

Zaprezentowane badania, przeprowadzone zostały na podstawie próby badawczej, skonstruowanej ze spółek, wchodzących w skład indeksu WIG20, w okresie 01.01.2013 - 12.12.2022. Do próby zakwalifikowano 30 spółek, przy czym spółki, które w analizowanym okresie dokonały *rebrandingu* lub okres ich notowania w indeksie charakteryzował się nieciągłością, traktowane były jako oddzielne podmioty, co rozszerzyło próbę do 35 podmiotów. Dla każdego z nich zgromadzone zostały anglojęzyczne artykuły prasowe, których datowanie odpowiadało okresowi notowania spółki we wskazanym indeksie. W ten sposób stworzono bazę, składającą się z ponad 285 tysięcy publikacji prasowych. Analizie podlegały jednak tylko te artykuły, w których tytułach pojawiły się nazwy analizowanych spółek. Po wyselekcjonowaniu artykułów o takiej charakterystyce liczba publikacji ograniczona została do 56 953.

Badania składały się z trzech etapów. Celem etapu pierwszego była weryfikacja pierwszej hipotezy pomocniczej:

H1: Anglojęzyczne artykuły prasowe, dotyczące spółek notowanych w indeksie WIG20 w okresie od 01.01.2013 do 31.12.2022, charakteryzują się przewagą sentymentu pozytywnego względem sentymentu negatywnego.

Proces weryfikacji hipotezy rozpoczął się od procedury ekstrakcji sentymentu, zawartego w zgromadzonych publikacjach. Do tego celu wykorzystany został słownik *Loughran-McDonald Master Dictionary*. Dzięki temu do każdego z artykułów przypisany został sentyment, którego wartości znajdowały się w przedziale od -1 (dla artykułów negatywnych) do 1 (dla artykułów pozytywnych).

Analiza struktury poszczególnych klas sentymentów dla wszystkich spółek pozwoliła stwierdzić, że dla 58 % podmiotów, dominującym sentymentem był sentyment neutralny. Dla 27 % spółek był to sentyment pozytywny. Sentyment negatywny natomiast dominował wśród 15 % badanych spółek. Wyłączenie z analizy artykułów, charakteryzujących się sentymentem neutralnym, pozwoliło wykazać, że dla 58% spółek, liczba artykułów cechujących się pozytywnym sentymentem, stanowi więcej niż połowę ogółu publikacji.

Można stwierdzić, że uzyskane wyniki zgodne są z wywodem, zaprezentowanym w rozdziale trzecim. Jednym z głównych – omawianych w nim – zagadnień, był problem kadrowania. Dokonany przegląd literatury przedmiotu pozwolił wskazać istotne czynniki, mogące mieć wpływ na obawy dziennikarzy finansowych, co do kadrowania zdarzeń gospodarczych jako negatywne. Część z nich opiera się na prostym założeniu, które zaproponowali Elitan Goldman, Jordan Martel i Jan Schneemeier. Stwierdzili oni, że spółki dążą do promowania pozytywnych i unikania negatywnych komunikatów. W związku z tym, zarządy firm stosują praktyki, które badacze nazywają "zaciemnianiem" (ang. *obfuscation*) niekorzystnych informacji. Dzieje się to na przykład poprzez używanie niejasnych sformułowań w raportach lub nadmierne podkreślanie pozytywnych aspektów działalności organizacji⁶⁵⁸. W tym miejscu między dziennikarzami a spółkami pojawia się konflikt interesów. Ponieważ efektywność pracy analitycznej zależy od możliwości uzyskania dostępu do danych, to – jak zaznacza Michael Borden – jeśli stanowiska dziennikarzy są sprzeczne z interesami konkretnej firmy, ryzykują oni pogorszeniem relacji z jej przedstawicielami i ewentualnym ograniczeniem dostępu do danych źródłowych⁶⁵⁹.

Innym czynnikiem, mogącym zniechęcać dziennikarzy do publikowania negatywnych informacji jest zjawisko, które Mark Boukes i Rens Vliegenthart nazywają „ramami konsekwencji ekonomicznych” (ang. *economic consequences-frame*)⁶⁶⁰. Alyt Damstra i Knut De Swert mówią wręcz o obawach, związanych z tak zwaną „samospełniającą się przepowiednią”. Autorzy ci wskazują na ogólnospołeczną świadomość roli, jaką na rynku finansowym odgrywają emocje. Podają w tym kontekście przykład liderów politycznych, którzy z reguły zachęcają ekspertów ekonomicznych do „dyskusji”, próbując tym samym łagodzić ton ich wypowiedzi⁶⁶¹.

⁶⁵⁸ Zob. Goldman, E., Martel, J., Schneemeier, J., A Theory of Financial Media, „Journal of Financial Economics”, 2022, nr 145/1, s. 239–258.

⁶⁵⁹ Borden, M. J., The Role of Financial Reporting in Corporate Governance. „Fordham Journal of Corporate & Financial Law”, 2007, nr XII, s. 367.

⁶⁶⁰ Boukes, M., Vliegenthart, R., A general pattern in the construction of economic newsworthiness? Analyzing news factors in popular, quality, regional, and financial newspapers, „Journalism”, 2020, nr 21/2, s. 289.

⁶⁶¹ Damstra, A., De Swert, K., The making of economic news: Dutch economic journalists contextualizing their work, „Journalism”, 2020, nr 22/12, s. 3090.

Mark Boukes i Rens Vliegenthart zwracają również uwagę na aspekt profesjonalizmu. Przeprowadzona przez nich analiza wykazała, że wartość informacyjna, którą nazwali „negatywnością”, w największym stopniu przejawia się w tak zwanych czasopismach popularnych. W najmniejszym stopniu wartość ta widoczna była w gazetach stricte finansowych. Zdaje się więc, że z perspektywy zawodu dziennikarskiego, nadmierny pesymizm może być postrzegany jako czynnik obniżający jakość informacji⁶⁶².

Scharakteryzowane czynniki stanowią przesłankę, by sądzić, że – co do zasady – w artykułach prasowych, opisujących działania spółek giełdowych, dominował będzie sentyment pozytywny. Konkluzja ta znalazła swoje odzwierciedlenie w wynikach badań, zaprezentowanych w niniejszej dysertacji. Z uwagi na powyższe, autorka nie znalazła podstaw do odrzucenia pierwszej hipotezy pomocniczej.

Kolejny etap badania służył weryfikacji drugiej hipotezy pomocniczej:

H2: Występuje asymetria w sposobie reakcji rynku na pozytywny i negatywny sentyment, zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20.

W tym celu zastosowana została metoda analizy zdarzeń, gdzie dzień zdarzenia zdefiniowany został jako dzień publikacji anglojęzycznego artykułu prasowego, charakteryzującego się skrajnym sentymentem pozytywnym lub negatywnym. Liczba artykułów, które spełniały to kryterium, wyniosła 5 681, przy czym 54% stanowiły artykuły, cechujące się sentymentem pozytywnym, a 46% - cechujące się sentymentem negatywnym. Reakcja rynku opisywana była przez ponadnormatywne średnie skumulowane stopy zwroty (CAAR), przy czym analizy dokonywano oddzielnie dla artykułów, cechujących się sentymentem pozytywnym i negatywnym.

Uzyskane wyniki wykazały, że w reakcji zarówno na zdarzenie zdefiniowane jako pozytywne, jak i negatywne, średnie skumulowane ponadnormatywne stopy zwrotu większości spółek, cechowały się statystyczną istotnością. Można było zaobserwować, że dla wartości $p\text{-value} < 0.01$ reakcja rynku na informacje, charakteryzujące się negatywnym sentymentem, widoczna była przede wszystkim w oknach zdarzenia $[-1:0]$

⁶⁶² Boukes, M., Vliegenthart, R., A General Pattern..., op.cit., s. 292.

oraz [-1:1]. W przypadku publikacji, cechujących się sentymentem pozytywnym, wśród największej liczby spółek, reakcja widoczna była w oknach [-1:4] oraz [-1:5]. Podobnie, dla wartości $p\text{-value} < 0.1$ – w przypadku sentymentu negatywnego – największa liczba spółek wykazała się statystycznie istotnymi średnimi skumulowanymi ponadnormatywnymi stopami zwrotu w oknach zdarzenia [-1:0] oraz [-1:1]. W przypadku sentymentu pozytywnego było to [-1:3]. Zaprezentowane wyniki uprawniają więc do stwierdzenia, że sposoby dyskontowania przez inwestorów rynkowych informacji, charakteryzujących się sentymentem pozytywnym oraz negatywnym, nie są takie same. Tym samym uzyskane rezultaty nie uprawniają do odrzucenia drugiej hipotezy pomocniczej.

Przeprowadzone badanie wpisuje się w szereg analiz, które dotąd wykonywane były na rynkach akcji. Uzyskiwane przez wielu badaczy rezultaty, potwierdzały odmienny sposób dyskontowania wiadomości tekstowych, w zależności od przypisywanego im sentymentu. Przykładowo Steven L. Heston i Nitish R. Sinha wykorzystując zbiór 900 tysięcy wiadomości z rynku amerykańskiego, wykazali, że pozytywne i negatywne wiadomości są uwzględniane w cenach akcji z różną szybkością. Wskazali również na relatywną krótkotrwałość tego efektu. W przypadku ich badania widoczny był on przez okres od jednego do dwóch dni. Co jednak ciekawe doszli do wniosku, że to reakcja na pozytywne wiadomości wywołuje gwałtowny zmiany w wycenie akcji. Reakcje na wiadomości o nacechowaniu negatywnym charakteryzowały się natomiast opóźnieniem⁶⁶³.

Być może jedną z przyczyn uzyskania odmiennych rezultatów, są wartości sentymentu, które brane były pod uwagę podczas przeprowadzania badania. Na problem ten uwagę zwrócili – wspominani w rozdziale piątym – David Ardia, Keven Bluteau oraz Kris Boudta. Poddali oni analizie spółki niefinansowe, notowane w indeksie S&P500. Dostrzegli istnienie pozytywnej zależności między skumulowaną ponadnormatywną stopą zwrotu, a tak zwanym skumulowanym ponadnormatywnym tonem wiadomości (RCAT). Badanie to koresponduje z wynikami, uzyskanymi w niniejszej dysertacji z dwóch powodów. Po pierwsze, w obu badaniach pod uwagę brane były skrajne

⁶⁶³ Zob. Heston, S. L., Sinha, N. R., News versus Sentiment: Predicting Stock Returns from News Stories, "Financial Analysts Journal", 2018, nr 73/3, s. 67-83.

wartości sentymentu. Po drugie reakcja inwestorów okazywała się silniejsza, jeśli dyskontowaniu podlegały wiadomości o negatywnym wydźwięku⁶⁶⁴. Na analogiczne właściwości mechanizmów, związanych z tempem dyskontowania informacji – w zależności od rodzaju sentymentu – uwagę zwrócili Aylt Damstra oraz Knut De Stewart, analizując rynek holenderski. Scharakteryzowali oni negatywne zmiany, mające miejsce na rynku finansowym, jako co do zasady – gwałtowne. Napływ na rynek pozytywnych informacji, związany był ich zdaniem z trendami o charakterze długofalowymi⁶⁶⁵.

Warto zauważyć, że omawiana asymetria widoczna jest nie tylko w ramach badań, których podstawą były artykuły prasowe. Przykładowo Allen H. Huang, Amy Y. Zang i Rong Zheng, zauważyli podobny wzorec zachowań inwestorów, przeprowadzając analizę raportów finansowych oraz zmian indeksu S&P500. Pozytywnie zweryfikowali oni hipotezę, mówiącą o tym, że reakcja inwestorów jest silniejsza w przypadku negatywnego wydźwięku tekstu. Zasugerowali oni również, że potencjalną przyczyną takiego stanu rzeczy, jest asymetria, która objawia się już na etapie upubliczniania informacji przez zarządy przedsiębiorstw. Ponieważ chętniej i częściej przekazują oni wiadomości korzystne z punktu widzenia spółki, to pozytywny sentyment nie stanowi dla inwestorów zaskoczenia. Odwrotnie sytuacja wygląda w przypadku sentymentu negatywnego⁶⁶⁶.

Dysproporcję w reakcji na wiadomości o odmiennym wydźwięku zauważyli również badacze, zajmujący się innymi rynkami niż rynek akcji. Przykładem są tutaj rezultaty uzyskane przez Lee Smalesa na rynku instrumentów pochodnych złota. Podobnie jak w przypadku badania, które przeprowadzili Allen H. Huang, Amy Y. Zang i Rong Zheng, pozytywnie zweryfikował on hipotezę o silniejszej reakcji inwestorów na wiadomości negatywne⁶⁶⁷.

⁶⁶⁴ Zob. Ardia, D., Bluteau, K., Boudt, K., *Media abnormal tone, earnings announcements, and the stock market*, "Journal of Financial Markets", 2022, nr 61, s. 1–19.

⁶⁶⁵ Zob. Damstra, A., De Swert, K., *The making of economic news: Dutch economic journalists contextualizing their work*, „Journalism”, 2020, nr 22/12, s. 3091.

⁶⁶⁶ Zob. Huang, A. H., Zang, A. Y., Zheng, R., *Evidence on the Information Content of Text in Analyst Reports*, "The Accounting Review", 2014, nr 89/6, s. 2151-2180.

⁶⁶⁷ Zob. Smales, L., *News sentiment in the gold futures market*, "Journal of Banking & Finance", 2014, nr 49, 275-286.

Należy również zwrócić uwagę, że uzyskane wyniki wpisują się w – prezentowaną w pierwszym oraz drugim rozdziale dysertacji – dyskusję, związaną z dyskontowaniem przez inwestorów informacji w świetle Hipotezy Efektywności Informacyjnej oraz paradygmatu finansów behawioralnych. Stanowią one przesłankę do stwierdzenia, że nie wszystkie informacje, które pojawiają się na rynku, dyskontowane są przez inwestorów w sposób natychmiastowy, jak również, że owa asymetria, związana jest z czynnikami pozafinansowymi. Odmienna długość reakcji na informacje ze względu na sentyment wspiera perspektywę behawioralną. Może mieć to bowiem związek z koncepcjami punktu referencyjnego oraz wag decyzyjnych. Efektem tych mechanizmów jest odczytywanie przez inwestorów wartości w sposób subiektywny, podczas gdy, zgodnie z paradygmatem neoklasycznym, powinny być one traktowane w sposób absolutny⁶⁶⁸. Zaprezentowane wyniki mogą wskazywać na to, że większą wagę przypisują zdarzeniom negatywnym niż pozytywnym, ponieważ zgodnie z twierdzeniem Daniela Kahnemana „strata boli bardziej, niż cieszy zysk”⁶⁶⁹. Należy jednak zauważyć, że większość spółek wykazywała istotność statystyczną CAAR w oknie zdarzenia [-1:0]. Oznacza to, że napływające na rynek informacje, były już zawarte w cenie, lub ich dyskontowanie następowało w relatywnie krótkim czasie. Fakt ten nie daje podstaw do odrzucenia Hipotezy Efektywności Informacyjnej w formie pólśilnej, jako adekwatnego opisu sytuacji spółek, znajdujących się w próbie badawczej⁶⁷⁰.

Celem ostatniego etapu badania, była weryfikacja trzeciej hipotezy pomocniczej:

H3: *Sentyment zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20, stanowi determinantę reakcji inwestorów giełdowych.*

Dla realizacji powyższego celu zbudowany został model logitowy. W jego ramach za zmienną objaśniającą przyjęto – obliczony w pierwszym etapie – sentyment tekstowy anglojęzycznych artykułów prasowych. Podobnie jak miało to miejsce w przypadku analizy zdarzeń, uwzględniono jedynie skrajne wyniki wskaźnika sentymentu, tj. takie, które mieściły się w przedziale od 0,9 do 1 (dla artykułów o sentymencie pozytywnym)

⁶⁶⁸ Zob. Zielonka, P., *Giełda i psychologia. Behawioralne aspekty inwestowania na rynku papierów wartościowych*, CeDeWu, Warszawa 2017, s. 79 - 87.

⁶⁶⁹ Kahneman, D., Tversky, A., *Prospect Theory...*, op.cit., s. 288.

⁶⁷⁰ Zob. Fama, E. F., *Efficient Capital Markets: A Review of Theory and Empirical Work*, „The Journal of Finance”, 1970, 25/2, 28–30

oraz od -1 do -0,9 (dla artykułów o sentymencie negatywnym). Rezultaty opisane zostały przez miary: pseudo-R2, p-value oraz współczynniki klasyfikacyjne.

Dla 71% podmiotów, znajdujących się w próbie badawczej, estymacja modelu nie pozwoliła uzyskać takich wartości pseudo-R2, dla których p-value wykazywałoby się statystyczną istotnością. Analiza współczynników klasyfikacji, takich jak: precyzja, czułość, swoistość oraz dokładność, wskazała na niezdolność modelu do prawidłowej predykcji ujemnych stóp zwrotu. Jednocześnie model osiągnął 100% poprawności, jeśli chodzi o przyporządkowanie dodatnich stóp zwrotu do właściwej kategorii. Mimo to, w kontekście prognozowania, nie można uznać jego wyników za satysfakcjonujące. Było to podstawą do odrzucenia trzeciej hipotezy pomocniczej.

Możliwość wykorzystania sentymentu tekstowego do predykcji cen akcji, potwierdzona została na rynkach niemal wszystkich kontynentów. Zaprezentowane zostało to w rozdziale piątym niniejszej dysertacji. Przykładem mogą być tutaj badania, które zaproponowali Ayman E. Khedr, S. E. Salama i Nagwa Yaseen, analizujący indeks NASDAQ. W modelu, w którym zmienną objaśniającą były wyłącznie wiadomości finansowe, osiągnęli oni wartości prognostyczne przekraczające 70%⁶⁷¹. Inny amerykański indeks (Dow Jones) przeanalizowali pod tym kątem: David E. Allen, Michael McAleer i Abhay K. Singh. W tym przypadku również udało się wykazać, że sentyment tekstowy ma znaczący, statystycznie istotny wpływ na stopy zwrotu spółek, wchodzących w skład indeksu⁶⁷².

Zdolność do wpływania przez media na dynamikę cen akcji, potwierdzana była również na rynku europejskim. Jan Kleinnijenhuis, Friederike Schultz, Dirk Oegema oraz Wouter van Atteveldt analizowali zachowanie inwestorów holenderskich podczas światowego kryzysu finansowego. Wskazali w tym kontekście na podwójną rolę mediów. Z jednej strony przekazywały one informacje o sytuacji ekonomicznej. Z drugiej – pogłębiały niekorzystne zmiany, które znajdowały swoje odzwierciedlenie w indeksie

⁶⁷¹ Zob. Khedr, A. E., Salama, S. E., Yaseen, N., *Predicting stock market behavior using data mining technique and news sentiment analysis*, „International Journal of Intelligent Systems and Applications”, 2017, nr 7, s. 22–30.

⁶⁷² Zob. Allen, D. E., McAleer, M., Singh, A. K., *Daily market news sentiment and stock prices*, „Applied Economics”, 2019, nr 51/30, s. 3212–3235.

AEX⁶⁷³. Vegard H. Larsen oraz Leif Anders Thorsrud zaprezentowali z kolei wyniki badań przeprowadzonych na rynku norweskim. W tym przypadku autorom udało się wykazać, że strategie inwestycyjne, których podstawą były wiadomości tekstowe, mogą przynieść ponadprzeciętne stopy zwrotu z akcji⁶⁷⁴.

Mimo że wyniki, uzyskane na podstawie trzeciego etapu, nie potwierdzają efektów prac wielu autorów, to należy zauważyć, że przeprowadzony przez autorkę przegląd literatury, pozwolił wskazać publikacje, w ramach których nie udało się zdefiniować sentymentu jako determinanty reakcji rynku. Przykładem mogą być badania zaprezentowane przez Adama Atkinsa, Mahesana Niranjana i Enrico Gerdinga, przeprowadzone na rynku amerykańskim. Uzyskane przez autorów rezultaty, wykazały niezdolność modelu klasyfikacyjnego do prawidłowej predykcji stóp zwrotu, na poziomie wyższym niż losowy⁶⁷⁵.

W tym kontekście, do interesujących wniosków doszli badacze holenderscy: Nadine Strauß, Rens Vliegthart oraz Piet Verhoeven. Poddali oni analizie spółki wchodzące w skład indeksu AEX w okresie 2002 – 2013. Zauważyli, że to raczej media odzwierciedlają dynamikę rynku, niż rynek reaguje na informacje medialne⁶⁷⁶. Wyniki te znalazły swoje potwierdzenie również na rynku brazylijskim. Badacze: Tanya Araújo, Samuel Eleutério oraz Francisco Louçã, sprawdzając możliwość wyjaśnienia dynamiki rynku za pomocą nastrojów wyrażanych w specjalistycznych mediach, analizowali spółki znajdujące się w indeksie⁶⁷⁷. Udowodnili oni jednak, że to przede wszystkim zmiany cen instrumentów wpływają na sentyment, a nie odwrotnie⁶⁷⁸.

Wskazać można również badania, których wyniki wprawdzie potwierdzają możliwość wykorzystania sentymentu w celach predykcyjnych, ale jedynie

⁶⁷³ *Amsterdam EXchange Market*.

⁶⁷⁴ Zob. Larsen, V. H., Thorsrud, L. A., *Asset returns, news topics, and media effects*, „CAMP Working Paper Series”, 2017, nr 5, s. 1–48.

⁶⁷⁵ Zob. Atkins, A., Niranjana, M., Gerding, E., *Financial news predicts stock market volatility better than close price*, „The Journal of Finance and Data Science”, 2018, nr 4, s. 120–137.

⁶⁷⁶ Strauß, N., Vliegthart, R., Verhoeven, P., *Lagging behind? Emotions in newspaper articles and stock market prices in the Netherlands*, „Public Relations Review”, 2016, nr 42, s. 548–555.

⁶⁷⁷ Zob. Araújo, T., Eleutério, S., Louçã, F., *Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood*, „Physica A.”, 2018, nr 505, s. 1139–1149.

⁶⁷⁸ ⁶⁷⁸ Zob. Araújo, T., Eleutério, S., Louçã, F., *Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood*, „Physica A.”, 2018, nr 505, s. 1139–1149.

w określonych interwałach. Na przykład włoscy badacze: Francesco Colasanto, Luca Grilli, Domenico Santoro i Giovanni Villani, doszli do wniosku, że przydatność tego typu zmiennej występuje jedynie w interwałach godzinowych⁶⁷⁹. Wydaje się więc, że przeprowadzane dotychczas analizy, związane z możliwością wykorzystania sentymentu tekstowego, do predykcji stóp zwrotu z akcji, nie pozwalają na wyciągnięcie jednoznacznych wniosków. Rezultaty, uzyskane przez autorkę niniejszej dysertacji, wskazują jednak na niezdolność modelu – zbudowanego na podstawie tej zmiennej – do prognozowania zmian w cenach akcji dla badanej próby w warunkach GPW w Warszawie.

Podsumowując, zaprezentowane wyniki badań, wskazują na odmienny sposób dyskontowania sentymentu, zawartego w artykułach prasowych. W przypadku sentymentu negatywnego reakcje, w postaci istotnych statystycznie średnich skumulowanych ponadnormatywnych stóp zwrotu, częściej charakteryzowały się krótkotrwałością i mieściły się w oknie zdarzenia od dnia przed publikacją artykułu do pierwszego dnia po jego publikacji. Reakcja na publikację, charakteryzującą się sentymentem pozytywnym, można było zaobserwować w oknach zdarzenia od dnia poprzedzającego publikację, do trzech, czterech lub pięciu dni po publikacji. Ponadto kierunek stóp zwrotu w reakcji na sentyment negatywny u dominującej części spółek był zgodny z oczekiwaniami. Oznaczało to, że ponadnormatywne stopy zwrotu przyjmowały wartości ujemne. Reakcja na sentyment pozytywny nie pozwala na wyciągnięcie tak jednoznacznego wniosku, ponieważ w przypadku części spółek ponadnormatywne stopy zwrotu również przyjęły wartości ujemne. Uprawnia to do stwierdzenia, że informacje o nacechowaniu pozytywnym dyskontowane są przez inwestorów w odmienny sposób niż informacje o nacechowaniu negatywnym. W związku z powyższym nie znaleziono podstaw do odrzucenia hipotezy głównej, mówiącej o tym, że rodzaj sentymentu, zawarty w danych tekstowych, jest odzwierciedlony w reakcji rynku.

6.3.2. Rekomendacje i dalsze kierunki badań

Z uwagi na potencjał aplikacyjny wykorzystywania analizy sentymentu do objaśnienia sposobów reagowania inwestorów giełdowych na napływające na rynek

⁶⁷⁹ Zob. Colasanto, F., Grilli, L., Santoro, D., Villani, G., *ALBERTino for stock price prediction: a Gibbs sampling approach*, „Information Sciences”, 2022, nr 597, s. 341–357.

informacje, pożądane jest kontynuowanie badań we wskazanym obszarze. Zasadne wydaje się zweryfikowanie – postawionych przez autorkę – hipotez przy użyciu innych narzędzi i procedur. Chodzi tu z jednej strony o zastosowanie alternatywnych słowników, umożliwiających ekstrakcję sentymentu z danych tekstowych. Należy jednak zwrócić uwagę na dynamiczny rozwój koncepcji, związanych z algorytmami uczenia maszynowego oraz sztucznej inteligencji, dzięki którym możliwa jest zarówno praca z danymi tekstowymi, jak i eksploracja sentymentu. Z drugiej strony więc istotne jest również zwrócenie się w stronę tego typu narzędzi.

Wykorzystane w badaniu źródło danych tekstowych, jakimi są artykuły prasowe, jest tylko jedną z możliwości. Przytoczone w rozdziale piątym badania wskazują na powszechność alternatywnych rozwiązań, takich jak media społecznościowe, czy sprawozdania finansowe. Poddanie analizie dynamiki stóp zwrotu na podstawie sentymentu, uzyskanego z innych źródeł danych, znajduje swoje uzasadnienie w światowej literaturze. Zaprezentowane wyniki badań uzyskane były na podstawie próby badawczej, w której znajdowały się wyłącznie spółki notowane w indeksie WIG20, dlatego warto byłoby również przeanalizować relację, zachodzącą między sentymentem tekstowym a stopami zwrotu spółek o niższej kapitalizacji.

Biorąc pod uwagę dotychczasowy dorobek badawczy w prezentowanym obszarze, istotne jest również podejmowanie dalszych prób, pozwalających zweryfikować możliwość zastosowania sentymentu tekstowego do predykcji stóp zwrotu akcji. Warto rozważyć tutaj dwie kwestie. Pierwszą jest zastosowanie innych modeli ekonometrycznych, które nie zamykają zmiennych objaśnianych w ramy dwóch klas, tak jak ma to miejsce w przypadku regresji logitowej. Druga kwestia natomiast ma związek z definiowaniem zmiennej objaśniającej. Być może wykorzystanie jedynie skrajnych sentymentów, jest koncepcją nadto rygorystyczną i pomocne byłoby zastosowanie całego przedziału wyników pozytywnych i negatywnych.

Autorka niniejszej pracy rekomenduje również szersze uwzględnianie paradygmatu behawioralnego, zarówno w kontekście dalszych badań nad zachowaniami inwestorów giełdowych, jak i w kontekście aktywności praktyków rynkowych, konstruujących modele prognostyczne. Uzyskane wyniki badań stanowią bowiem przesłankę,

do stwierdzenia, że perspektywa ta, stanowi teoretyczny fundament, pozwalający wyjaśnić zjawiska, które dotychczas – na gruncie finansów neoklasycznych – definiowane były jako anomalie w dyskontowaniu informacji przez inwestorów giełdowych. Analogiczna rekomendacja wiąże się z włączaniem do badań i analiz, związanych z oceną bieżącej sytuacji rynkowej, czynników o charakterze niefinansowym. Rosnąca liczba danych tekstowych, co podkreślane było w teoretycznej części pracy, kreuje pilną potrzebę wypracowania metod i narzędzi, pozwalających na eksplorację tego typu danych oraz ekstrakcję z nich wartościowych treści informacyjnych.

Zakończenie

W obszarze ekonomii i finansów informacja oraz dostęp do niej odgrywają kluczową rolę. Informacja umożliwia szeroko pojętym uczestnikom życia gospodarczego podejmowanie określonych decyzji, których konsekwencją będzie zaangażowanie w daną aktywność lub nie. Fundamentalną kwestią, dla omawianego problemu, jest zdefiniowanie pojęcia informacji. Autorka postuluje, aby w – kontekście rynków finansowych – informację rozumieć nie jako samą treść, ale jako proces, którego nieodłącznymi elementami są nadawca oraz odbiorca. Taka perspektywa pozwala uchwycić problem dyskontowania informacji przez inwestorów giełdowych w możliwie szerokiej perspektywie. Należy zwrócić uwagę, że w ostatnich latach, katalog nadawców treści informacyjnych uległ rozszerzeniu. Obecnie w ten sposób określić można nie tylko spółki, czy publiczne instytucje, ale również dziennikarzy, zajmujących się tematyką ekonomiczno-finansową. Z uwagi na powszechny dostęp do Internetu, możliwość publikowania oraz odbierania tego typu treści, staje się niemal nieograniczona. Dlatego też istotne staje się poszukiwanie odpowiedzi na pytanie o możliwy wpływ nowych typów nadawców informacji, na reakcję rynków finansowych.

Celem dysertacji w wymiarze poznawczym było zidentyfikowanie oraz scharakteryzowanie kluczowych paradygmatów, opisujących sposoby dyskontowania informacji przez uczestników rynku finansowego. Cel ten zrealizowany został w części teoretycznej niniejszej pracy. Autorka, postulując procesowe ujęcie informacji, scharakteryzowała główne ogniwa procesu komunikacji, czyli nadawcę, komunikat oraz odbiorcę. Nadawca zdefiniowany został jako dziennikarz, specjalizujący się w tematyce rynków finansowych. W pracy wskazane zostały dwa kluczowe elementy, wpływające na sposób prezentacji treści w artykułach prasowych, tj. kadrowanie i czynniki selekcji informacji.

Opisanie drugiego elementu, wspomnianego procesu, wymagało podkreślenia znaczenia informacji na rynkach finansowych. Dokonano tego między innymi poprzez przeanalizowanie jej rodzajów, co obejmowało: mechanizm cenowy, informację publiczną, prywatną oraz poufną. W tym kontekście przedstawiono pojęcia, dotyczące nowego typu danych, z jakimi mamy do czynienia w praktyce badawczej, to jest

Big Data oraz *Big Text*, przez co rozumieć należy zbiory nieustrukturyzowanych danych tekstowych⁶⁸⁰. Na tym tle omówiona została koncepcja analizy sentymentu, który – na gruncie światowej literatury – definiowany jest jako stopień pozytywności lub negatywności tekstu⁶⁸¹. Autorka podkreśliła potrzebę diachronicznego rozumienia tego zjawiska w ramach mechanizmu rynkowego. Idea ta stanowi unikalny wkład w rozwój wiedzy na temat możliwości zastosowania analizy sentymentu w obszarze ekonomii i finansów.

Ostatnim opisanym elementem procesu komunikacji był odbiorca informacji, czyli sam inwestor giełdowy. W tym kontekście wyróżnione zostały dwa kluczowe paradygmaty. Pierwszy to finanse neoklasyczne oraz Hipoteza Efektywności Informacyjnej. W świetle tej perspektywy inwestor scharakteryzowany został jako osoba racjonalna, zdolna do adekwatnej interpretacji, otrzymywanych informacji, porównania dostępnych alternatyw oraz podejmowania optymalnych decyzji⁶⁸². Drugim paradygmatem było podejście behawioralne. Jego przedstawiciele podkreślają, że sposób przetwarzania informacji przez inwestorów, obarczony jest szeregiem uproszczeń i błędów poznawczych, co przekłada się na podejmowane przez nich decyzje. Powyższa dyskusja stanowiła przesłankę do stwierdzenia, że mechanizmy, związane z dyskontowaniem informacji przez inwestorów giełdowych, wymagają dalszych badań.

Dlatego też celem pracy w wymiarze badawczym było: *zbadanie i ocena sposobu dyskontowania informacji przez inwestorów giełdowych, z punktu widzenia dodatkowej perspektywy, jaką jest sentyment artykułów prasowych, na temat spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie*. Cel ten zrealizowany został w empirycznej części pracy, gdzie weryfikacji podlegała hipoteza główna:

H0: Rodzaj sentymentu, zawarty w danych tekstowych, jest odzwierciedlony w reakcji rynku.

⁶⁸⁰ Zob. S. R. Das, *Text and Context: Language Analytics in Finance, Foundations and Trends® in Finance*, 2014, VIII, s. 3-4.

⁶⁸¹ Kearney, C., Liu, S., *Textual sentiment in finance: A survey of methods and models*, „International Review of Financial Analysis”, 2014, nr 33, s. 172.

⁶⁸² Szyszka, A.: *Behavioral Finance and Capital Markets: How Psychology Influences Investors and Corporations*. Palgrave Macmillan. New York, 2013, s. 9.

Badaniu poddano 35 podmiotów, notowanych na Giełdzie Papierów Wartościowych w Warszawie w indeksie WIG20, w okresie od 1 stycznia 2013 roku do 31 grudnia 2022 roku. Hipoteza główna wsparta została trzema hipotezami pomocniczymi, które odpowiadały następującym po sobie etapom badania.

H1: Anglojęzyczne artykuły prasowe, dotyczące spółek notowanych w indeksie WIG20 w okresie od 01.01.2013 do 31.12.2022, charakteryzują się przewagą sentymentu pozytywnego względem sentymentu negatywnego.

W ramach pierwszego etapu przeprowadzona została ekstrakcja sentymentu z ponad 57 tysięcy anglojęzycznych artykułów prasowych, dotyczących spółek znajdujących się w próbie badawczej. Pozwoliło to wykazać, że wśród zgromadzonych artykułów dominuje sentyment neutralny. Ponadto, analizując relację między sentymentem pozytywnym i negatywnym, zaobserwowano, że ten pierwszy charakteryzuje ponad połowę spółek. W związku z tym autorka nie znalazła podstaw do odrzucenia pierwszej hipotezy pomocniczej.

H2: Występuje asymetria w sposobie reakcji rynku na pozytywny i negatywny sentyment, zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20 polskiej giełdy papierów wartościowych.

W drugim etapie badania analizie podlegał sposób reakcji rynku na pojawiający się sentyment pozytywny i negatywny. Do tego celu zastosowana została metoda analizy zdarzeń. Uzyskane wyniki uprawniały do stwierdzenia, że inwestorzy, związani w okresie badawczym ze spółkami notowanymi w indeksie WIG20, dyskutowali wiadomości w różny sposób, w zależności od tego, jaki sentyment reprezentował dany artykuł. Dlatego też nie znaleziono podstaw do odrzucenia drugiej hipotezy pomocniczej.

H3: Sentyment zawarty w anglojęzycznych artykułach prasowych, dotyczących spółek notowanych w indeksie WIG20, stanowi determinantę reakcji inwestorów giełdowych.

Do zrealizowania ostatniego etapu badania zbudowany został model logitowy, który zweryfikować miał możliwość zdefiniowania sentymentu jako zmiennej determinującej reakcję rynku. Uzyskane wyniki charakteryzowały się niskim stopniem

dopasowania modelu. Większość nie uzyskała również rezultatów, które określić można jako statystycznie istotne. W konsekwencji odrzucono trzecią hipotezę pomocniczą.

Przeprowadzenie całego procesu badawczego pozwoliło jednak pozytywnie zweryfikować hipotezę główną i przyjąć, że zawarty w anglojęzycznych artykułach prasowych sentyment, jest odzwierciedlony w reakcji rynku, reprezentowanej przez dzienne stopy zwrotu spółek, wchodzących w skład indeksu WIG20 w okresie od 1 stycznia 2013 roku do 30 grudnia 2022 roku. Podsumowując, można stwierdzić, iż przyjęte w niniejszej dysertacji cele, zostały zrealizowane zarówno w warstwie teoretycznej, jak i empirycznej.

Dokonany przez autorkę przegląd dotychczasowego piśmiennictwa, pozwolił przedstawić liczne badania, wskazujące na możliwość zastosowania analizy tekstu oraz sentymentu w obszarze tematyki, związanej z rynkami finansowymi. Stwierdzić można, że idea eksploracji sentymentu, jak dotąd, weryfikowana była na rynku amerykańskim, azjatyckim, australijskim, afrykańskim, jak i europejskim. Powszechność tego typu analiz wydaje się wynikać z zasadniczego waloru aplikacyjnego. Objasnienie, w jaki sposób sentyment informacji tekstowej wpływa na reakcję rynku, może dawać podstawy do uwzględniania tej zmiennej w modelach prognozowania stóp zwrotu. Tematyka ta stanowi więc przedmiot uwagi, nie tylko naukowców, ale również analityków oraz inwestorów giełdowych. Można jednak zauważyć, że o ile w krajach Europy Zachodniej, badacze wykazują zainteresowanie wykorzystaniem analizy sentymentu do oceny reakcji inwestorów na informacje, to w krajach Europy Środkowo-Wschodniej zagadnienie to cieszy się mniejszą popularnością. Zgodnie z wiedzą autorki, prezentowana dysertacja, jest pierwszą pracą, oceniającą znaczenie sentymentu, zawartego w artykułach prasowych, dla objaśnienia dziennych stóp zwrotu spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie.

Z uwagi na powyższe, podejmowanie dalszych analiz, pozwalających oszacować znaczenie sentymentu tekstowego dla dynamiki rynków finansowych, znajduje swoje uzasadnienie, zwłaszcza w krajach rozwijających się. Wydaje się, że dynamiczny rozwój algorytmów uczenia maszynowego oraz sztucznej inteligencji, otwiera w tym kontekście nowe możliwości, które pozwoliłyby ponownie zweryfikować, postawione w pracy

hipotezy badawcze. Mowa tu o alternatywnych sposobach eksploracji danych tekstowych oraz ekstrakcji sentymentu. Zwłaszcza jednak chodzi o możliwości wykorzystania tego typu danych do predykcji stóp zwrotu z różnego typu walorów. To ten aspekt bowiem znaczenie nie tylko z punktu widzenia akademickiego, ale również praktycznego. W tym celu nieodzowne jest jednak zwrócenie się ku paradygmatowi finansów behawioralnych, co pozwala dostrzec znaczenie czynników niefinansowych w procesie podejmowania decyzji przez inwestorów giełdowych.

Bibliografia

1. Agrawal A., Jaffe J.F., Mandelker G.N., *The post-merger performance of acquiring firms: A re-examination of an anomaly*, „The Journal of Finance” R.1992. nr XLVII/4, s.1605–1621.
2. Ahern, K. R., Sosyura, D., *Rumor Has It: Sensationalism in Financial Media*, „The Review of Financial Studies”, R. 2015, nr 28/7.
3. Aite Group, *Big Data in Capital Markets: At the Start of the Journey*, 2014.
4. Akerlof, G.A., *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, [w:] (red.), Akerlof, G.A., *An economic theorist's book of tales*, Cambridge University Press, Cambridge, 1984.
5. Allen, D. E., McAleer, M., Singh, A. K., *Daily market news sentiment and stock prices*, „Applied Economics”, R. 2019, nr 51/30, s. 3212–3235.
6. Alomari, M., Al Rababa'a, A. R., El-Nader, G., Alkhataybeh, A., Rehman, U. M., *Examining the effects of news and media sentiments on volatility and correlation: Evidence from the UK*, „The Quarterly Review of Economics and Finance”, R. 2021, nr 82, s. 280–297.
7. American Association for Public Opinion Research, *AAPOR Report on Big Data*, AAPOR Big Data Task Force, 2015.
8. Araújo, T., Eleutério, S., Louçã, F., *Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood*, „Physica A.”, R. 2018, nr 505, s. 1139–1149.
9. Atkins, A., Niranjana, M., Gerding, E., *Financial news predicts stock market volatility better than close price*, „The Journal of Finance and Data Science”, R. 2018, nr 4, s. 120–137.
10. Audrino, F., Sigrist, F., Ballinari, D., *The impact of sentiment and attention measures on stock market volatility*, „International Journal of Forecasting”, R.2020, nr 36/2, s.334–357.
11. Bageho, W., *The only game in town*, „Financial Analysts Journal”, R.1971, s.12–22.
12. Baker, H. K., Nofsinger, J. R., *Behavioral finance: Investors, corporations, and markets*, John Wiley & Sons, New Jersey 2010.
13. Banarescu, A., *Detecting and Preventing Fraud with Data Analytics*, „Procedia Economics and Finance.”, R. 2015, nr 32.
14. Banaszczyk-Soroka, U., *Rynek papierów wartościowych. Inwestorzy, instrumenty finansowe i metody wyceny*, C.H.Beck, Warszawa 2016.
15. Bank Światowy, *World Development Indicators*, <https://databank.worldbank.org/reports.aspx?source=2&series=IT.NET.USER.ZS&country=#>. Dostęp: 1 kw. 2023.
16. Bankier, <https://www.bankier.pl/>, Dostęp: 10 sty. 2024.
17. BBC News, *City Slickers guilty of tip scam*, http://news.bbc.co.uk/2/hi/uk_news/4507448.stm. Dostęp: 20 lip. 2022.

18. Bednarek, M., Caple, H., *The Discourse of News Values. How news organizations create newsworthiness*, Oxford University Press, New York 2017.
19. Będowska-Sójka, B., *Wpływ informacji na ceny instrumentów finansowych. Analiza danych śróddziennych*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2014.
20. Biblioteka Uniwersytetu Warszawskiego, *Bazy online*, <https://www.buw.uw.edu.pl/zasoby-online/bazy-online/>, Dostęp: 10 sty. 2024.
21. Biddle, G., Hilary, G., Verdi, R.S., *How does financial reporting quality relate to investment efficiency?*, „Journal of Accounting and Economics”, R. 2009, nr 48/2–3, s. 112–131.
22. Biktimirov, E. N., Sokolyk, T., Ayanso, A., *Sentiment and hype of business media topics and stock market returns during the COVID-19 pandemic*, „Journal of Behavioral and Experimental Finance”, R.2021, nr 31, s.1–6.
23. Bird, S., Klein, E., Loper, E., *Natural Language Processing with Python*, O'Reilly Media, Sebastopol 2009.
24. Black, F., *Noise*, „Journal of Finance”, R. 1986, nr 41/3, s. 529–43.
25. Blazquez, D., Domenech, J., *Big Data sources and methods for social and economic analyses*, „Technological Forecasting & Social Change”, R. 2017, nr 130, s. 99–113.
26. Bloomfield, R. J., *The “Incomplete Revelation Hypothesis” and financial reporting*, „Accounting Horizons”, R. 2002, nr 16/3, s. 233–243.
27. Boisot, M.H., *Knowledge Assets: Securing Competitive Advantage in the Information Economy*, Oxford University Press, Oxford 1998.
28. Borden, M.J., *The Role of Financial Reporting in Corporate Governance*, „Fordham Journal of Corporate & Financial Law”, R. 2007, nr XII, s. 311–369.
29. Borovkova, S., Lammiman, A., *The impact of news sentiment on energy futures returns*, Working Paper, R.2023.
30. Borowski, K., *Finanse behawioralne. Modele*, Wydawnictwo Difin, Warszawa 2014.
31. Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., *Which News Moves Stock Prices? A Textual Analysis*. National Bureau of Economic Research, Working Paper, R. 2013, nr 18725, s. 1–45.
32. Boukes, M., Jones, N.P., Vliegthart, R., *Newsworthiness and story prominence: How the presence of news factors relates to upfront position and length of news stories*, „Journalism”, R. 2022, nr 23/1, s. 98–116.
33. Boukes, M., Jones, Natalie P., Vliegthart, R., *Newsworthiness and story prominence: How the presence of news factors relates to upfront position and length of news stories*, „Journalism”, R. 2020, s. 1–19.
34. Boukes, M., Vliegthart, R., *A general pattern in the construction of economic newsworthiness? Analyzing news factors in popular, quality, regional, and financial newspapers*, „Journalism”, R.2020, nr 21/2.

35. Bouteska, A., Mefteh-Wali, A., Dang, T., *Predictive power of investor sentiment for Bitcoin returns: Evidence from COVID-19 pandemic*, „Technological Forecasting and Social Change”, R.2022, nr 184, s.1–13.
36. Bouteska, A., *The effect of investor sentiment on market reactions to financial earnings restatements: Lessons from the United States*, „Journal of Behavioral and Experimental Finance”, R.2019, nr 24, s.1–11.
37. Bowers, S., *Three face jail over Mirror share fraud*, <https://www.theguardian.com/media/2005/dec/08/pressandpublishing.mirror1>. Dostęp: 20 lip. 2022.
38. Brigham, E.F., Gapenski, L.C., *Zarządzanie Finansami*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2000.
39. Bukovina, J., *Social media big data and capital markets – An overview*, „Journal of Behavioral and Experimental Finance”, R. 2016, nr 11.
40. Bulski, D., Górski, M., *Wyniki testów hipotezy pól silnej efektywności informacyjnej Giełdy Papierów Wartościowych w Warszawie w latach 2005–2011*, „Problemy Zarządzania”, R. 2012, nr 10/4/39, s. 141–168.
41. Burgin, M., *Theory Of Information: Fundamentality, Diversity And Unification*, World Scientific, Los Angeles 2009.
42. Butterick, K. J., *A Churnalism, Complacency and Collusion: A Critical Introduction to Business and Financial Journalism*, Pluto Press, Londyn, 2015.
43. Buzala, P., *Reakcja inwestorów na rekomendacje giełdowe. Implikacje dla efektywności rynku akcji*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia”, R. 2012, nr 689/50, s. 11–20.
44. Buzala, P., *Silna a pól silna forma efektywności informacyjnej rynku kapitałowego – doświadczenia z GPW w Warszawie na podstawie zaleceń analityków*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego. Finanse, Rynki Finansowe, Ubezpieczenia”, R. 2015, nr 862/75, s. 97–109.
45. Cabena P., Hadjinian, P., Stadler, R., Verhees, J., Zanasi, A., *Discovering Data Mining: From Concept to Implementation*, Prentice Hall, New Jersey 1998.
46. Caiffa, M., Farina, V., Fattobene, L., *CEO Duality: Newspapers and Stock Market Reactions*, „Journal of Risk and Financial Management”, R.2021, nr 14/35, s.1–18.
47. Calomiris, Ch. W., Mamaysky, H., *How news and its context drive risk and returns around the world*, „Journal of Financial Economics”, R.2019, nr 133. s.299–336.
48. Cameron, A., Palan, R., *The imagined economies of globalization*, SAGE Publications, Londyn 2004.
49. Camiciottoli, B. C., *Using English as a lingua franca to engage with investors: An analysis of Italian and Japanese companies' investor relations communication policies*, „English for Specific Purposes”, R.2020, nr 58, s.90–101.
50. Carmona, J., Delcoure, N., Fernandez, F., H., *Social media sentiment, tariffs, and international equity pricing*, „International Journal of Electronic Finance”, R.2020, nr 10/1–2, s.93–115.

51. Cecchini, M., Aytug, H., Koehler, G. J., Pathak, P., *Making words work: Using financial text as a predictor of financial events*, „Decision Support Systems”, R. 2010, nr 50, s. 164–175.
52. Cernat, M., *The Ethics of financial journalism. A critical perspective*, „Cogito – Multidisciplinary Research Journal”, R. 2014, nr VI/3.
53. Cheikh, E., Cullot, N., Nicolle, Ch., *Understandable Big Data: A Survey*, „Computer Science Review”, R. 2015, nr 17, s. 70–81.
54. Chen W., Lai, K. K., Cai, Y., *Exploring public mood toward commodity markets: a comparative study of user behavior on Sina Weibo and Twitter*, „Internet Research” R.2020. nr 31/3, s.1102–1119.
55. Chen, G., Kim, K.A., Nofsinger, J.R., Rui, O.M., *Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors*, „Journal of Behavioral Decision Making”, R. 2007, nr 20, s. 425–451.
56. Chen, H., De, P., Hu, Y., Hwang, B-H., *Wisdom of crowds: The value of stock opinions transmitted through social media*, „The Review of Financial Studies”, R. 2014, nr 27/5, s. 1367–1403.
57. Chen, H., Zimbra, D., *AI and Opinion Mining*, „IEEE Intelligent Systems”, 2010, nr 25/3, s. 74–80.
58. Chen, J., *Information Theory and Market Behavior*, „CFAI Journal of Behavioral Finance”, R. 2005, nr 2/4, s. 25–45.
59. Chen, K., Luo, P., Liu, L., Zhang, W., *News, search and stock co-movement: Investigating information diffusion in the financial market*, „Electronic Commerce Research and Applications”, R. 2018, nr 28, s. 159–171.
60. Chen, P., Chun-Yang, Z., *Data-intensive applications, challenges, techniques and technologies: A survey on Big Data*, „Information Sciences”, R. 2014, nr 275, s. 314–347.
61. Chen. H., Chiang. R., Storey. V., *Business intelligence and analytics: from big data to big impact*, „MIS Quarterly”, R. 2012, nr 36/4, s. 1165–1188.
62. Choi, S., *Measuring economic diplomacy using event study method: the case of EU-China summit talks and Airbus stock price changes*, „Asia Europe Journal”, R.2023, nr 21, s.155–171.
63. Chopra, A., Prashar, A., Sain, Ch., *Natural Language Processing*, „International Journal of Technology Enhancements and Emerging Engineering Research”, nr 1/4, s. 131–134.
64. Chou, R.K., Wang, Y-Y., *A Test of the Different Implications of the Overconfidence and Disposition Hypotheses*, „Journal of Banking and Finance”, R. 2011, nr 35/8, s. 2037–2046.
65. Chowdhury, G. G., *Natural language processing*, „Annual Review of Information Science and Technology”, R. 2003, nr 37/1, s. 51–89.
66. Chuang, W-I., Susmel, R., *Who Is the More Overconfident Trader? Individual vs. Institutional Investors*, „Journal of Banking and Finance”, 2011, nr 35/7, s. 1626–1644.

67. Chun, S. H., Kim, S. H., *Data mining for financial prediction and trading: Application to single and multiple markets*, „Expert Systems with Applications”, R. 2004, nr 26/2, s. 131–139.
68. Ciolek, M., *Efektywność informacyjna polskiego rynku akcji*, CeDeWu, Warszawa 2015.
69. Cohen, R., Gompers, P., Vuolteenaho, T., *Who underreact to cash-flow news? Evidence from trading between individuals and institutions*, „Journal of Financial Economics”, R. 2002, nr 66, s. 409–462.
70. Colasanto, F., Grilli, L., Santoro, D., Villani, G., *ALBERTino for stock price prediction: a Gibbs sampling approach*, „Information Sciences”, R. 2022, nr 597, s. 341–357.
71. Columbia University Libraries, *Research Guides. A-Z Databases*, <https://guides.library.columbia.edu/az.php?a=all>, Dostęp: 10 sty. 2024.
72. Czekaj, J., *Efektywność giełdowego rynku akcji w Polsce. Z perspektywy dwudziestolecia*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2014.
73. Czekaj, J., Woś, M., Żarnowski J., *Efektywność giełdowego rynku akcji w Polsce. Z perspektywy dziesięciolecia*, Polskie Wydawnictwo Naukowe, Warszawa 2001.
74. Czerwonka, M., Gorlewski, B., *Finanse behawioralne. Zachowania inwestorów i rynku*, Oficyna Wydawnicza Szkoły Głównej Handlowej, Warszawa 2012.
75. Czupryna, M., Żarnowski, J., Efektywność pólsilna, [w:] (red.) Czekaj, J., *Efektywność giełdowego rynku akcji w Polsce z perspektywy dwudziestolecia*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2014.
76. Da. Z., Engelberg. J., Gao. P., *The sum of all FEARS: Investor Sentiment and Asset Prices*, „SSRN Electronic Journal”, R. 2013, s. 1–42.
77. Damodaran, A., *Finanse korporacyjne. Teoria i praktyka*, Wydawnictwo Helion, Gliwice, 2017.
78. Damstra, A., De Swert, K., *The making of economic news: Dutch economic journalists contextualizing their work*, „Journalism”, R. 2020, nr 22/12, s. 3083–3100.
79. Das, S. R., Chen, M.Y., *Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web*, „Management Science”, R. 2007, nr 53/9, s. 1375–1388.
80. Das. S. R., *Text and Context: Language Analytics in Finance. Foundations and Trends in Finance*, now Publishers Inc, Hanover 2014.
81. Davis, A., *Public relations, business news and the reproduction of corporate elite power*, „Journalism”, R. 2000, nr 1/3, s. 282–304.
82. De Bondt, W. F. M., Thaler R., *Does the Stock Market Overreact?*, „The Journal of Finance”, R. 1985, nr 40/3, s. 793–805.
83. De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., *Noise Trader Risk in Financial Markets*, „Journal of Political Economy”, R. 1990, nr 98/4, s. 703–738.
84. De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., *Positive feedback investment strategies and destabilizing rational speculation*, „Journal of Finance”, R. 1990, nr 45/2, s. 379–395.

85. de Saussure, F., *Kurs językoznawstwa ogólnego*, Państwowe Wydawnictwo Naukowe, Warszawa 1961.
86. de Vreese, C.H., *News framing, Theory and typology*, „Information Design Journal”, R. 2005, nr 13/1, s. 51–62.
87. deHaaff, M., *Sentiment Analysis, Hard But Worth It!*, https://customerthink.com/sentiment_analysis_hard_but_worth_it. Dostęp: 20 lip. 2022.
88. Delcey, T., *Samuelson vs Fama on the efficient market hypothesis: The point of view of expertise*, „Economia – History / Methodology / Philosophy”, R. 2019, nr 9/1, s. 37–57.
89. Deng, S., Huang, Z. J., Sinha, A. P., & Zhao, H., *The interaction between microblog sentiment and stock returns: An empirical examination*, „MIS Quarterly: Management Information Systems”, R.2018, nr 42/3, s.895–918.
90. Deutsche Bank, *Big Data How it can become a differentiator*, 2014.
91. Dębski, W., *Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki*, Polskie Wydawnictwo Naukowe, Warszawa 2014.
92. Digital Marketing Community, *How People Use Their Devices: What Marketers Need to Know*, 2016 | *Think With Google*, <https://www.digitalmarketingcommunity.com/researches/how-people-use-their-devices-what-marketers-need-to-know-sep-2016-think-with-google/>. Dostęp: 1 kw. 2023.
93. Dobrowolski, K., *Teoria rynków efektywnych i model racjonalnego inwestora – od warunków ryzyka do warunków konfliktu*, „Współczesna Gospodarka”, R. 2015, nr 6/2, s. 11–19.
94. Doman, M., *Mikrostruktura Gield Papierów Wartościowych*, Wydawnictwo Uniwersytetu w Poznaniu, Poznań 2011.
95. Dong, W., Liao, S., Liang, L., *Financial Statement Fraud Detection using Text Mining: a systemic functional linguistics theory perspective*, PACIS 2016, Proceedings 2016, Chiayi, Taiwan
96. Donohue, B., *How to Fail at Corporate Fraud*, <https://threatpost.com/how-fail-corporate-fraud-010813/77378/>. Dostęp: 1 kw. 2023.
97. Dougal, C., Engelberg, J., García, D., Parsons, Ch. A., *Journalists and the stock market*, „The Review of Financial Studies”, R. 2012, nr 25/3, s. 639–679.
98. Doyle, G., *Financial news journalism. A post-Enron analysis of approaches towards economic and financial news production in the UK*, „Journalism”, R.2006, nr 7/4, s. 433–452.
99. Dretske, F., *Knowledge and the Flow of Information*, MIT Press, Cambridge 1981.
100. Duarte, F., *Amount of Data Created Daily (2023)*, <https://explodingtopics.com/blog/data-generated-per-day#how-much>. Dostęp: 1 kw. 2023.
101. Dzielinski, M., Hasseltoft, H., *Aggregate News Tone, Stock Returns, and Volatility*, „SSRN Electronic Journal”, R. 2013, s. 1–62.

102. Eilders, Ch., *News factors and news decisions. Theoretical and methodological advances in Germany*, „Communications”, R. 2006, nr 31/3, s. 5–24.
103. El-Haj, M., Rayson, P., Walker, M., Young, S., Simaki, V., *In search of meaning: Lessons, resources and next steps for computational analysis of financial discourse*, „Journal of Business Finance and Accounting”, R. 2019, nr 46, s. 265–306.
104. Emerging Markets Information Service, <https://www.emis.com/pl>, Dostęp: 10 sty. 2024.
105. Entman, R. M., *Framing: Toward Clarification of a Fractured Paradigm*, „Journal of Communication”, R. 1993, nr 43/1, s. 51–58.
106. Erdas, M. L., *Validity of Weak-Form Market Efficiency in Central and Eastern European Countries (CEECs): Evidence from Linear and Nonlinear Unit Root Tests*, „Review of Economic Perspectives”, R. 2019, nr 19/4, s. 399–428.
107. Ernst&Young, *Forensic data analytics – Globally integrated compliance review, litigation support and investigative services*, [http://www.ey.com/Publication/vwLUAssets/EY_-_Forensic_Data_Analytics_\(FDA\)/\\$FILE/EY-forensic-data-dnalytics.pdf](http://www.ey.com/Publication/vwLUAssets/EY_-_Forensic_Data_Analytics_(FDA)/$FILE/EY-forensic-data-dnalytics.pdf). Dostęp: 20 sie. 2017.
108. Fama, E. F., *Efficient Capital Markets: A Review of Theory and Empirical Work*, „The Journal of Finance”, R. 1970, nr 25/2, s. 383–417.
109. Fama, E. F., French, K. R., *A five-factor asset pricing model*, „Journal of Financial Economics”, R. 2015, nr 116, s. 1–22.
110. Fama, E. F., French, K. R., *Common risk factors in the returns on stocks and bonds*, „Journal of Financial Economics”, R. 1993, nr 33/1, s. 3–56.
111. Fama, E. F., *Market efficiency, long-term returns, and behavioral finance*, „Journal of Financial Economics”, R. 1998, nr 49, s. 283–306.
112. Fama, E. F., *Random Walks in Stock Market Prices*, „Financial Analysts Journal”, R. 1965, nr 21/5, s. 55–59.
113. Fama, E. F., *Risk, Return and Equilibrium: Some Clarifying Comments*, „The Journal of Finance”, R. 1968, nr 23/1, s. 29–40.
114. Fama, E.F., *The Behavior of Stock-Market Prices*, „The Journal of Business”, R. 1965, nr 38/1, s. 34–105.
115. Fawcett, T., *An introduction to ROC analysis*, „Pattern Recognition Letters”, R.2006, nr 27, s.861–874.
116. Fedorova, E., Druchok, S., Drogozov, P., *Impact of news sentiment and topics on IPO underpricing: US evidence*, „International Journal of Accounting and Information Management”, R.2022, nr 30/1, s.73–94.
117. Feuerriegel, S., Gordon, J., *Long-term stock index forecasting based on text mining of regulatory disclosures*, „Decision Support Systems”, R. 2018, nr 112, s. 88–97.
118. Fishman, M. J., Hagerty, K. M., *Insider Trading and the Efficiency of Stock Prices*, „The RAND Journal of Economics Journal of Economics”, R. 1992, nr 23/1, s. 106–122.

119. Fleischer, M., *Zarys ogólnej teorii komunikacji*, [w:] (red.) Graszewicz, M., Jastrzębski, J., *Teorie komunikacji i mediów*, Oficyna Wydawnicza ATUT, Wrocław 2009.
120. Flynn, J., *How Many People Use The Internet? [2023]: 35 Facts About Internet Usage In America And The World*, <https://www.zippia.com/advice/how-many-people-use-the-internet/>. Dostęp: 1 kw. 2023.
121. Fowler, R., *Language in the news: Discourse and ideology in the press*, Routledge, Nowy Jork 1991.
122. Francis, J. C., *Inwestycje. Analiza i zarządzanie*, WIG-Press, Warszawa 2000.
123. Fung, G. P. Ch., Yu, J. X., Lam, W., *Stock prediction: Integrating text mining approach using real-time news*, IEEE International Conference on Computational Intelligence for Financial Engineering, Hong Kong, Chiny 2003.
124. Gaikwad, S. V., Chaugule, A., Patil, P., *Text Mining Methods and Techniques*, „International Journal of Computer Applications”, R. 2014, nr 85/17, s. 42–45.
125. Gajdka, J., *Behawioralne finanse przedsiębiorstw. Podstawowe podejścia i koncepcje*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2013.
126. Galtung, J., Ruge, M. H., *The Structure of Foreign News*, „Journal of Peace Research”, R. 1965, nr 2/1, s. 64–9.
127. Gantz, J., Reinsel, D., *IDC Iview. Extracting Value from Chaos*, 2011.
128. Garcia, D., *Sentiment during recessions*, „The Journal of Finance”, R. 2013, nr 68/3, s. 1267–1300.
129. Garcia-Lopez, F. J., Batyrshin, I., Gelbukh, A., *Analysis of relationships between tweets and stock market trends*, „Journal of Intelligent & Fuzzy Systems”, R. 2018, nr 34, s. 3337–3347.
130. Gartner, Inc., *Data Mining*, <http://www.gartner.com/it-glossary/data-mining>. Dostęp: 20 lip. 2022.
131. General Inquirer, <https://inquirer.sites.fas.harvard.edu/>. Dostęp: 1 kw. 2023.
132. Ghailan, O., Mokhtar, H., Hegazy, O., *Improving Credit Scorecard Modeling Through Applying Text Analysis*, „International Journal of Advanced Computer Science and Applications”, R. 2016, nr 7/4.
133. Giełda Papierów Wartościowych w Warszawie, *Udział inwestorów w obrotach giełdowych - dane zagregowane 2009 – I półrocze 2023*, https://www.gpw.pl/pub/GPW/analizy/Udzial_inwestorow_1H2023.pdf, Dostęp: 10 sty. 2024.
134. Giełda Papierów Wartościowych w Warszawie, *Zagregowane wyniki badania dot. udziału inwestorów w obrotach instrumentami finansowymi*, https://www.gpw.pl/pub/GPW/analizy/2015_P2_Udzialy_20160222.pdf, Dostęp: 10 sty. 2024.
135. Giglio, S., Kelly, B., *Excess volatility: Beyond discount rates*, „The Quarterly Journal of Economics”, R. 2018, nr 133/1, s. 71–127.

136. Glancy, F. H., Yadav, S. B., *A computational model for financial reporting fraud detection*, „Decision Support Systems”, R. 2011, nr 50, s. 595–601.
137. Główny Urząd Statystyczny, *Spółeczeństwo informacyjne w Polsce w 2022 r.*, 2022.
138. Główny Urząd Statystyczny, *Spółeczeństwo informacyjne w Polsce. Wyniki badań statystycznych za lata 2012–2016*, 2016.
139. Goel, S., Gangolly, J., Faerman, S.R., Uzuner, O., *Can Linguistic Predictors Detect Fraudulent Financial Filings?*, „Journal of Emerging Technologies in Accounting”, R. 2010, nr 7, s. 25–46.
140. Goldman, E., Martel, J., Schneemeier, J., *A Theory of Financial Media*, „Journal of Financial Economics”, R. 2022, nr 145/1, s. 239–258.
141. Goodfellow, Ch., Bohl, M.T., Gebka, B., *Together we invest? Individual and institutional investors' trading behaviour in Poland*, „International Review of Financial Analysis”, R. 2009, nr 18, s. 212–221.
142. Gordon, B., Rittenberg, L., *The Warsaw Stock Exchange: A Test of Market Efficiency*, „Comparative Economic Studies”, R. 1995, nr 37/2, s. 1–27.
143. Goswami, P., Madan, S., *A Survey on Big Data & Privacy Preserving Publishing Techniques*, „Advances in Computational Sciences and Technology”, R. 2017, nr 10/3, s.395–408.
144. Graszewicz, M., Jastrzębski, J., (red) *Teorie komunikacji i mediów*, Oficyna Wydawnicza ATUT, Wrocław 2009.
145. Griffin, D., Tversky, A., *The weighing of evidence and the determinants of confidence*, „Cognitive Psychology”, R. 1992, nr 24, s. 411–435.
146. Groth, S. S., Muntermann, J., *An intraday market risk management approach based on textual analysis*, „Decision Support Systems”, R.2011, nr 50, s.680–691.
147. Gruszczyński, M., Kuszewski, T., Podgórska, M., *Ekonometria i badania operacyjne*, Wydawnictwo Naukowe PWN, Warszawa 2009.
148. Gu, C., Kurov, A., *Informational role of social media: Evidence from Twitter sentiment*, „Journal of Banking and Finance”, R.2020, nr 121, s.1–13.
149. Guo, Li., Shi, F., Tu, J., *Textual analysis and machine leaning: Crack unstructured data in finance and accounting*, „The Journal of Finance and Data Science”, R. 2016, nr 2/3, s. 153–170.
150. Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna a Wolters Kluwer business, Warszawa 2012.
151. Gurgul, H., *Analiza zdarzeń na rynkach akcji. Wpływ informacji na ceny papierów wartościowych*, Oficyna Ekonomiczna, Kraków 2006.
152. Hajek, J., *Weak-form Efficiency Test in the Central European Capital Markets*, „Politická Ekonomie”, R. 2007, nr 6, s. 773–791.
153. Hajek, P., Henriques, R., *Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods*, „Knowledge-Based Systems”, R. 2017, nr 128, s. 139–152.

154. Hall, S., Critcher, C., Jefferson, T., Clarke, J., Roberts, B., *Policing the Crisis: Mugging, the State, and Law and Order*, Macmillan, Londyn 1978.
155. Han, J., Kamber, M., *Data Mining Concepts and Techniques*, Elsevier, San Francisco 2006.
156. Hanna, A. J., Turner, J. D., Walker, C. B., *News media and investor sentiment during bull and bear markets*, „The European Journal of Finance”, R.2020, nr 26/14, s.1377–1395.
157. Harcup, T., O'Neill, D., *What Is News? Galtung and Ruge revisited*, „Journalism Studies”, R. 2010, nr 2/2, s. 261–280.
158. Harvard Library, *A guide to library resources in support of research in urban planning and design*, <https://guides.library.harvard.edu/gsd/urban>, Dostęp: 10 sty. 2024.
159. Hayek Fredrich A., *The Use of Knowledge in Society*, „The American Economic Review”, R. 1945, nr 35/4, s. 519–530.
160. Henry, E., *Are investors influenced by how earnings press releases are written?*, „Journal of Business Communication”, R.2008, nr 45/4, s.363–407.
161. Heston, S. L., Sinha, N. R., *News versus Sentiment: Predicting Stock Returns from News Stories*, „Finance and Economics Discussion Series. Washington: Board of Governors of the Federal Reserve System”, R. 2016, nr 2016–048, s. 1–35.
162. Holton, C., *Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem*, „Decision Support Systems Journal”, R. 2009, nr 46, s. 853–864.
163. Huang, W., *The Role of Financial Media in Corporate Financing*, „SSRN Electronic Journal”, R. 2010, s. 1–36.
164. Hurwitz, J., Nugent, A., Halper, F., Kaufman, M., *Big Data for Dummies*, John Wiley & Sons Inc, New Jersey 2013.
165. Iatridis, G. E., *Financial reporting language in financial statements: Does pessimism restrict the potential for managerial opportunism?*, „International Review of Financial Analysis”, R. 2016, nr 45, s. 1–17.
166. Im, L. T., San, W. P., Anthony, P., On, K. C., *Improving polarity classification for financial news using semantic similarity techniques*, „International Journal of Intelligent Information Technologies”, R. 2018, nr 14/4, s. 39–54.
167. International Data Corporation, *Asia/Pacific Big Data and Cognitive/Artificial Intelligence Systems and Content Analytics*, https://www.idc.com/getdoc.jsp?containerId=IDC_P24232. Dostęp: 20 lip. 2022.
168. International Telecommunication Union, *ICT Facts And Figures 2017*, 2017.
169. Jackson, P., Moulinier, I., *Natural Language Processing for Online Applications: Text Retrieval, Extraction, and Categorization*, John Benjamins Publishing, Amsterdam 2007.
170. Jajuga, K., Jajuga, T., *Inwestycje. Instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa*, Wydawnictwo Naukowe PWN, Warszawa 2008.

171. Jajuga, K., *Równowaga rynku kapitałowego – pół wieku historii rodziny CAPM*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego, Finanse, Rynki Finansowe, Ubezpieczenia”, R. 2013, nr 768/63, s. 181–192.
172. Jancenelle, V. E., Buccieri, D., *Signaling downsizing intentions after a major crisis: does managerial authenticity matter?*, „International Journal of Organizational Analysis”, R.2022, nr 31/1, s.7–26.
173. Jancenelle, V. E., *Moral foundations and financial performance: the effect of moral cues during times of earnings uncertainty*, „International Journal of Organizational Analysis”, R.2020, nr 30/2, s.207–222.
174. Jancenelle, V. E., *Organizational Psychological Capital During Earnings Conference Calls: Mitigating Shareholders' Sell-Off in the Face of Earnings Surprises?*, „Journal of Leadership & Organizational Studies”, R.2018, nr 25/4, s.469–480.
175. Jancenelle, V. E., Storrud-Barnes, S. F., Iaquinto, A., *Making investors feel good during earnings conference calls: The effect of warm-glow rhetoric*, „Journal of General Management”, R.2019, nr 44/2, s.63–72.
176. Jegadeesh, N., *Evidence of Predictable Behavior of Security Returns*, „Journal of Finance”, R. 1990, nr 45/3, s. 881–898.
177. Jegadeesh, N., Titman, S., *Overreaction, Delayed Reaction, and Contrarian Profits*, „Review of Financial Studies”, R. 1995, nr 8/4, s. 973–993.
178. Jegadeesh, N., Titman, S., *Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency*, „The Journal of Finance”, R. 1993, nr 48/1, s. 65–91.
179. Jiang, W., *A Model of Competitive Stock Trading Volume*, „Journal of Political Economy”, R. 1994, nr 102/1, s. 127–168.
180. Johnman, M., Vanstone, B. J., Gepp, A., *Predicting FTSE 100 returns and volatility using sentiment analysis*, „Accounting and Finance”, R.2018, nr 58, s.253–274.
181. Justia U.S. Supreme Court Center: *Carpenter v. United States*, 484 U.S. 19 (1987). No. 86–422, <https://supreme.justia.com/cases/federal/us/484/19>. Dostęp: 20 lip. 2022.
182. Kahneman, D., *Pułapki myślenia. O myśleniu szybkim i wolnym*, Media Rodzina, Poznań 2011.
183. Kahneman, D., Tversky, A., *Prospect Theory: An Analysis of Decision under Risk*, „The Econometric Society”, R. 1979, nr 47/2, s. 263–292.
184. Kallen, M. J., Scholtens, B., *Movers and Shakers: Stock Market Response to Induced Seismicity in Oil and Gas Business*, „Energies”, R.2021, nr 14/23.
185. Kantola, A., *On the dark side of democracy: the global imaginary of financial journalism*, [w:] (red.) Cammaerts, E., Carpentier, N., *Reclaiming the Media. Communication Rights and Democratic Media Roles*, European Communication Research and Education Association, Bristol 2007.
186. Kavšek, B., *Using Words from Daily News Headlines to Predict the Movement of Stock Market Indices*, „Managing Global Transitions”, R. 2017. nr 15/2, s. 109–121.
187. Kearney, C., Liu, S., *Textual sentiment in finance: A survey of methods and models*, „International Review of Financial Analysis”, R. 2014, nr 33.

188. Kelly, S., Ahmad, K., *Estimating the impact of domain-specific news sentiment on financial assets*, „Knowledge-Based Systems”, R.2018, nr 150, s.116–126.
189. Kepios Pte., Ltd., *Digital 2023: Poland. The essential guide to the latest connected behaviours*, <https://indd.adobe.com/view/f37abde9-e4ac-4df0-8eef-3733deaa1cc2?allowFullscreen=true&wmode=opaque>. Dostęp: 1 kw. 2023.
190. Khedr, A. E., Salama, S. E., Yaseen, N., *Predicting stock market behavior using data mining technique and news sentiment analysis*, „International Journal of Intelligent Systems and Applications”, R. 2017, nr 7, s. 22–30.
191. Kim, Y., Jeong, R.S., Ghani, I., *Text Opinion Mining to Analyze News for Stock Market Prediction*, „International Journal of Advances in Soft Computing and Its Applications”, R. 2014, nr 6/1, s. 1–13.
192. Kleinnijenhuis, J., Schultz, F., Oegema, D., van Atteveldt, W., *Financial news and market panics in the age of high-frequency sentiment trading algorithms*, „Journalism”, R. 2013, nr 14/2, s. 271–291.
193. Knowles, S., Philips, G., Lidberg, J., *Reporting The Global Financial Crisis: A longitudinal tri-nation study of mainstream financial journalism*, „Journalism Studies”, R. 2017, nr 18/3, s. 322–340.
194. Koc, J. O., *Do Institutional Investors Get Affected by Biases in Their Individual Investments*, „International Journal of Business & Management Studies”, R. 2022, nr 3/1, s. 44–58.
195. *Kodeks Dobrych Praktyk Wydawców Prasy*, Izba Wydawców Prasy.
196. *Kodeks Etyki Stowarzyszenia Dziennikarzy Polskich*, Stowarzyszenie Dziennikarzy Polskich.
197. Kompa, K., Matuszewska-Janicka, A., *The Warsaw Stock Exchange: A Test of Market Efficiency*, „International Advances in Economic Research”, R. 2009, nr 15, s. 59–70.
198. Kowalczyk, E., *O istocie informacji*, Wydawnictwo Komunikacji i Łączności, Warszawa 1986.
199. Kuhn, T. S., *Struktura rewolucji naukowych*, Fundacja Aletheia Warszawa 2001.
200. Kurtz, H., *Is the economy suffering from media malady?*, <https://www.washingtonpost.com/archive/business/1990/10/28/is-the-economy-suffering-from-media-malady/52d9122b-792d-480d-9658-483e334f31e2>. Dostęp: 9 mar. 2023.
201. Lakonishok, J., Shleifer, A., Thaler, R., Vishny, R., *Window Dressing by Pension Fund Managers*, „Behavioral Finance”, R. 1991, nr 81/2, s. 227–231.
202. Larose, D.T., *Discovering Knowledge In Data An Introduction to Data Mining*, John Wiley & Sons Inc, New Jersey 2005.
203. Larsen, V. H., Thorsrud, L. A., *Asset returns, news topics, and media effects*, „CAMP Working Paper Series”, R. 2017, nr 5, s. 1–48.
204. Lawrence, A., *Individual investors and financial disclosure*, „Journal of Accounting & Economics”, R. 2013, nr 56, s. 130–147.

205. Lazaro, D.C., Medalla, E. M., English as the Language of Trade, *Finance and Technology in APEC: An East Asia Perspective*, „Philippine Journal of Development”, R.2004, nr 58/XXXI/2.
206. Lazzini A., Lazzini S., Balluchi F., Mazza M., *Emotions, moods and hyperreality: social media and the stock market during the first phase of COVID-19 pandemic*, „Accounting, Auditing and Accountability Journal”, R.2022, nr 35/1, s.199–215.
207. Lehmann, B., *Fads, Martingales, and Market Efficiency*, „Quarterly Journal of Economics”, R. 1990, nr 105/1, s. 1–28.
208. Leibenstein, H., *General X-efficiency Theory and Economic Development*, Oxford University Press, Nowy Jork 1978.
209. Li, F., *Annual report readability, current earnings, and earnings persistence*, „Journal of Accounting and Economics”, R. 2008, nr 45, s. 221–247.
210. Li, F., *Do stockmarket investors understand the risk sentiment of corporate annual reports?*, „SSRN Electronic Journal”, R.2006, s. 1–53.
211. Li, F., *Textual Analysis of Corporate Disclosures: A Survey of the Literature*, „Journal of Accounting Literature”, R. 2010, nr 29.
212. Li, J. i inni, *A novel text-based framework for forecasting agricultural futures using massive online news headlines*, „International Journal of Forecasting”, R.2022, nr 38, s.35–50.
213. Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., Chen, Y., *The effect of news and public mood on stock movements*, „Information Sciences”, R. 2014, nr 278, s. 826–840.
214. Li, X., Xie, H., Chen, L., Wang, J., Deng, X., *News impact on stock price return via sentiment analysis*, „Knowledge-Based Systems”, R. 2014, nr 69, s. 14–23.
215. Li, Y., Bu, H., Li, J., Wu, J., *The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning*, „International Journal of Forecasting”, R.2020, nr 36/4, s.1541–1562.
216. Lillo, F., Micciché, S., Tumminello, M., Piilo, J., Mantegna, R. N., *How news affects the trading behaviour of different categories of investors in a financial market*, „Quantitative Finance”, R. 2015, nr 15/2, s. 213–229.
217. Lim, Y.X., Tan, C., *Do negative events really have deteriorating effects on stock performance? A comparative study on Tesla (US) and Nio (China)*, „Journal of Asian Business and Economic Studies” R.2021. nr 29/2. s.105–119.
218. Liu, B., *Sentiment Analysis and Subjectivity*, [w:] (red.) Herlich, R., Graepel, T., *Handbook of natural language processing*, Boca Raton: Taylor & Francis Group, 2010.
219. Loughran, T., McDonald, B., *Barron's Red Flags: Do They Actually Work?*, „Journal of Behavioral Finance”, R. 2011, nr 12/2, s. 90–97.
220. Loughran, T., McDonald, B., *Measuring Readability in Financial Disclosures*, „The Journal of Finance”, R. 2014, nr LXIX/4, s. 1643–1671.
221. Loughran, T., McDonald, B., *Textual Analysis in Accounting and Finance: A Survey*, „Journal of Accounting Research”, R. 2016, nr 54/4, s.1187–1230.

222. Loughran, T., McDonald, B., *When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks*, „The Journal of Finance”, R.2011, nr 66/1, s.35–65.
223. Lu, Y-Ch., Shen, Ch-H., Wei, Y-Ch., *Revisiting early warning signals of corporate credit default using linguistic analysis*, „Pacific-Basin Finance Journal”, R.2013, nr 24, s.1–21.
224. Lutje, T., *To be good or to be better: Asset managers' attitudes towards herding*, „Applied Financial Economics”, R. 2009, nr 19, s. 825–839.
225. Machlarz, A., *O ogólnej definicji pojęcia informacji semantycznej*, „Filozofia Nauki”, R. 2011, nr 19/1, s. 71–93.
226. Magnusson, C., Arppe, A., Eklund, T., Back, B., Vanharanta, H., Visa, A., *The language of quarterly reports as an indicator of change in the company's financial status*, „Information & Management”, R. 2005, nr 42, s. 561–574
227. Malkiel, B. G., *Błądząc po Wall Street. Dlaczego nie można wygrać z rynkiem*, WIG-Press, Warszawa 2003.
228. Manning, P., *Financial journalism, news sources and the banking crisis*, „Journalism”, R. 2012, nr 14/2, s. 173–189.
229. Maqsood, H. i inni, *A local and global event sentiment based efficient stock exchange forecasting using deep learning*, „International Journal of Information Management”, R.2020, nr 50, s.432–451.
230. Markowitz, H. M., *Portfolio Selection*, „The Journal of Finance”, R. 1952, nr 7/1, s. 77–91.
231. McFadden, D., *Conditional logit analysis of qualitative choice behavior*, [w:] (red.) Zarembka, P. *Frontiers in Econometrics*, Academic Press, Nowy Jork 1974.
232. McKee, A., *Textual Analysis: A Beginner's Guide*, SAGE, London 2003.
233. McKinley, K., *Stock Market Efficiency and Insider Trading*, „Issues in Political Economy”, R. 1999, nr 8, s. 1–14.
234. McKinsey Global Institute, *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, 2011.
235. McKinsey Global Institute, *The age of analytics: competing in a data-driven world*, 2016.
236. McQuail, D., *McQuail's Mass Communication Theory*, SAGE Publications, Londyn 2010.
237. McQuail, D., *Teoria komunikowania masowego*, Polskie Wydawnictwo Naukowe, Warszawa 2007.
238. Menkhoff, L., Schmidt, U., Brozynski, R., *The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence*, „European Economic Review”, R. 2006, nr 50, s. 1753–1766.
239. Merrill, G.J., *The Political Content of British Economic*, „Business and Financial Journalism”, Palgrave Macmillan, Cham 2019.

240. Merrill, G.J., *The revolution must wait: Economic, business and financial journalism beyond the 2008 crisis*, „Ethical Space: The International Journal of Communication Ethics”, R. 2012, nr 9/1, s. 41–51.
241. Mielcarz, P., Podgórski, B., Wermczuk, P., *Positive recommendations and abnormal returns on the Warsaw Stock Exchange in 2005–2006*, [w:] (red.) Urbańczyk, E., „The Problems of Company Value Management”, Publishing House Print Group, 2007.
242. Mikołajek-Gocejna, M., Urbaś, T., *Racjonalni inwestorzy zamiast racjonalnych oczekiwań w hipotezie rynku efektywnego*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, R. 2017, nr 481, s. 33–53.
243. Miller, B., *The effects of reporting complexity on small and large investor trading*, „Accounting Review”, R. 2010, nr 85, s. 2107–2143.
244. Miller, M. H., Modigliani, F., *Dividend Policy, Growth, and the Valuation of Shares*, „The Journal of Business”, R. 1961, nr 34/4, s. 411–433.
245. Mittermayer, M-A., *Forecasting Intraday Stock Price Trends with Text Mining Techniques*, 37th Hawaii International Conference on System Sciences, Big Island, HI, USA 2004.
246. Mittermayer, M-A., Knolmayer, G. F., *Text mining systems for market response to news: a survey* Institute of Information Systems, University of Bern, „Working Paper”, R. 2006, nr 184.
247. Mrozowski, M., *Media masowe. Władza, rozrywka, i biznes*, Oficyna Wydawnicza ASPRA-JR, Warszawa 2001.
248. Mukhopadhyay, S., *Opinion mining in management research: the state of the art and the way forward*, „Opsearch”, R. 2018, nr 55, s. 221–250.
249. Mynarski, S., *Elementy teorii systemów i cybernetyki*, Polskie Wydawnictwo Naukowe, Warszawa 1979.
250. Nassirtoussi. A.K., Aghabozorgi, S., Wah, T.Y., Ngo, D. C. L., *Text mining for market prediction: A systematic review*, „Expert Systems with Applications”, R. 2014, nr 41, s. 7653–7670.
251. Nassirtoussi. A.K., Aghabozorgi, S., Wah, T.Y., Ngo, D. C. L., *Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment*, „Expert Systems with Applications”, R. 2015, nr 42, s. 306–324.
252. Nauta, D., *The Meaning of Information*, Mouton & Co, The Hague, Paryż 1972.
253. Nikfarjam, A., Emadzadeh, E., Muthaiyah, S., *Text mining approaches for stock market prediction*, 2010, The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010 Singapore.
254. Nisar, T. M., Yeung, M. S., *Twitter as a tool for forecasting stock market movements: A short-window event study*, „The Journal of Finance and Data Science”, R.2018, nr 4/2, s.101–119.
255. Nizer, P. S. M, Nievola, J. C., *Predicting published news effect in the Brazilian stock market*, „Expert Systems with Applications”, R. 2012, nr 39, s. 10674–10680.

256. Nobelprize.org, *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2002*, <https://www.nobelprize.org/prizes/economic-sciences/2002/summary>. Dostęp: 7 mar. 2023.
257. Nogueira Reis, P. M., Pinho, C., *A new European investor sentiment index (EURsent) and its return and volatility predictability*, „Journal of Behavioral and Experimental Finance”, R. 2020, nr 27, s. 1–13.
258. Nowicka, E., *Świat człowieka – świat kultury*, Polskie Wydawnictwo Naukowe, Warszawa, 2006.
259. O'Connell, P.G.J., Teo, M., *Institutional Investors, Past Performance, and Dynamic Loss Aversion*, „The Journal of Financial and Quantitative Analysis”, R. 2009, nr 44/1, s.155–188.
260. Oliveira, N., Cortez, P., Areal, N., *Stock market sentiment lexicon acquisition using microblogging data and statistical measures*, „Decision Support Systems”, R.2016, nr 85, s.62–73.
261. Oncu, E., *The impact of COVID-19 on health sector stock returns*, „MPRA Paper” R.2021.
262. Oracle Polska, *Czym jest relacyjna baza danych (RDBMS)?*, <https://www.oracle.com/pl/database/what-is-a-relational-database/>. Dostęp: 1 kw. 2023.
263. O'Reilly Media, Inc., *Big Data Now: 2014 Edition*, O'Reilly Media, Inc, Sebastopol 2015.
264. Palczewski, M., *News values w amerykańskich i europejskich teoriach selekcji informacji*, [w:] (red.) Graszewicz, M., Jastrzębski, J., *Teorie komunikacji i mediów*, Oficyna Wydawnicza ATUT, Wrocław 2009.
265. Pan, Z., Kosicki, G.M., *Framing analysis: An approach to news discourse*, „Political Communication”, R. 1993, nr 10, s. 55–75.
266. Parkiet, <https://www.parkiet.com/>, Dostęp: 10 sty. 2024.
267. Pfarrer, M. D., Polloc, T. G., Rindova, V. P., *A tale of two assets: the effects of firm reputation and celebrity on earnings surprises and investors' reactions*, „Academy of Management Journal”, R.2010, nr 53/5, s.1131–1152.
268. Piao, J., *Financial Media, Globalisation, and China's Economic Integration: Comparing Narrative Construction of The Economist and Caijing*, Rozprawa doktorska, University of Westminster, 2015.
269. Picasso, A., Merello, S., Ma, Y., Cambria, E., *Technical analysis and sentiment embeddings for market trend prediction*, „Expert Systems with Applications”, R. 2019, nr 135, s. 60–70.
270. Pixley, J., *Emotions in Finance: Booms, Busts and Uncertainty*, Cambridge University Press, Cambridge 2012.
271. Podstawka, M., *Finanse. Instytucje, instrumenty, podmioty, rynki, regulacje*, Polskie Wydawnictwo Naukowe, Warszawa 2017.

272. Polak, K., *Wpływ opublikowanej rekomendacji na reakcje inwestorów*, „Studia i Materiały Wydziału Zarządzania Uniwersytetu Warszawskiego”, R. 2018, nr 27/2, s. 127–135.
273. Poland A.M. Daily, <https://polandam.pl/>, Dostęp: 10 sty. 2024.
274. Polish News Bulletin, <https://pnb.pl/>, Dostęp: 10 sty. 2024.
275. Powers, D., *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation*, „International Journal of Machine Learning Technology”, R.2011, nr 45293, s.37–63.
276. Purda, L., Skillicorn, D., *Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection*, „Contemporary Accounting Research”, R. 2014, nr 32/3, s. 1193–1223.
277. PwC, *Where have you been all my life? How the financial services industry can unlock the value in Big Data*, 2013.
278. Rasheed, H., Ahmad, H., Javid, A. Y., *Gold vs. PSX sectors during political uncertainties: an event study analysis*, „Business Review”, R.2022, nr 16/2, s.1–20.
279. Ravisankar, P., Ravi, V., Raghava Rao, G., Bose, I., *Detection of financial statement fraud and feature selection using data mining techniques*, „Decision Support Systems Journal”, R. 2011, nr 50, s. 491–492.
280. Reboredo, J. C., Ugolini, A., *The impact of Twitter sentiment on renewable energy stocks*, „Energy Economics”, R.2018, nr 76, s.153-169.
281. Rocca, L., Giacomini, D., Zola, P., *Environmental disclosure and sentiment analysis: state of the art and opportunities for public-sector organisations*, „Meditari Accountancy Research”, R.2021, nr 29/3, s.617–646.
282. Rochester, J.B., *Using Computers and Information*, Education and Training, Indianapolis 1996.
283. Rostek, K., Młodzianowski, P., *Współzależność informacji sieciowych oraz zmian indeksów zachodzących na Gieldzie Papierów Wartościowych w Warszawie*, „Zeszyty Naukowe Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach”, R. 2017, nr 115, s. 249–263.
284. Rouhani, S., Abedin, E., *Crypto-currencies narrated on tweets: a sentiment analysis approach*, „International Journal of Ethics and Systems”, R.2020, nr 36/1, s.58–72.
285. Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 596/2014 z dnia 16 kwietnia 2014 r. w sprawie nadużyć na rynku (rozporządzenie w sprawie nadużyć na rynku) oraz uchylające dyrektywę 2003/6/WE Parlamentu Europejskiego i Rady i dyrektywy Komisji 2003/124/WE, 2003/125/WE i 2004/72/WE. Dziennik Urzędowy Unii Europejskiej L 173/1.
286. Samuelson, P.A., Nordhaus, W.D., *Ekonomia*, Dom Wydawniczy Rebis, Poznań 2017.
287. Samunderu, E., Yordanova, A., *The Effect of Public Contract Winning Announcements on Share Prices: An Event-Based Study on the Pharmaceutical Industry*, „Journal of Economic Analysis”, R.2024, nr 3/1, s.1–26.

288. Saurabh S., Dey K., *Unraveling the relationship between social moods and the stock market: Evidence from the United Kingdom*, „Journal of Behavioral and Experimental Finance”, R.2020, nr 26, s.1–9.
289. Schumaker, R. P., Chen, H., *Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFinText System*, „ACM Transactions on Information Systems”, R. 2009, nr 27/2, s. 1–19.
290. Schumaker, R. P., Zhang, Y., Huang, Ch-N., Chen, H., *Evaluating sentiment in financial news articles*, „Decision Support Systems”, R. 2012, nr 53, s. 458–464.
291. Schumaker, R.P., Chen, H., *A Quantitative Stock Prediction System Based on Financial News*, „Information Processing and Management”, R. 2009, nr 45/5, s. 571–583.
292. Schuster, T., *News Events and Price Movements. Price Effects of Economic and Non-Economic Publications in the News Media*, „Leipzig University Working Paper”, R. 2003, nr 03–02, s. 1–34.
293. Schuster, T., *News Events and Price Movements. Price Effects of Economic and Non-Economic Publications in the News Media*, „SSRN Electronic Journal”, R. 2003, s. 3–34.
294. Sekuła, P., *Empiryczny test strategii fundamentalnej*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, R. 2012, nr 271, s. 280–289.
295. Shaharuddin, N. i inni, *The Role of English Proficiency and Financial Development in Stimulating Economic Growth*, „Journal of Entrepreneurship, Business and Economics” R.2022. nr 10/2S2. s.18–37.
296. Shannon, C.E., *A Mathematical Theory of Communication*, „The Bell System Technical Journal”, R. 1948, nr 27/3, s. 379–423.
297. Sharpe, W.F., *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, „The Journal of Finance”, R. 1964, nr 19/3, s. 425–442.
298. Shefrin, H., *A Behavioral Approach to Asset Pricing*, Elsevier Inc, Burlington 2008.
299. Shefrin, H., *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing*, Oxford University Press, Oxford 2002.
300. Shefrin, H., Statman, M., *The disposition to sell winners too early and ride losers too long, Theory and evidence*, „The Journal of Finance”, R. 1985, nr XL/3, s. 777–790.
301. Shefrin, H., *The behavioral Paradigm Shift*, „RAE-Revista de Administração de Empresas”, R. 2015, nr 55/1, s. 95–98.
302. Shiller, R.J., *Irrational Exuberance*, Princeton University Press, New Jersey 2015.
303. Shoemaker, P.J., *Gatekeeping. Communication Concepts*, Sage Publications, Londyn 1991.
304. Shoemaker, P.J., Vos, T., *Gatekeeping Theory*, Routledge, Nowy Jork 2009.
305. Sias, W.R., *Institutional Herding*, „Review of Financial Studies”, R. 2004, nr 17/1, s. 165–206.

306. Siganos, A., Vagenas-Nanos, E., Verwijmeren, P., *Facebook's daily sentiment and international stock markets*, „Journal of Economic Behavior & Organization”, R. 2014, nr 107/B, s. 730–743.
307. Simon, H., *A Behavioral Model of Rational Choice*, „The Quarterly Journal of Economics”, R. 1955, nr 69/1, s. 99–118.
308. Sinayobye, J. O., Kiwanuka, F., Kaawaase Kyanda S., *A state-of-the-art review of machine learning techniques for fraud detection research*, ACM/IEEE Symposium on Software Engineering in Africa 2018, Gothenburg, Szwecja.
309. Smailovic, J., Grear, M., Lavrac, N., Znidarsic, M., *Stream-based active learning for sentiment analysis in the financial domain*, „Information Sciences”, R. 2014, nr 285, s. 181–203.
310. Snarska, M., *Efektywność informacyjna giełdowych rynków akcji*, [w:] (red.) Czekaj, J., *Efektywność giełdowego rynku akcji w Polsce z perspektywy dwudziestolecia*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2014.
311. Sobczyk, M., *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2007.
312. Song, M., Shin, K-S., *Forecasting economic indicators using a consumer sentiment index: Survey-based versus text-based data*, „Journal of Forecasting”, R.2019, nr 38, s.504–518.
313. Starzeński, O., *Analiza rynków finansowych*, Wydawnictwo C.H. Beck, Warszawa 2011.
314. Statman, M., *Behavioral Finance: Finance with Normal People*, „Borsa Istanbul Review”, R. 2014, nr 14/2, s. 65–73.
315. Stiglitz, J. E., Grossman, S.J., *On the Impossibility of Informationally Efficient Markets*, „The American Economic Review”, R. 1980, nr 70,3. s. 393–408.
316. Stiglitz, J. E., *Informacja i zmiana paradygmatu w ekonomii*, „Gospodarka Narodowa”, R. 2004, nr 3, s. 95–114.
317. Storkenmaier, A., *Financial Markets and Public Information*, KIT Scientific Publishing, Karlsruhe 2011.
318. Stowarzyszenie Inwestorów Indywidualnych, *Oto najpopularniejsze źródła informacji dla inwestorów [OBI 2022]*, <https://www.sii.org.pl/15945/aktualnosci/badania-i-rankingi/oto-najpopularniejsze-zrodla-informacji-dla-inwestorow-obi-2022.html>. Dostęp: 8 mar. 2023.
319. Strauß, N., *Financial journalism in today's high-frequency news and information era*, „Journalism”, R. 2019, nr 20/2, s. 274–291.
320. Strauss, N., Smith, Ch. H., *Buying on rumors: how financial news flows affect the share price of Tesla*, „Corporate Communications: An International Journal”, R.2019, nr 24/4, s.593–607.
321. Strauß, N., Vliegthart, R., Verhoeven, P., *Lagging behind? Emotions in newspaper articles and stock market prices in the Netherlands*, „Public Relations Review”, R. 2016, nr 42, s. 548–555.

322. Suto, M., Toshino, M., *Behavioural biases of Japanese institutional investors: Fund management and corporate governance*, „Corporate Governance: An International Review”, R. 2005, nr 13/4, s. 466–477.
323. Szczepański, R., *Semantyczne Teorie Informacji*, „Studia Metodologiczne”, R. 2015, nr 34, s. 61–81.
324. Szyszka, A., *Behavioral Finance and Capital Markets: How Psychology Influences Investors and Corporations*, Palgrave Macmillan, Nowy Jork 2013.
325. Szyszka, A., *Finanse behawioralne. Nowe podejście do inwestowania na rynku kapitałowym*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2009.
326. Szyszka, A., *Zjawisko Kontynuacji Stóp Zwrotu Na Giełdzie Papierów Wartościowych w Warszawie*, „Bank i Kredyt”, R. 2005, nr 8, s. 37–49.
327. Tambini, D., *What are financial journalists for?*, „Journalism Studies”, R. 2010, nr 11/2, s. 158–174.
328. Teo, M., O'Connell, P.G. J., *Prospect Theory and Institutional Investors*, „SSRN Electronic Journal”, R. 2005, s. 1–39.
329. Tetlock, P. C., *Giving Content to Investor Sentiment: The Role of Media in the Stock Market Published*, „The Journal of Finance”, R. 2007, nr LXII/3, s. 1139–1168.
330. Thaler, R., *Judgment and Decision Making under Uncertainty: What Economists Can Learn from Psychology*, „Risk Analysis in Agriculture Research and Educational Developments”, January 16–18, 1980, Tucson, Arizona.
331. Thompson, P. A., *Market Manipulation? Applying the Propaganda Model to Financial Media Reporting*, „Westminster Papers in Communication and Culture”, R. 2009, nr 6/2, s. 73–96.
332. Thomson Reuters, *Starmine Quantitative Models*, http://professional.starmine.com/static_content/StarMine_Text_Mining_Credit_Risk_Model_overview.pdf. Dostęp: 20 sie. 2017.
333. Throckmorton, Ch. S., Mayew, W. J., Venkatachalam, M., Collins, L. M., *Financial fraud detection using vocal, linguistic and financial cues*, „Decision Support Systems Journal”, R. 2015, nr 74, s. 78–87.
334. Timo Korkeamäki, Tuomas Takalo, *Valuation of Innovation and Intellectual Property: The Case of iPhone*, „European Management Review”, R.2013, nr 10/4, s.197–210.
335. Tsai, C-W., Lai, C-F., Chao, H-C., Vasilakos, A. V., *Big data analytics: a survey*, „Journal of Big Data”, R. 2015, nr 2/21, s. 1–32.
336. Tsai, F-T, Lu, H-M., Hung, M-W., *The impact of news articles and corporate disclosure on credit risk valuation*, „Journal of Banking & Finance”, R.2016, nr 68, s.100–116.
337. Tsukioka, Y., Yanagi, J., Takada, T., *Investor sentiment extracted from internet stock message boards and IPO puzzles*, „International Review of Economics and Finance”, R.2018, nr 56, s.205–217.
338. Tutorials Point, *Data Mining data pattern evaluation*, 2017.

339. Tversky, A., Kahneman, D., *Advances in Prospect-Theory – Cumulative Representation of Uncertainty*, „Journal of Risk and Uncertainty”, R. 1992, nr 5/4, s. 297–323.
340. Tversky, A., Kahneman, D., *The Framing of Decisions and the Psychology of Choice*, „Science”, R. 1981, nr 211, s. 453–458.
341. Tymiakin, L., *O triadzie komunikacyjnej: wartościowanie – emocje – ekspresja (On the Communication Trifecta: Valuating – Emotions – Expression)*, „Annales Universitatis Mariae Curie-Skłodowska Lublin – Polonia”, R. 2017, nr II, s. 199–216.
342. Umar, M., i inni, *Crypto swings and the performance of carbon-intensive equity funds in China*, „Resources Policy” R.2022. nr 78. s.1-6.
343. University of Oxford, *Reuters Institute for the Study of Journalism, Media Coverage of Banking and Financial News*, 2014.
344. Usher, N., *Making business news: A production analysis of The New York Times*, „International Journal of Communication”, R. 2017, nr 11, s. 363–382.
345. Ustawa z dnia 26 stycznia 1984 r. – *Prawo prasowe*. Dziennik Ustaw 1984 Nr 5 poz. 24.
346. Ustawa z dnia 29 lipca 2005 r. – *O obrocie instrumentami finansowymi*. Dziennik Ustaw 2005 Nr 183 poz. 1538.
347. Van de Kauter, M., Breesch, D., Hoste, V., *Fine-grained analysis of explicit and implicit sentiment in financial news articles*, „Expert Systems with Applications”, R. 2015, nr 42, s. 4999–5010.
348. Verma, R., Soydemir, G., *The Impact of U.S. Individual and Institutional Investor Sentiment on Foreign Stock Markets*, „Journal of Behavioral Finance”, R. 2006, nr 7/3, s. 128–144.
349. Vijayarani, S., Ilamathi, J., Nithya, S., *Preprocessing Techniques for Text Mining – An Overview*, „International Journal of Computer Science & Communication Networks”, R. 2015, nr 5/1, s. 7–16.
350. Wang, B., Huang, H., Wang, X., *A novel text mining approach to financial time series forecasting*, „Neurocomputing”, R. 2012, nr 83, s. 136–145.
351. Wang, J., *A Model of Competitive Stock Trading Volume*, „Journal of Political Economy”, R. 1994, nr 102/1, s. 127–68.
352. Wang, J-Z., *Forecasting stock indices with back propagation neural network*, „Expert Systems with Applications”, R. 2011, nr 38, s. 14346–14355.
353. Waszkowski, A., *Hipoteza efektywności rynku-weryfikacja dla indeksu WIG-Spożywczy*, „Problemy Rolnictwa Światowego”, R. 2011, nr 11/26/4, s. 169–176.
354. West, J., Bhattacharya, M., *Intelligent financial fraud detection: A comprehensive review*, „Computers and Security”, R. 2016, nr 57.
355. Wędrowska, E., *Ilościowe i Jakościowe Koncepcje Informacji*, „Ekonomiczne Problemy Usług”, R.2010, nr 57, s. 49–56.
356. Whittlesey, Ch. R., *The Banking System and War Finance*, National Bureau of Economic Research, Cambridge, 1943.

357. Wiseman, R.M., Gomez-Mejia, L.R., *A behavioral agency model of management risk taking*, „Academy of Management Review”, R. 1998, nr 23/1, s. 133–153.
358. Witkowska, D., Żebrowska-Sychodolska, D., *Badanie słabej formy efektywności informacyjnej GPW*, „Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania”, R. 2008, nr 9, s. 155–165.
359. Woleński, J., *Informacja i semantyka*, „Filozofia Nauki”, R. 1997, nr 5/1, s. 59–64.
360. Wójtowicz, T., *Efekt momentum na GPW w Warszawie w latach 2003–2010*, „Ekonomia Menedżerska”, R. 2011, nr 9, s. 63–74.
361. Wujec, M., *Analysis of the Financial Information Contained in the Texts of Current Reports: A Deep Learning Approach*, „Journal of Risk and Financial Management”, R.2021, nr 14/582, s.1–17.
362. Wuthrich, B., Cho, V., Leung, S., Permuntilleke, D., Sankaran, K., Zhang, J., Lam, W., *Daily stock market forecast from textual web data*, „Conference Paper”, San Diego CA, USA 1998.
363. Xing, F. Z., Cambria, E., Zhang, Y., *Sentiment-aware volatility forecasting*, „Knowledge-Based Systems”, R.2019, nr 176, s.68–76.
364. Yoshimori, M., *Shadow Exchange Rates – Changing the Winds with Headwinds and Tailwinds*, „SocioEconomic Challenges”, R.2019, nr 3, s.78–88.
365. You, H., Zhang, X., *Financial reporting complexity and investor underreaction to 10-K information*, „Review of Accounting Studies”, R. 2009, nr 14/4, s. 559–586.
366. Yu, J-R., i inni, *Dynamic rebalancing portfolio models with analyses of investor sentiment*, „International Review of Economics and Finance”, R.2022, nr 77, s.1–13.
367. Yu, L.C., Wu, J.L., Chang, P.C., Chu, H.S., *Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news*, „Knowledge-Based Systems”, R.2013, nr 41, s.89–97.
368. Yu, Y., Duan, W., Cao, Q., *The impact of social and conventional media on firm equity value: A sentiment analysis approach*, „Decision Support Systems”, R.2013, nr 55/4, s.919–926.
369. Zielonka, P., *Giełda i psychologia. Behawioralne aspekty inwestowania na rynku papierów wartościowych*, CeDeWu, Warszawa 2017.
370. Zikopoulos P.C., Eaton, Ch., deRoos, D., Deutsch, T., Lapis, G., *Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data*, McGraw-Hill, Nowy Jork 2011.
371. Zubair, S., Cios, K. J., *Extracting News Sentiment and Establishing its Relationship with the S&P500 Index*, 48th Hawaii International Conference on System Sciences Extracting, 2015 Kauai, HI, USA.
372. Żebrowska-Suchodolska, D., *Weryfikacja słabej formy efektywności za pomocą testów statystycznych na GPW w Warszawie*, „Optimum Studia Ekonomiczne”, R. 2009, nr 1/41, s. 55–68.
373. Żurek, M., *Inklinacje behawioralne na rynkach kapitałowych w świetle modeli SEM*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2016.

Spis tabel

Tabela 1. Stosunek do ryzyka. Teoria perspektywy.....	69
Tabela 2. Wybrane błędy poznawcze i motywacyjne inwestorów giełdowych.....	74
Tabela 3. Przykładowe klasyfikacje emocji na gruncie literatury światowej w kontekście wykorzystania metod eksploracji tekstu w obszarze finansów i ekonomii	140
Tabela 4. Zdefiniowanie sentymentu na gruncie literatury światowej w kontekście wykorzystania metod eksploracji tekstu w obszarze finansów i ekonomii, jako stopnia pozytywności i negatywności treści	142
Tabela 5. Wybrane badania, wykorzystujące oficjalne dokumenty spółek jako źródło danych w obszarze zastosowania analizy tekstu w finansach.....	151
Tabela 6. Wybrane badania, wykorzystujące publikacje prasowe jako źródło danych w obszarze zastosowania analizy tekstu w finansach.....	153
Tabela 7. Wybrane przykłady badań, wykorzystujących słownik Loughran-McDonald Master Dictionary	187
Tabela 8. Zastosowanie analizy zdarzeń w ramach badań związanych z analizą tekstu w kontekście rynków finansowych	189
Tabela 9. Zastosowanie miar uśredniających do prezentacji wyników analizy zdarzeń dla pojedynczych instrumentów.....	192
Tabela 10. Zastosowanie modeli klasyfikacyjnych, w tym regresji logitowej, do badań nad reakcją rynków finansowych i ich podmiotów na otrzymanie informacji niefinansowych.....	197
Tabela 11. Spółki zakwalifikowane do próby badawczej.....	201
Tabela 12. Spółki niewchodzące w skład próby badawczej	202
Tabela 13. Liczba artykułów przypisanych do poszczególnych klas sentymentu.....	207
Tabela 14. Wartości CAAR oszacowane w wyniku reakcji na artykuły prasowe, charakteryzujące się negatywnym sentymentem.....	211
Tabela 15. Wartości CAAR oszacowane w wyniku reakcji na artykuły prasowe, charakteryzujące się pozytywnym sentymentem	212
Tabela 16. Wartości CAAR dla spółek, które w oknie zdarzenia, nie wykazały istotnych statystycznie stóp zwrotu reakcji na sentyment zawarty w artykułach prasowych	213
Tabela 17. Wyniki estymacji modelu logitowego.....	215
Tabela 18. Wyniki klasyfikacji dziennych stóp zwrotu dla spółek w próbie badawczej	216
Tabela 19. Wyniki współczynników klasyfikacji	218

Spis wykresów

Wykres 1. Osoby korzystające z Internetu (% populacji).....	118
Wykres 2. Dynamika miesięcznych stóp zwrotu indeksów WIG20 i WIG w okresie 01.2013 – 12.2022	200
Wykres 3. Liczba zgromadzonych artykułów dla spółek, znajdujących się w próbie badawczej.....	205
Wykres 4. Struktura artykułów zaklasyfikowanych jako pozytywne lub negatywne	206
Wykres 5. Struktura sentymentu artykułów, wykorzystanych w analizie zdarzeń.....	209

Spis rysunków

Rysunek 1. Typy informacji na rynku finansowym.....	36
Rysunek 2. Funkcja użyteczności, a stosunek inwestora do ryzyka.....	52
Rysunek 3. Krzywa wartości dla zysków i strat teorii perspektywy. Teoria perspektywy	66
Rysunek 4. Funkcja wag decyzyjnych	68
Rysunek 5. Eksploracja danych jako element procesu odkrywania wiedzy.....	126
Rysunek 6. Sentyment tekstowy w procesie kształtowania sentymentu rynkowego	137
Rysunek 7. Procedura badawcza	180
Rysunek 8. Udział inwestorów krajowych i zagranicznych w obrocie Rynku Głównego GPW w Warszawie.....	183
Rysunek 9. Ogólny podział metod analizy tekstu.....	185

Załączniki

Załącznik 1. Statystyki opisowe dziennych stóp zwrotu spółek uwzględnionych w próbie badawczej

Spółka	Średnia	Błąd standardowy	Mediana	Odchylenie standardowe	Wariancja próbki	Kurtoza	Skośność	Zakres	Minimum	Maksimum	Suma	Licznik
PZU	0,0003	0,0003	0,0002	0,0169	0,0003	2,6951	-0,0378	0,1971	-0,1113	0,0857	0,8089	2497
ACP	0,0005	0,0003	0,0000	0,0171	0,0003	2,1474	0,2476	0,1741	-0,0727	0,1013	1,3701	2497
CPS	0,0001	0,0004	0,0000	0,0182	0,0003	1,4655	0,0263	0,1680	-0,0898	0,0782	0,1713	1946
BZW	0,0004	0,0005	0,0000	0,0195	0,0004	1,0224	0,2546	0,1457	-0,0692	0,0766	0,5612	1377
OPL	0,0001	0,0004	0,0000	0,0195	0,0004	6,4658	0,5502	0,2739	-0,1176	0,1563	0,1465	2194
PKO	0,0002	0,0004	-0,0002	0,0196	0,0004	5,5646	-0,0895	0,3009	-0,1682	0,1327	0,4673	2497
PEO	0,0001	0,0004	0,0000	0,0199	0,0004	9,7291	-0,3226	0,3673	-0,1959	0,1714	0,3465	2497
BRE	0,0018	0,0012	0,0001	0,0201	0,0004	2,0139	0,4152	0,1530	-0,0556	0,0974	0,5469	303
GTC	-0,0002	0,0011	0,0000	0,0210	0,0004	1,9251	0,2296	0,1672	-0,0813	0,0859	-0,0794	363
ENG	-0,0006	0,0007	0,0000	0,0212	0,0005	1,8257	0,0565	0,1988	-0,1067	0,0921	-0,6372	997
PGN	0,0002	0,0004	0,0000	0,0214	0,0005	5,0407	0,5866	0,2757	-0,0961	0,1797	0,5373	2427
PKN	0,0005	0,0004	0,0000	0,0215	0,0005	1,8253	0,1264	0,2402	-0,0963	0,1439	1,1728	2497
ENA	-0,0003	0,0010	-0,0008	0,0218	0,0005	2,2710	0,3824	0,1889	-0,0884	0,1004	-0,1696	501
LTS	0,0006	0,0005	-0,0002	0,0221	0,0005	2,7689	0,2379	0,2646	-0,1370	0,1277	1,3439	2392
Spółka	Średnia	Błąd standardowy	Mediana	Odchylenie standardowe	Wariancja próbki	Kurtoza	Skośność	Zakres	Minimum	Maksimum	Suma	Licznik

BHW	0,0006	0,0013	0,0000	0,0224	0,0005	0,9413	-0,1435	0,1510	-0,0841	0,0669	0,1918	302
EUR	-0,0002	0,0007	0,0000	0,0230	0,0005	2,0384	-0,2243	0,1758	-0,0992	0,0766	-0,2382	997
ALR	-0,0003	0,0006	-0,0010	0,0244	0,0006	7,8502	0,1634	0,3581	-0,1988	0,1593	-0,5523	1746
KGH	0,0002	0,0005	0,0000	0,0247	0,0006	2,5372	0,0383	0,2767	-0,1599	0,1168	0,5802	2497
TPE	0,0000	0,0005	0,0000	0,0247	0,0006	5,8902	0,8405	0,3363	-0,1365	0,1998	0,0706	2497
DNP	0,0016	0,0008	0,0006	0,0250	0,0006	3,6757	0,1475	0,2558	-0,1379	0,1179	1,5094	949
PGE	0,0000	0,0005	-0,0009	0,0258	0,0007	7,3699	0,8241	0,3877	-0,1912	0,1965	0,0248	2497
TPS	0,0001	0,0015	0,0000	0,0259	0,0007	45,1617	-4,3593	0,3484	-0,2792	0,0692	0,0319	303
LWB	-0,0014	0,0010	-0,0008	0,0260	0,0007	35,1546	-0,8530	0,5378	-0,2967	0,2412	-1,0044	740
MBK	0,0001	0,0006	-0,0005	0,0262	0,0007	7,3242	0,6235	0,4042	-0,1633	0,2409	0,2375	2194
SPL	0,0001	0,0008	-0,0009	0,0264	0,0007	5,9725	0,1508	0,3378	-0,1750	0,1629	0,0790	1004
LPP	0,0005	0,0006	0,0000	0,0264	0,0007	8,4470	0,0486	0,4122	-0,2519	0,1602	1,0884	2195
KER	-0,0005	0,0012	0,0000	0,0286	0,0008	6,0567	-0,3656	0,3269	-0,2056	0,1214	-0,3058	611
PLY	0,0015	0,0014	0,0000	0,0294	0,0009	56,2549	4,5616	0,5022	-0,1350	0,3673	0,6826	441
CDR	0,0007	0,0009	0,0018	0,0312	0,0010	2,9570	-0,1182	0,3196	-0,1638	0,1558	0,7988	1196
CCC	-0,0001	0,0008	-0,0007	0,0325	0,0011	12,6051	0,8330	0,5641	-0,2364	0,3277	-0,1981	1756
ALE	-0,0019	0,0015	-0,0023	0,0327	0,0011	2,2883	0,3646	0,2954	-0,1377	0,1578	-0,9592	508
Spółka	Średnia	Błąd standardowy	Mediana	Odchylenie standardowe	Wariancja próbki	Kurtoza	Skośność	Zakres	Minimum	Maksimum	Suma	Licznik

JSW	0,0002	0,0008	-0,0014	0,0359	0,0013	8,6331	1,2536	0,4980	-0,1585	0,3395	0,5386	2248
SNS	0,0003	0,0014	0,0000	0,0393	0,0015	8,3512	0,6461	0,4167	-0,1667	0,2500	0,2169	801

Załącznik 2. Udział badanych spółek w indeksie WIG20 w latach 2013 – 2022 [%].

rok	2022				2021				2020			
rewizja	12	9	6	3	12	9	6	3	12	9	6	3
ACP	1,98	2,23	2,16	1,88	2,12	1,84	1,70	1,73	1,97			
ALE	5,16	5,88	5,09	5,31	7,19	10,24	9,52	9,81	12,38			
ALR									0,77	0,71	0,78	1,18
BHW												
BRE												
BZW												
CCC	0,72	0,81	0,86	0,94	1,63	1,77	1,95	1,81	1,13	1,24	0,88	1,30
CDR	5,10	3,23	4,17	5,19	5,48	4,77	5,27	8,84	12,80	15,00	15,00	11,15
CPS	2,18	2,54	2,75	3,20	3,81	3,83	3,36	3,67	3,74	4,52	4,41	3,94
DNP	8,62	8,73	7,20	5,85	6,61	6,27	5,70	5,65	6,66	6,54	5,46	4,00
ENA												
ENG												
EUR												
GTC												
JSW	1,23	1,43	1,75	0,99	0,89	0,74	0,72	0,99	0,56	0,46	0,45	0,42
KER												
KGH	7,84	7,14	9,10	8,79	8,00	9,59	11,59	13,30	9,93	10,46	7,09	6,20
LPP	4,88	5,22	5,33	6,60	5,60	5,54	4,63	4,01	3,88	5,42	5,28	5,36
LTS			2,88	1,98	2,03	2,01	1,89	1,69	1,47	2,07	3,30	3,24
LWB												
MBK	2,11	1,64	1,70	2,49						1,43	1,55	2,44
OPL	2,15	2,09	1,94	2,26	2,20	2,16	1,86	1,84	2,11	2,80	2,70	2,41
PEO	7,45	6,54	8,03	9,52	8,64	7,35	6,94	5,57	5,14	5,36	5,81	8,90
PGE	2,84	3,97	3,85	2,53	3,04	3,01	3,13	2,48	2,47	3,08	2,17	2,26
PGN		5,23	4,91	3,65	4,07	4,17	4,53	4,48	3,97	4,94	3,97	2,95
PKN	15,00	14,18	10,12	8,64	9,16	8,43	9,43	8,24	7,39	8,78	11,75	10,60
PKO	12,13	10,92	13,00	15,00	15,00	13,59	13,62	12,35	11,75	10,85	11,79	14,76
PLY										2,70	2,91	1,99
PZU	8,48	8,85	8,61	8,26	8,37	9,25	8,64	8,16	7,70	9,17	10,71	11,28
SNS												
SPL	4,57	4,13	4,12	4,87	4,70	3,68	3,66	3,39	3,03	2,96	3,22	4,93
TPE					1,29	1,47	1,42	1,32	1,15	1,51	0,78	0,70
TPS												
Suma	92	95	98	98	100	100	100	99	100	100	100	100

rok	2019				2018				2017			
rewizja	12	9	6	3	12	9	6	3	12	9	6	3
ACP									1,48	1,54	1,82	2,08
ALE												
ALR	1,31	1,90	2,20	2,32	2,24	2,84	3,15	3,32	2,88	2,54	3,13	2,98
BHW												
BRE												
BZW						5,62	5,56	5,30	5,39	4,98	5,00	5,44
CCC	1,53	1,88	2,17	2,21	2,57	3,00	3,68	3,41	3,37	3,04	2,69	2,72
CDR	8,71	8,31	7,14	5,57	4,62	6,60	4,64	3,38				
CPS	3,83	3,90	3,32	3,11	3,00	2,91	3,31	2,86	2,44	2,66	2,68	2,52
DNP	3,12	3,62	2,94	2,03								
ENA												
ENG					0,91	0,80	0,99	1,13	1,09	1,28	0,95	1,03
EUR					0,65	0,65	0,90	0,92	1,00	1,23	1,17	1,46
GTC												
JSW	0,53	0,86	1,34	1,50	1,87	1,85	2,20	2,38	2,20	2,24	1,82	1,82
KER												
KGH	5,99	5,30	6,24	5,90	6,00	5,70	6,28	6,80	7,02	7,58	7,45	8,81
LPP	5,54	4,82	4,58	4,59	5,09	5,56	5,78	5,57	4,84	4,66	3,90	3,54
LTS	3,67	3,57	3,46	3,70	3,24	2,89	2,42	2,25	2,47	2,07	2,31	1,98
LWB												
MBK	2,46	2,17	2,46	2,59	2,55	2,47	2,74	2,80	2,75	2,53	2,63	2,66
OPL	2,09	1,99	1,53	1,60	1,55	1,49	1,60	1,73	1,62	1,71	1,43	1,58
PEO	8,12	8,01	8,07	8,01	8,16	8,89	9,06	9,34	9,06	9,16	10,60	11,40
PGE	3,30	3,06	3,50	4,35	4,65	3,29	4,03	4,00	4,39	5,25	4,37	4,69
PGN	3,67	3,92	4,37	5,08	4,98	4,19	4,85	4,64	4,55	4,98	4,92	4,74
PKN	13,23	13,34	12,65	13,55	14,43	13,51	11,95	12,34	15,00	14,67	15,00	13,51
PKO	15,00	15,00	15,00	15,00	15,00	15,00	15,00	15,00	15,00	14,10	14,07	14,78
PLY	1,77	1,40	1,12	1,05								
PZU	10,48	10,94	11,05	10,99	11,47	11,73	10,79	11,48	11,97	11,95	12,43	10,70
SNS												
SPL	4,76	5,23	6,01	5,75	5,92							
TPE	0,89	0,79	0,84	1,10	1,11	1,00	1,09	1,35	1,48	1,83	1,65	1,57
TPS												
Suma	100	100	100	100	100	100	100	100	100	100	100	100

rok	2016				2015				2014			
rewizja	12	9	6	3	12	9	6	3	12	9	6	3
ACP	2,44	2,56	2,54	2,64	2,37	2,24	2,09	2,05	1,87	1,70	1,80	1,90
ALE												
ALR	2,46	2,96	1,75	2,10	2,44	2,44	2,20	1,97	1,72	1,78	1,90	2,01
BHW												
BRE												
BZW	5,71	5,75	5,02	5,32	5,23	4,75	5,20	5,21	5,78	5,50	5,55	6,11
CCC	3,03	2,93	2,63	2,08	2,37							
CDR												
CPS	3,25	3,15	2,99	3,01	2,95	2,57	2,48	2,56				
DNP												
ENA	1,24	1,33	1,37	1,45	1,52	1,64	1,61	1,75				
ENG	0,90	1,02	1,26	1,58	1,83	2,12	2,31	2,10				
EUR	1,84	2,22	2,45	2,66	2,25	1,76	1,32	1,43	1,29	1,49	1,51	1,76
GTC												
JSW									0,76	1,19	1,17	1,34
KER									0,64	0,77	0,73	1,03
KGH	7,05	6,48	5,33	5,61	6,78	6,73	7,94	7,14	8,84	9,56	7,95	8,24
LPP	4,21	3,67	4,14	4,04	5,16	5,12	4,51	4,77	6,41	3,96	3,88	4,51
LTS	1,83	1,54	1,70	1,36					0,79	1,20	1,26	1,18
LWB						0,92	1,38	1,65	1,90	2,13	2,24	2,34
MBK	2,75	2,88	2,46	2,71	2,59	2,53	2,76	3,00	3,20	3,28	3,38	3,61
OPL	2,13	2,08	2,45	2,47	2,54	2,73	3,04	2,88	3,26	3,64	3,56	3,75
PEO	11,87	12,16	12,14	12,25	10,88	10,83	11,31	11,65	11,51	11,87	13,52	13,37
PGE	4,80	5,78	5,90	6,46	6,19	7,16	7,44	7,52	8,58	8,83	7,97	6,48
PGN	4,79	5,84	5,79	5,23	6,18	5,32	4,87	3,55	4,09	4,30	4,08	4,16
PKN	14,11	12,37	13,42	12,26	10,76	12,26	9,78	8,45	6,49	6,26	7,48	6,63
PKO	13,87	13,51	12,68	13,34	13,58	13,02	14,26	14,64	15,00	15,00	15,00	15,00
PLY												
PZU	10,08	9,89	12,31	11,78	11,63	12,55	12,11	14,04	14,10	13,44	12,79	12,73
SNS					1,01	1,22	1,08	1,06	1,04	1,22	1,26	1,37
SPL												
TPE	1,65	1,88	1,67	1,66	1,75	2,06	2,32	2,59	2,74	2,89	2,97	2,46
TPS												
Suma	100	100	100	100	100	100	100	100	100	100	100	100

rok	2013				2012
rewizja	12	9	6	3	12
ACP	2,00	1,97	1,85	1,90	1,78
ALE					
ALR					
BHW	2,06	1,82	1,77	1,76	1,95
BRE	3,47	3,19	2,54	2,33	2,47
BZW	5,58	5,26	4,24		
CCC					
CDR					
CPS					
DNP					
ENA					
ENG					
EUR	1,95	2,67	2,53	2,19	
GTC	0,98	1,08	1,04	1,12	1,22
JSW	1,86	2,03	1,81	1,99	2,10
KER	1,08	1,42	1,62	1,93	2,11
KGH	9,01	8,73	11,61	14,49	14,08
LPP					
LTS	1,22	1,16	1,33	1,37	1,31
LWB	2,06	2,07	2,29	2,61	2,62
MBK					
OPL					
PEO	13,41	12,39	11,41	11,21	10,52
PGE	6,80	6,22	6,72	6,83	7,91
PGN	4,90	5,82	5,00	5,03	4,13
PKN	7,15	7,77	8,73	8,54	8,71
PKO	15,00	15,00	15,00	15,00	15,00
PLY					
PZU	13,92	14,64	14,01	12,82	13,40
SNS	1,36	1,26	1,35	1,47	1,69
SPL					
TPE	2,78	2,65	2,53	2,77	2,95
TPS	3,43	2,84	2,62	4,37	5,01
Suma	100	100	100	100	99