

ЭМОЦИОНАЛЬНЫЙ ТОН РОССИЙСКОГО, ИТАЛЬЯНСКОГО, НЕМЕЦКОГО И ФРАНЦУЗСКОГО НОВОСТНОГО ИНТЕРНЕТ-КОНТЕНТА В ПЕРИОД РАЗВОРАЧИВАНИЯ ПАНДЕМИИ COVID-19

А.С. ПАНФИЛОВА^a, Д.В. УШАКОВ^{a,b}

^a ФГБУН Институт психологии РАН, 129366, Москва, ул. Ярославская, д. 13, к. 1

^b Московский государственный университет имени М.В. Ломоносова, 119991, Россия, Москва, Ленинские горы, 1

Sentiment Analysis of Russian, Italian, German and French Internet News Content during the Spread of the Coronavirus Pandemic

A.S. Panfilova^a, D.V. Ushakov^{a,b}

^a Institute of Psychology, Russian Academy of Sciences, 13 build. 1, Yaroslavskaia Str., Moscow, 129366, Russian Federation

^b Lomonosov Moscow State University, 1 Leninskie Gory, Moscow, 119991, Russian Federation

Резюме

В работе исследуются показатели эмоциональной окраски газетных новостей по тематике коронавирусной эпидемии, публиковавшихся в различных странах в наиболее яркий период начала эпидемии. Общая выборка составляет 12 000 новостных записей, опубликованных в период с 1 по 30 марта 2020 г. на итальянском, немецком, французском и русском языках. Среди рассматриваемых стран наиболее стабильной тональностью публикуемых новостей наблюдалась в России, преимущественно негативная тональность новостных записей наблюдалась во Франции, а позитивная — в Италии. Показано, что тональность новостей в наибольшей степени связана с ростом пандемии, а не с абсолютными показателями

Abstract

This paper describes the indicators of the emotional tone of newspapers daily news about coronavirus epidemic, published in various countries at the beginning of the epidemic. The total sample is 12,000 news entries published between 1st and 30th March 2020 in Italian, German, French and Russian languages. Among the countries under consideration, the most stable tone of the published news was observed in Russia, the predominantly negative tone of news records was in France, and positive tone was found in Italy. It is shown that the emotional tone of news was strongly related with the increase in the pandemic indicators, and was less

Исследование выполнено при финансовой поддержке РФФИ, проект № 20-04-60156 Вирусы.

The study was funded by RFBR, project number 20-04-60156.

телями ее распространения. Эмоциональный тон прессы в большей степени определялся общемировыми событиями, связанными с пандемией, чем распространением заболевания внутри собственной страны. Величина «временного окна», в период которого пресса сохраняла тональность публикуемых новостей, составляла 3–4 дня. С использованием метода иерархической регрессии были построены модели, предсказывающие тональность прессы следующего дня для каждой из анализируемых стран: Франция ($R^2 = 0.78$), Италия ($R^2 = 0.75$), Германия ($R^2 = 0.76$), Россия ($R^2 = 0.85$). Для Франции и Италии наиболее значимый вклад в эту тональность вносят параметры распространения эпидемии в мире в предшествующие 3 дня и эмоциональная тональность локальных СМИ в предыдущие 3 дня. Для Германии большую объяснительную силу имеет показатель эмоционального тона прессы в предыдущие 4 дня, а также эмоциональный тон итальянских СМИ в предшествующие 1–2 дня. В России можно отметить влияние общемировых показателей распространения эпидемии и эмоциональный тон СМИ в предыдущие дни.

Ключевые слова: COVID-19, коронавирус, сентимент-анализ, новости, глобализация, иерархическая регрессия, СМИ.

Панфилова Анастасия Сергеевна — научный сотрудник, лаборатория психологии и психофизиологии творчества, ФГБУН «Институт психологии Российской академии наук», кандидат технических наук.

Сфера научных интересов: когнитивная психология, методы машинного обучения.

Контакты: panfilova87@gmail.com

Ушаков Дмитрий Викторович — директор, Институт психологии РАН; заведующий кафедрой общей психологии, факультет психологии МГУ имени М.В. Ломоносова, академик РАН, доктор психологических наук, профессор.

Сфера научных интересов: структурно-динамическая теория интеллекта, психология творчества и одаренности, социальный и эмоциональный интеллект, макропсихология, методологические проблемы психологии.

Контакты: ushakovdv@ipran.ru

related with the absolute indicators of its spread. The emotional tone of the press was stronger determined by the global events associated with the pandemic than by the spread within their own country. Three to 4 days was the size of the sliding window during which the press maintained the tone of the published news. Hierarchical regression method was used for built models for predicting the tone of the next day's press for each of the analyzed countries: France ($R^2 = 0.78$), Italy ($R^2 = 0.75$), Germany ($R^2 = 0.76$), Russia ($R^2 = 0.85$). The most relevant predictors were identified for each model. The most significant contribution for France and Italy was made by the parameters of the spread of the epidemic in the world in the previous 3 days and the emotional tone of the local media in the previous 3 days. The indicator of the emotional tone of the press in the previous 4 days and the emotional tone of the Italian media in the previous 1–2 days had a great explanatory power for Germany. In Russia, one can note the impact of global indicators of the spread of the epidemic and the emotional tone of the media in previous days.

Keywords: COVID-19, coronavirus, sentiment analysis, newspaper, globalization, hierarchical regression method, mass media.

Anastasia S. Panfilova — Research Fellow, Institute of Psychology, Russian Academy of Sciences, PhD in Engineering.

Research Area: cognitive psychology, machine learning methods.

E-mail: panfilova87@gmail.com

Dmitry V. Ushakov — Director, Institute of Psychology of the Russian Academy of Sciences; Head of the Department, Department of Psychology, Lomonosov Moscow State University, Member of the Russian Academy of Sciences, DSc in Psychology, Professor.

Research area: New Investment Theory of intelligence, creativity, giftedness, social and emotional intelligence, macropsychology, philosophy of psychology.

E-mail: ushakovdv@ipran.ru

Пандемия — явление, находящее сильный эмоциональный отклик даже у тех, кого она непосредственно не затронула. Тревога, страх, надежда, благодарность, ревность соперничества и т.д. — вот далеко не полный перечень эмоций, которые переживали люди по всему миру в период развития коронавирусной инфекции. Переживание эмоций современным обществом — очень сложный процесс. Дело не только в том, что эмоции огромного количества поразному связанных между собой людей действуют друг на друга, резонируя или попадая в противофазу. Дело еще и в том, что средства массовой информации, обладая возможностью донести свою позицию до очень многих, выступают усилителем и проводником тех или иных эмоциональных посланий в общественном поле. Именно СМИ выполняют глобализационную функцию, мгновенно донося раскаты грома эпидемии из места, где она набирает силу, до еще не испытывавших ее точек планеты.

Эмоции, выражаемые СМИ, не могут не опираться на эмоции населения, поскольку только при таком условии СМИ интересны и принимаются населением. Журналисты, испытывая эмоции вместе с другими людьми, подхватывая их, развивают в ярких образах и вбрасывают в общественное мнение, встречая одобрение или неприятие, но редко оставляя людей равнодушными.

Однако эмоции, реально переживаемые народом, отражаются СМИ в измененном и переработанном виде. По законам жанра, средства массовой информации не могут не использовать «громких» заголовков для привлечения внимания аудитории (Ecker et al., 2014) и не концентрироваться на негативных новостях (Soroka, McAdams, 2015). Отрицательная новость получит большее число комментариев и вызовет большую вовлеченность пользователей по сравнению с положительной, а эмоциональная окраска реакции читателей в виде комментариев коррелирует с эмоциональной окраской самой новости (Kumar et al., 2018).

Кроме того, СМИ всегда имеют функцию вести за собой общественное мнение, что заставляет отфильтровывать в том числе и сообщения с эмоциональной окраской, не соответствующей стремлениям авторов или лиц, которые контролируют соответствующие информационные органы. В данной статье не имеет смысла обсуждать причины, которые определяют выбор средством массовой информации той или иной эмоциональной позиции, будь то ответственность за происходящее, видение желательного будущего, корыстные политические или финансовые интересы. Однако в целом эмоциональные реакции людей по всему миру во многом опосредованы деятельностью СМИ.

Само собой разумеется, что средства массовой информации далеки от того, чтобы быть монолитом, согласованно ведущим за собой общественное мнение. Современные исследования показывают, что эмоциональная окраска новостей может отличаться в зависимости от типа СМИ; существуют значительные различия и в способах представления одних и тех же новостей с точки зрения позитивной и негативной окраски в зависимости от типа источника. Пандемия коронавируса дает хороший материал для исследования описанных процессов. Она представляет собой яркое эмоциональное событие, и в

то же время ее сюжеты и угрозы поддаются квантификации. Можно проанализировать связь между объективными, формально оцененными параметрами эпидемии в разных странах, такими как число заболевших и умерших, и эмоциональной реакцией прессы на них. Это позволяет рассмотреть ряд вопросов.

Во-первых, интересно исследовать характер связи между объективно оцененной силой эпидемии и силой эмоциональной реакции на нее. Закон Вебера—Фехнера гласит, что такого рода связь может иметь логарифмический характер. Жизненный опыт подсказывает, что в случае эмоциональной реакции на события действие логарифмического закона правдоподобно: эмоциональная реакция зависит от того, насколько сильно вызывающее эмоции событие отличается от того, что ему предшествовало. Так, тяжелые исторические периоды типа войн и эпидемий притупляют эмоциональную реакцию, переживание отдельных событий становится менее острым.

В случае эпидемии коронавируса можно предположить, что эмоциональный тон прессы будет зависеть не столько от количества вновь заболевших или недавно умерших, сколько от отношения этого количества к числу заболевших и умерших в предыдущий момент времени. Другими словами, мы предполагаем, что эмоциональную реакцию вызывает больше изменение интенсивности эпидемии, чем сила эпидемии сама по себе.

Во-вторых, важно определить интервал событий эмоционального «настоящего времени». Эмоциональную реакцию мы можем испытывать по отношению к событиям, происходящим в данный момент, а можем вновь переживать, возвращаясь к прошлому. Однако настоящий момент, вызывающий эмоциональную реакцию, тоже имеет определенную длительность. В отношении некоторых событий эта длительность минимальна — минуты и даже секунды, но иногда она растягивается на часы, дни, месяцы. Анализ разворачивания эпидемии в комплексе с его эмоциональным отражением в прессе открывает возможность объективно оценить продолжительность воздействия эмоциогенных событий на эмоциональный фон переживаний людей. Конечно, сам факт эпидемии оставит глубокий эмоциональный след надолго. В то же время мы предполагаем, что продолжительность эмоциональной реакции прессы на отдельные события разворачивания эпидемии — такие как ее спады и подъемы — может составлять несколько дней.

В-третьих, люди переживают как глобальное, общемировое развитие эпидемии, так и ее локальное состояние в своей стране и местности. Начало XXI в. характеризовалось бурной глобализацией, которая проявляется и в том, что люди по всему миру начинают жить одними и теми же событиями. Если пару столетий назад люди прежде всего были информированы о событиях в своем ближнем и ближайшем окружении, а о происходящем в других странах узнавали с запозданием и в редуцированном виде, то ныне технические возможности позволяют информации очень быстро циркулировать по всему миру, но взамен индивидуализация личной жизни, возможно, сокращает информационные потоки относительно ближайшего окружения. Возникает, однако, вопрос, сопровождается ли перераспределение объемов поступающей информации

соответствующим эмоциональным отношением? Сколь существенный эмоциональный отклик получают дальние события в сравнении с ближними? Эти вопросы можно задавать относительно конечных «пользователей» информации, но можно и относительно эмоционального тона информационных текстов, который доходит до пользователей.

В-четвертых, можно оценить стабильность эмоционального тона, установившегося в прессе, независимо от происходящего разворачивания эпидемии. Здесь можно предположить полярные варианты. Первый вариант — эмоциональная инертность прессы, при которой эмоциональный тон каждого следующего дня имеет тенденцию к сохранению близости с тоном предыдущего даже при изменении текущей ситуации. Второй вариант заключается в эмоциональной лабильности — склонность к изменению эмоционального тона следующего дня по сравнению с предыдущим.

Сентимент-анализ

Для ответа на поставленные выше вопросы требуется по возможности лишенное субъективности, т.е. формализованное, выявление эмоционального тона больших текстовых массивов из прессы разных стран. Это достижимо с помощью развившихся в последнее время методов автоматического анализа текстов, получивших название сентимент-анализа.

Наиболее широкое распространение сентимент-анализ получил в связи с изучением реакции рынка на новости (Афанасьев и др., 2019; Федорова и др., 2018; Перминов, Ахмедова, 2014; Kartick, Banerjee, 2019; Chen, Chen, 2019; Rognone et al., 2020; Li et al., 2020). Первые исследования П. Тетлока (Tetlock, 2007) обнаружили, что оценка эмоционального фона новостей может предсказать прибыль компаний и доходность акций. Анализируя взаимосвязь между настроениями в новостях и доходностью фондового рынка за период 2004–2010 гг., Л. Смейлс показал значимую связь между агрегированными показателями настроения новостей и доходностью акций, которая колеблется во времени и в зависимости от рассматриваемой отрасли, причем данная взаимосвязь асимметрична: реакция на негативные новости проявляется в большей степени (Smales, 2016).

Значительное увеличение числа пользователей социальных сетей за последнее десятилетие привело к тому, что реакцию общества на те или иные события, происходящие в мире, возможно изучать, применяя сентимент-анализ к записям в социальных медиа (Богданов, Дуля, 2019; Гималетдинова, Довтаева, 2020; Baoqing et al., 2020; Hui et al., 2017; Rahab et al., 2020). В 2013 г. один из двух ведущих поставщиков данных для профессиональных участников финансовых рынков Bloomberg добавил анализ настроений записей в Twitter в реализуемое программное обеспечение (Yadav et al., 2019), что характеризует данную сеть как один из источников для принятия решений, отражающий реакцию общества на те или иные события, происходящие в мире. Анализ социальных медиа используется в контексте стихийных бедствий и значимых политических событий как для быстрого выявления, так и для прогнозирования

эмоционального состояния в ответ на события (Gavilanes et al., 2019; Ruz et al., 2020).

Технология sentiment-анализа может быть реализована с использованием различных методов (Kataayama et al., 2019). Первоначальный алгоритм предполагал применение словаря тональных слов (lexicon-based approach), далее были добавлены комплексы правил построения фраз (rule-based approach). Развитие методов машинного обучения позволило повысить точность оценок эмоциональной окраски текста, однако данный метод крайне чувствителен к обучающему набору данных, что могло представлять проблему некоторое количество лет назад.

Метод

Данные

Для формирования выборки новостей использовался API¹ специализированного интернет-сервиса <https://newsapi.org>. В отличие от остальных ресурсов, созданных крупными агентствами и компаниями, предоставляющими материалы только на одном языке либо материалы одного новостного интернет-издания, данный сервис содержит обширную подборку источников на разных языках и предоставляет возможность осуществлять поиск по названию и тексту новости по ключевому слову с указанием языка поиска за заданный период времени.

Для получения русскоязычных новостей использовалось ключевое слово «коронавирус» с указанием языка поиска (ru), для иностранных новостей были заданы слово «coronavirus» и указатели языка (it, de, fr). Запрос осуществлялся по каждому дню в отдельности с 1 по 30 марта 2020 г., по каждому языку и включал в себя указание сортировки новостей по актуальности (relevancy) с лимитом в 100 штук.

Таким образом, была сформирована выборка данных из 3000 новостей на каждом из четырех языков. Каждая новость содержит следующие атрибуты: заголовок, краткое описание, дата и время публикации, название интернет-ресурса.

Для исследования использовались языки, которые в большей степени соответствуют странам, где они являются официальными. К примеру, испанский язык, исключенный из анализа, является официальным в 19 странах мира на четырех континентах, что делает невозможным установить основную страну с наибольшим числом населения, которое его использует. По аналогичной причине не рассматривались новости на английском языке. Немецкий язык является официальным языком не только в Германии, но и в Австрии и Швейцарии, однако численность населения последних двух стран составляет

¹ Application programming interface (API) сервиса предоставляет набор готовых процедур, функций и структур, с помощью которых разработчики могут создавать свои программы, приложения для работы с сервисом.

суммарно лишь 21% населения Германии, поэтому немецкий язык был включен в анализ.

Россию в наших данных представляют 200 информационных источников, Италию — 58, Германию — 70, Францию — 84. В Приложении 1 представлены источники, опубликовавшие более 20 новостных сообщений по тематике коронавирусной инфекции в течение изучаемого периода (показаны не все источники, используемые в анализе). Среди наиболее представленных на различных языках СМИ можно выделить следующие:

- на русском языке: Вести, Лента, Лайф, Тасс, Фонтанка;
- на немецком языке: Spiegel, Welt, Sueddeutsche, Faz, Heise, Die Zeit;
- на французском языке: LeFigaro, L'Equipe, Huffingtonpost, Liberation;
- на итальянском языке: La Repubblica, Corriere, Lifattoquotidiano, ANSA.

Ежедневные отчеты по распространению пандемии во всех странах и мире публикуются Всемирной организацией здравоохранения на интернет-ресурсе, посвященном коронавирусной инфекции². В связи с тем, что в данных ВОЗ по рассматриваемым странам наблюдаются пропуски значений по некоторым датам, возникла необходимость сопоставления их с регулярно обновляемым сторонним источником — платформой Coronavirus Resource Center³, представленной Johns Hopkins University School of Medicine.

Сентимент-анализ

Существующие предобученные модели для различных языков могут значительно отличаться по объему и качеству, а комбинации используемых методов демонстрируют разную точность для используемых в исследовании языков, что не позволит выполнить сравнительный анализ полученных оценок тональности материалов. С целью преодоления данной проблемы был использован API Cognitive Services, предоставляемый Microsoft Azure⁴, — облачная платформа, реализующая доступ к различным сервисам, в частности, для работы с текстом. Данный сервис адаптирован для работы в мультиязычном режиме, позволяя указывать название языка анализируемого фрагмента текста. В качестве результата своей работы сервис возвращает оценку тональности краткого описания новости в диапазоне от 0 до 1, где минимальное значение соответствует негативному эмоциональному тону, максимальное значение — позитивному, а значение 0.5 соответствует нейтральной эмоциональной окраске.

Подготовка данных

С целью анализа динамики настроений СМИ показатели эмоциональной окраски новостных описаний агрегированы по дням. Также введены два показателя: доля новых случаев заражения (*proportion new cases*) и доля новых

² <https://virusncov.com>

³ <https://coronavirus.jhu.edu>

⁴ <https://azure.microsoft.com/ru-ru/>

смертей (*proportion new deaths*). Доля новых случаев заражения рассчитывается как отношение числа новых инфицированных за текущий день к общему числу зараженных в предыдущий день, для показателя смертности значения высчитываются аналогичным образом. Данные характеристики рассчитаны для каждой страны в отдельности и для мира в целом.

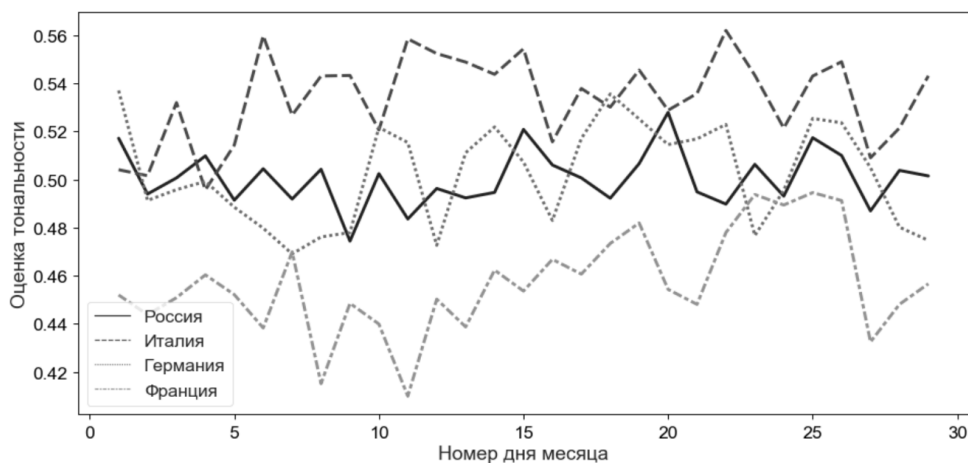
Обращает внимание тот факт, что данные по новым выявленным случаям заражения и смерти появляются в отчетах на следующий день и, соответственно, в представленной публикации рассматриваются данные, опираясь на дату их объявления общественности. Например: утром 10 апреля 2020 г. опубликована информация о 1786 случаях новых заражений в России, которые произошли 9 апреля. В анализируемых данных показатель 1786 относился бы к 10 апреля, как и во всех общедоступных отчетах. Объявление новых данных по распространению эпидемии логичным образом может повлиять на публикации интернет-СМИ в этот день. Однако, предполагая, что в разных странах настроение сетевых изданий может определяться событиями дня не только сегодняшнего, но и предыдущих, мы ввели в данные временные лаги, т.е. временные сдвиги, от 1 до 5 дней по характеристикам распространения пандемии. Также в процессе анализа использовался метод скользящего окна шириной от 15 до 29 дней, позволяющий выявить периоды схожего/противоположного направления изменений функций.

Результаты исследования

На рисунке 1 представлен график изменения тональности новостей сетевых СМИ с 1 по 30 марта для России, Италии, Франции и Германии.

Рисунок 1

График эмоциональной окраски новостей сетевых СМИ по тематике коронавирусной инфекции в период с 1 по 30 марта 2020 г.



Визуальный анализ графиков свидетельствует о том, что, во-первых, в разных странах большие или меньшие различия уровней оценок эмоциональной окраски новостных блоков наблюдаются на всем протяжении периода анализа; во-вторых, можно заметить признаки как сходства, так и своеобразия траекторий изменения анализируемых оценок на некоторых участках.

Очевидно, что средняя эмоциональная тональность прессы для рассмотренных стран различается. Это подтверждается данными, представленными на рисунке 2.

Все агрегированные по дням оценки настроений СМИ значительно различаются между собой по критерию Тьюки для всех пар сравниваемых параметров за исключением России и Германии. Наиболее отрицательный тон новостей характеризует Францию, наиболее положительный, как это ни странно на первый взгляд, — Италию, более всего страдавшую от эпидемии в рассматриваемый период. Значения характеристик приведены в таблице 1.

Рисунок 2

Распределение усредненных по дням оценок эмоциональной окраски новостных блоков для каждой страны

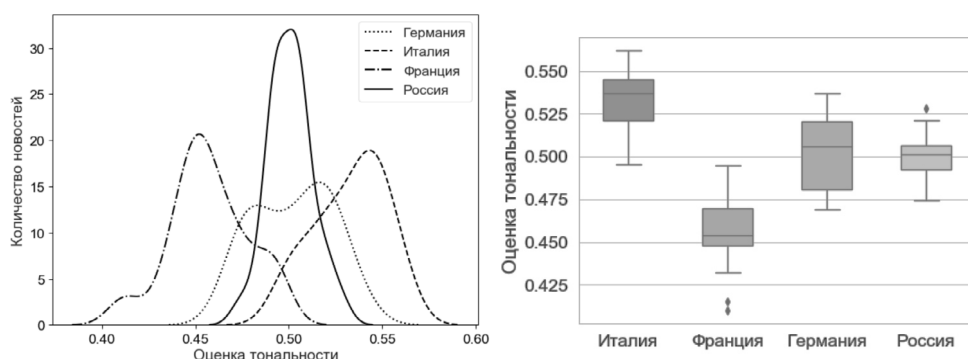


Таблица 1

Результаты попарного сравнения средних оценок тональности новостных блоков для каждой страны по критерию Тьюки

| Группа 1 | Группа 2 | meandiff | p adj | Lower | Upper | Reject |
|----------|----------|----------|-------|---------|--------|--------|
| Германия | Италия | 0.0306 | 0.001 | 0.0182 | 0.0430 | True |
| Германия | Россия | -0.0015 | 0.900 | -0.0139 | 0.0109 | False |
| Германия | Франция | -0.0448 | 0.001 | -0.0573 | 0.0324 | True |
| Италия | Россия | -0.0321 | 0.001 | -0.0445 | 0.0197 | True |
| Италия | Франция | -0.0754 | 0.001 | -0.0878 | 0.0630 | True |
| Россия | Франция | -0.0433 | 0.001 | -0.0557 | 0.0309 | True |

Для России характерен наименьший разброс оценок по дням. Другими словами, эмоциональный тон прессы России в рассматриваемый период был более стабильным, чем в трех других странах.

Данный вид анализа показывает, что различие в эмоциональном тоне прессы разных стран не связано ни с силой эпидемии в стране в исследуемый период, ни со скоростью ее распространения, так как наблюдаемые статистически значимые различия в показателях средних значений тональности новостных сообщений между рассматриваемыми странами показывают, к примеру, преобладание позитивного тона итальянской прессы при наиболее стремительном и масштабном развитии эпидемии в этой стране по сравнению с остальными странами.

Зависимость эмоционального тона прессы от хода пандемии

Был проведен анализ взаимосвязи динамики агрегированных оценок эмоциональной окраски новостных блоков и динамики показателей эпидемии коронавируса для всех четырех стран за период с 1 по 30 марта 2020 г.

Мы использовали несколько параметров для оценки разворачивания эпидемии. Эти параметры делятся на абсолютные и относительные. Абсолютные выражают просто число заболевших и умерших, в то время как относительные — отношение числа заболевших и умерших к общему числу заболевших и умерших к предыдущему дню. Таким образом, абсолютные значения характеризуют силу эпидемии в соответствующий момент времени, а относительные — скорость ее прироста. В соответствии с поставленным исследовательским вопросом нас интересовало, на что больше реагирует эмоциональный тон прессы — на абсолютные показатели эпидемии или на скорость ее прироста.

Было оценено восемь параметров эпидемии, для каждого из которых в дальнейшем будет применяться сокращенное обозначение. Первые четыре параметра относятся к разворачиванию эпидемии в мире:

1. Общее число умерших в день в мире (АУМ).
2. Общее число вновь заболевших в мире (АЗМ).
3. Отношение числа умерших в мире к общему числу умерших в мире к предыдущему дню (ОУМ).
4. Отношение числа вновь заболевших в мире к общему числу заболевших в мире к предыдущему дню (ОЗМ).

Оставшиеся четыре переменные относятся к разворачиванию эпидемии в одной стране:

1. Общее число умерших в день в стране (АУС).
2. Общее число вновь заболевших в стране (АЗС).
3. Отношение числа умерших в стране к общему числу умерших в стране к предыдущему дню (ОУС).
4. Отношение числа вновь заболевших в стране к общему числу вновь заболевших в стране к предыдущему дню (ОЗС).

Эмоциональный тон прессы и ход эпидемии в мире

Был проведен регрессионный анализ для всех четырех исследуемых стран, где в качестве зависимой переменной выступал эмоциональный тон прессы, а независимых — четыре параметра разворачивания эпидемии в мире.

Значимые результаты получены лишь для Франции: на эмоциональный тон прессы отрицательно влияла переменная ОУМ — прирост числа умерших в мире ($R^2 = 0.42$).

Эмоциональный тон прессы и ход эпидемии в своей стране

Аналогичный анализ был проведен для определения влияния характеристик эпидемии в стране на эмоциональный тон прессы в этот день. Значимые связи были найдены только для Италии для переменной ОЗС — отношения числа вновь заболевших в стране к числу заболевших к предыдущему дню ($R^2 = 0.52$).

Таким образом, мы видим, что на эмоциональный тон прессы оказывают влияние только параметры прироста эпидемии, а не ее абсолютные размеры сами по себе.

Эмоциональный тон прессы и разворачивание эпидемии в течение предыдущих дней

Еще один поставленный выше исследовательский вопрос касался временного «окна», в течение которого события эпидемии оказывают влияние на эмоциональный тон прессы. Для анализа этого вопроса было изучено влияние восьми перечисленных выше переменных на эмоциональный тон прессы с отсрочкой от одного до четырех дней. Ниже отсрочка будет обозначаться индексом переменной, равным числу дней отсрочки. Например, переменная АУС2 обозначает число умерших в стране от ковида за два дня до момента, в который оценивается эмоциональный тон прессы.

Изучение взаимосвязи эмоциональной окраски новостей с учетом временного лага для показателей распространения эпидемии в стране проводилось для различных временных периодов (начало периода: 2–10 марта; окончание периода: 21–30 марта).

Рассмотрим результаты, полученные для разных стран.

Для Франции на период 11–28 марта обнаружено влияние на эмоциональный тон прессы переменной ОУС2 — прироста числа умерших внутри страны за 2 дня до изучаемого момента ($R^2 = 0.38$). Также в период 3–28 марта во Франции значимое влияние на эмоциональный тон прессы оказывала переменная ОУМ3 — рост числа умерших в мире за три дня до изучаемого момента, причем с тем же коэффициентом детерминации $R^2 = 0.38$.

Для Италии и Германии значимой оказалась переменная ОЗМ1 — рост числа заболевших в мире за один день до изучаемого момента. Для Италии эта

закономерность действовала на протяжении периода 1–28 марта ($R^2 = 0.32$), а для Германии – 8–29 марта ($R^2 = 0.41$).

Для России значимого влияния отсроченных переменных, с учетом поправки Бонферрони, не наблюдалось.

В целом можно констатировать, что и при рассмотрении отсроченного влияния имели значения только переменные прироста эпидемии, а не ее абсолютные значения. Наиболее значимое влияние наблюдается при отсрочке от одного до трех дней.

Автокорреляция эмоционального тона прессы

Воздействие на прессу оказывают не только разворачивающиеся события эпидемии, но и она сама: предыдущие публикации в СМИ могут оказывать воздействие на эмоциональный тон последующих. Здесь можно представить два вида зависимости. Первая – инерционная: взятый тон имеет тенденцию воспроизводиться и дальше. Вторая – колебательная: следующие публикации стремятся компенсировать некоторую эмоциональную односторонность предыдущих, в результате чего эмоциональный маятник как бы раскачивается.

Для оценки силы связи значений ряда оценок тональностей с предыдущими в зависимости от величины временного сдвига была использована автокорреляция, т.е. последовательный анализ корреляций одного и того же временного ряда, сдвинутого на определенный промежуток времени – лаг. Проверка с использованием обобщенного теста Дики–Фуллера продемонстрировала стационарность оценок тональности новостных блоков по тематике коронавируса для показателей каждой из анализируемых стран в период с 1 по 30 марта 2020 г. Это означает, что во всех странах эмоциональная тональность прессы изменялась на протяжении измеряемого периода в соответствии с определенными внутренними закономерностями. Результаты дополненного теста Дики–Фуллера для рассматриваемых стран приведены в таблице 2.

При этом наблюдались различия между странами. Для Италии (см. рисунок 3а) характерна зависимость, которая выше была обозначена как инерционная: эмоциональный тон прессы имеет тенденцию к сходству с эмоциональным тоном не менее чем трех предыдущих дней. Аналогичным образом

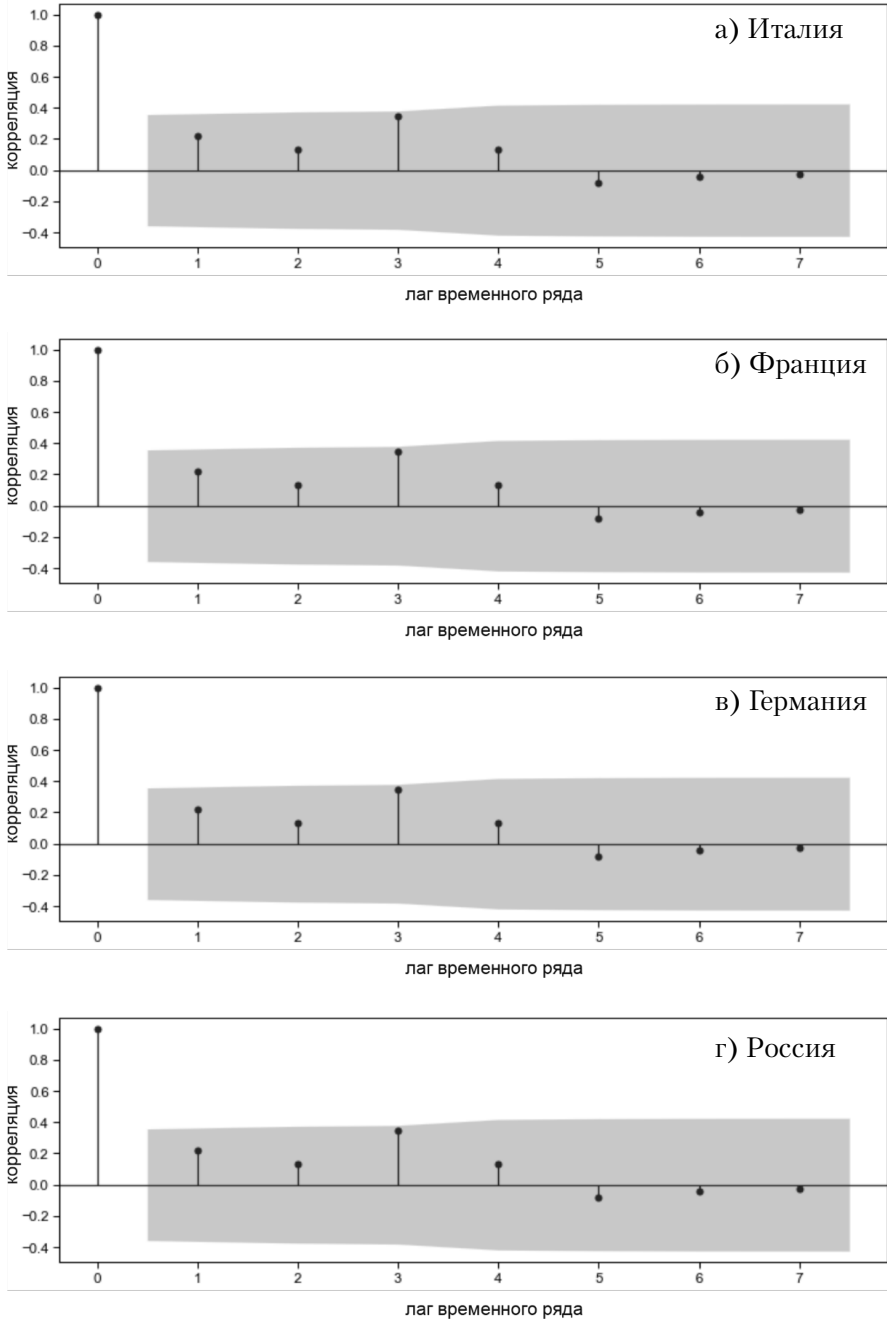
Таблица 2

Значение характеристик результатов анализа временных рядов оценок тональности новостей рассматриваемых стран по тесту Дики–Фуллера

| Страна | Augmented Dickey-Fuller test statistic | Critical value | <i>p</i> -value |
|----------|--|----------------|-----------------|
| Россия | 5.87 | 3.68 | < 0.01 |
| Германия | 3.75 | 3.68 | < 0.01 |
| Италия | 4.38 | 3.68 | < 0.01 |
| Франция | 3.45 | 2.97 | < 0.05 |

Рисунок 3

Коррелограмма оценок тональности новостных блоков изданий разных стран по тематике коронавирусной инфекции в период с 1 по 30 марта 2020 г. (лаг временного ряда измеряется в днях)



для Франции и Германии (см. рисунки 3б и 3в) эмоциональный тон прессы имеет тенденцию к сходству с эмоциональным тоном предыдущего дня.

Для России (см. рисунок 3г) в рассматриваемый промежуток времени была свойственна колебательная зависимость, которая, однако, не является статистически значимой.

Общая регрессионная модель эмоциональной тональности прессы

Далее был проведен анализ сравнительной значимости влияния на эмоциональный тон прессы основных рассмотренных факторов: показателей распространения эпидемии внутри страны и в мире, с отсрочкой и без нее, а также эмоционального тона прессы в предыдущие дни. Для этого для каждой из четырех изучаемых стран строились иерархические регрессионные модели с пошаговым добавлением независимых переменных.

В моделях прогноза эмоционального тона прессы текущего дня, построенных для Франции, показатели распространения эпидемии внутри страны (модель 1) объясняют 18% дисперсии тональности прессы, однако добавление общемировых параметров распространения эпидемии (модель 2) увеличивает объясняемую дисперсию до 51%. Еще 17% добавляют параметры распространения эпидемии в мире в предшествующие три дня (модель 3). Дополнительный учет эмоциональной тональности СМИ в предыдущие три дня (модель 4) доводит объясняемую дисперсию до 78% (см. таблицу 3).

Таким образом, для Франции наибольшее влияние на тон прессы оказывали самые последние данные о распространении эпидемии в мире. Они влияли больше, чем свежие данные об эпидемии в самой Франции, и больше, чем данные об эпидемии предыдущих дней.

В моделях прогноза эмоционального тона прессы текущего дня, построенных для Италии, показатели распространения эпидемии внутри страны оказали несколько большее влияние, чем для Франции, и объяснили 21% дисперсии (модель 1), в то время как добавление показателей разворачивания эпидемии в мире (модель 2) увеличило детерминацию на 23%, что меньше, чем во Франции. Также не очень значительный прирост дает прибавление параметров распространения эпидемии в мире в предшествующие три дня (модель 3) —

Таблица 3

Параметры иерархической регрессии модели прогнозирования тональности новостных блоков интернет-СМИ по тематике коронавирусной инфекции во Франции

| Модель | R^2 | Статистика изменений | | | | ANOVA |
|--------|-------|----------------------|-----------|-----------|---------------------|---------|
| | | Изменение F | Ст. св. 1 | Ст. св. 2 | Значим. изменения F | p-value |
| 1 | 0.18 | 1.72 | 3 | 23 | 0.19 | 0.19 |
| 2 | 0.51 | 7.12 | 2 | 21 | 0.004 | 0.007 |
| 3 | 0.68 | 1.27 | 6 | 15 | 0.33 | 0.03 |
| 4 | 0.78 | 1.94 | 3 | 12 | 0.18 | 0.03 |

всего 6%. Однако 25% объясненной дисперсии добавляет введение в модель тональности локальных СМИ в предыдущие три дня (модель 4). В Италии, таким образом, большую роль играли процессы «эмоциональной инертности» прессы (см. таблицу 4).

В моделях прогноза эмоционального тона прессы текущего дня, построенных для Германии, показатели распространения эпидемии внутри страны (модель 1) объясняют 14% дисперсии эмоционального тона прессы, а добавление показателей распространения эпидемии в мире (модель 2) увеличивает объясняемую дисперсию лишь на 8%. Однако большую объяснительную силу имеет добавление эмоционального тона прессы в предыдущие четыре дня (модель 3) — объясняемая дисперсия увеличивается до 55%. К тому же эмоциональный тон прессы Германии оказался связан с эмоциональным тоном итальянских СМИ в предшествующие один-два дня. Добавление этого предиктора (модель 4) доводит объясняемую дисперсию до 76% (см. таблицу 5).

В моделях прогноза эмоционального тона прессы текущего дня, построенных для России, показатели распространения эпидемии внутри страны (модель 1) объясняют 16% дисперсии эмоционального тона прессы, а добавление общемировых показателей (модель 2) доводит этот процент до 39%. В то же время из всех рассмотренных стран Россия больше всего реагирует на

Таблица 4

Параметры иерархической регрессии модели прогнозирования тональности новостных блоков интернет-СМИ по тематике коронавирусной инфекции в Италии

| Модель | R^2 | Статистика изменений | | | | ANOVA |
|--------|-------|----------------------|-----------|-----------|---------------------|---------|
| | | Изменение F | Ст. св. 1 | Ст. св. 2 | Значим. изменения F | p-value |
| 1 | 0.21 | 1.10 | 5 | 21 | 0.39 | 0.39 |
| 2 | 0.44 | 3.85 | 2 | 19 | 0.04 | 0.10 |
| 3 | 0.50 | 0.50 | 4 | 15 | 0.74 | 0.28 |
| 4 | 0.75 | 6.46 | 2 | 13 | 0.01 | 0.03 |

Таблица 5

Параметры иерархической регрессии модели прогнозирования тональности новостных блоков интернет-СМИ по тематике коронавирусной инфекции в Германии

| Модель | R^2 | Статистика изменений | | | | ANOVA |
|--------|-------|----------------------|-----------|-----------|---------------------|---------|
| | | Изменение F | Ст. св. 1 | Ст. св. 2 | Значим. изменения F | p-value |
| 1 | 0.14 | 0.83 | 4 | 21 | 0.52 | 0.52 |
| 2 | 0.22 | 0.65 | 3 | 18 | 0.59 | 0.65 |
| 3 | 0.55 | 3.61 | 3 | 15 | 0.038 | 0.14 |
| 4 | 0.76 | 5.07 | 2 | 12 | 0.025 | 0.038 |

Таблица 6

Параметры иерархической регрессии модели прогнозирования тональности новостных блоков интернет-СМИ по тематике коронавирусной инфекции в России

| Модель | R^2 | Статистика изменений | | | | ANOVA |
|--------|-------|----------------------|-----------|-----------|---------------------|---------|
| | | Изменение F | Ст. св. 1 | Ст. св. 2 | Значим. изменения F | p-value |
| 1 | 0.16 | 0.96 | 4 | 20 | 0.45 | 0.45 |
| 2 | 0.39 | 0.85 | 6 | 14 | 0.55 | 0.57 |
| 3 | 0.85 | 7.74 | 4 | 10 | 0.004 | 0.016 |

эмоциональный тон СМИ в предыдущие дни, введение показателей которого в качестве предикторов (модель 3) увеличивает объясняемую дисперсию более чем вдвое — сразу на 46% (см. таблицу 6).

Для каждой из стран получены модели, где в качестве независимых характеристик методом иерархической регрессии отобраны параметры распространения эпидемии и тональность новостных сообщений СМИ в предыдущие дни в стране и в мире, а в качестве зависимой переменной выступает оценка тональности новостей за текущий день. Значимые параметры моделей приведены в Приложении 2.

Обсуждение результатов

Проведенный анализ позволяет дать ответы на поставленные в начале статьи исследовательские вопросы.

Показатели абсолютной силы эпидемии (число заболевших и умерших в своей стране и в мире) в рассматриваемый период оказывали малое влияние на эмоциональный тон прессы. В значительно большей степени влияли относительные показатели — рост числа заболевших и умерших в день. Этот факт имеет вполне логичное объяснение: мы склонны эмоционально реагировать на изменения среды, в то время как к хорошему и плохому, если они не изменяются, мы постепенно привыкаем.

Пресса всех четырех стран эмоционально реагировала на изменение показателей эпидемии как в своей стране, так и в мире в целом. При этом реакция на общемировой ход эпидемии была сильнее, чем на ход эпидемии в своей стране, во всех странах, за исключением, пожалуй, Германии. Эти результаты могут вызвать удивление в отношении Италии, которая на момент исследования была наиболее страдавшей от эпидемии страной мира. Почему же там пресса более эмоционально реагировала на ход эпидемии в мире, чем в собственной стране? Ответ на этот вопрос может быть дан на основании полученных данных лишь гипотетически. Возможно, речь идет о своего рода коллективных защитных механизмах, которые предохраняют от осознания чрезмерной силы страданий, испытываемых страной. С этих же позиций может быть

объяснен и в целом позитивный эмоциональный тон итальянской прессы в этот тяжелый период.

В целом реакция человека на непосредственные жизненные ситуации меняется с развитием глобализации. В предыдущие столетия люди получали в большей степени информацию из своего ближайшего окружения и непосредственно реагировали на нее, однако, по мере возникновения СМИ, то, что происходит в мире, становится важнее, приобретает все большую актуальность. Информационная ситуация в связи с развитием эпидемии служит хорошей иллюстрацией этой закономерности.

Полученные данные свидетельствуют и о величине «временного окна», в течение которого события эпидемии влияли на эмоциональный тон периодики. В целом наибольшее влияние оказывали новейшие события, самые последние опубликованные данные. В то же время, хотя и в более легкой степени, на эмоциональность прессы влияли и события предшествующих трех-четырёх дней.

Полученные данные позволяют нарисовать своеобразные портреты всех четырех изученных стран в их эмоциональной реакции на удар эпидемии.

Италия, страна, находившаяся в исследованный период под наибольшим давлением эпидемии, демонстрировала при этом среди четырех изученных стран наиболее позитивный эмоциональный тон. При этом эмоциональный тон итальянских СМИ определялся распространением эпидемии в мире даже чуть больше, чем в своей стране. Пресса показала большую эмоциональную инертность — она была склонна поддерживать эмоциональный тон, установившийся на протяжении нескольких последних дней. Возможное объяснение всех перечисленных особенностей итальянской прессы таково: формирование реакций совладания, предполагающее недопущение панических эмоций и когнитивную защиту в виде объективации и децентрированного рассмотрения ситуации — не сконцентрированного на проблемах собственной страны, а разомкнутого на общемировую ситуацию.

Франция, напротив, хотя и не была еще сильно затронута эпидемией, позволяла себе наибольший пессимизм. Как и в Италии, эмоциональный тон французских СМИ обладал высокой инертностью, пресса была склонна воспроизводить эмоциональные паттерны нескольких предшествующих дней. В то же время, по сравнению с Италией, предыдущие публикации в прессе влияли меньше, чем другие факторы. Французская пресса сильнее зависела от показателей распространения эпидемии, чем пресса остальных рассмотренных стран. При этом в большей степени влияли показатели распространения эпидемии в мире, а не в самой Франции.

Германия, занимая вместе с Россией промежуточное положение по эмоциональному тону СМИ среди четырех стран, оказалась менее всего зависимой от текущих показателей эпидемии в своей стране и особенно — в мире. Однако в германской прессе прослеживается интересный феномен: ее эмоциональный тон коррелировал с тоном итальянской прессы за 1–2 дня до этого. Складывается впечатление, что немецкие СМИ реагировали на то, как освещает события пресса Италии, находившейся в то время под наибольшим

ударом эпидемии. При этом немецкая пресса демонстрировала на изученном временном промежутке склонность к автоколебаниям эмоционального фона с периодичностью в один день.

Россия показала средние значения эмоционального тона — примерно на одном уровне с Германией. При этом в России эмоциональный тон был наиболее стабильным из всех четырех стран (наименьшие изменения от одного дня к другому). В России есть СМИ, которые в основном преподносят информацию в негативном свете, в то время как другие — в более позитивном: в других странах такого выраженного уклона сетевых изданий не наблюдается. В нашей стране в рассматриваемый период эмоциональный тон новостных интернет-СМИ на тему коронавируса при всей относительной незначительности его изменений имел тенденцию к колебаниям с периодичностью в пять дней. Параметры эпидемии в мире оказывали на российскую прессу большее влияние, чем разворачивание эпидемии внутри страны. Надо иметь в виду, что в этот период эпидемия в стране была еще слаба, поэтому ориентация на регионы мира с более острой ситуацией вполне понятна.

Полученные результаты подтверждаются аналогичными исследованиями (Alamoodi et al., 2020; Garcia, Berton, 2021; Melo, Figueiredo, 2021), в которых значимые различия в тональности прессы и социальных медиа обнаруживаются для множества языков (португальский, английский, японский, корейский, итальянский, немецкий и пр.) и стран. Данные результаты справедливы как для постов в социальной сети Twitter, так и для текстов и заголовков новостных сообщений (Garcia, Berton, 2021; Ghasiya, Okamura, 2021; Melo, Figueiredo, 2021). Можно добавить, что тональность российских СМИ находится ближе к германским из анализируемых. Стоит отметить, что построение моделей прогноза развития пандемии, учитывающих влияние в том числе и тональности новостей (Chalkiadakis et al., 2021), обоснованно стоит дополнить и показателями развития пандемии в мире, и показателями эмоционального фона предшествующих дней не только в данной стране, но и в других странах. Ряд исследований (Alamoodi et al., 2020; Garcia, Berton, 2021; Melo, Figueiredo, 2021), посвященных анализу тональности сообщений, опубликованных в социальной сети Twitter в связи с COVID-19 и отражающих результаты изучения динамики эмоционального фона сообщений в различных странах и на различных языках и проблем, обсуждаемых в связи с эпидемией (политика, экономика, спорт, здравоохранение), демонстрируют различия в том, каким образом СМИ освещают проблему COVID-19 в зависимости от языка и страны. Существующие работы не в полной мере охватывают национальные СМИ, которые в большей степени являются источником информации для населения, не учитывают временной сдвиг, с которым освещаемые события могут влиять на поведение людей, а также не анализируют влияние новостного фона других стран, динамики численности заболевших и умерших в других странах и в мире и оперируют исключительно абсолютными показателями по численности заболевших и умерших.

Заключение: ограничения исследования

Полученные результаты требуют интерпретации с учетом понятных ограничений.

Прежде всего надо отметить, что исследование относится к периоду, ограниченному всего одним месяцем. Это был месяц, когда эпидемия во всем мире (за исключением не включенного в исследование Китая) только нарастала, успев затронуть в полной мере одну из исследованных стран (Италию), но еще не охватив другие. Как установленные для этого временного промежутка закономерности ведут себя в другие периоды, в том числе — на этапе затухания эпидемии в стране или появления второй волны, составляет предмет отдельных исследований.

Как уже отмечалось, СМИ не являются беспристрастным зеркалом народных эмоций, — скорее, они ведут их за собой. В этом плане более полную картину эмоциональной реакции страны можно получить, если дополнить анализ эмоционального тона СМИ сентимент-анализом содержания соцсетей за аналогичный период. Подобный анализ может показать, в какой степени эмоциональный контент соцсетей сходен со СМИ в разных странах, где наблюдается опережение или запаздывание, и т.д.

Практическую значимость авторы аналогичных исследований видят в том, что их результат может быть использован как механизм сдерживания развития пандемии на начальном этапе, благодаря усиленному негативному эмоциональному фону новостных лент, стимулирующему население принимать большие меры предосторожности (Garcia, Berton, 2021; Desai, 2021; Melo, Figueiredo, 2021), что может быть справедливо и для России.

Исследование было проведено на материале стран, где использовались различные языки. Сегодня сравнительный сентимент-анализ текстов, написанных на разных языках, считается корректным. Исторически развитие подходов к эмоциональной оценке текстов относилось исключительно к английскому языку. В этих условиях использование автоматизированного машинного перевода приводило к ряду искажений. Поэтому в настоящее время применяется подход с использованием параллельных корпусов текста, в которых эмоционально-окрашенным словам из словаря одного языка ставятся в соответствие слова другого языка. Этот подход позволяет получать наборы данных со схожими тональностями в различных языках. Методы, основанные на глубоком обучении (LSTM, CNN), продемонстрировали многообещающие результаты в области анализа настроений за счет включения анализа контекста (Matheus et al., 2020).

Межкультурные сравнения эмоционального фона новостей с использованием сентимент-анализа (Pessutto et al., 2020) были проведены М. Баутином, Л. Виджаярену и С. Скиеной (Bautin et al., 2008) и продемонстрировали наличие значимых корреляций оценок полярностей новостей на различных языках, а наблюдаемые отклонения были интерпретированы локальными событиями, происходившими в соответствующих странах за рассматриваемый период. Также проведен анализ субъектов, упоминание о которых в новостях

и блогах носит преимущественно позитивный или негативный оттенок (Godbole et al., 2007).

Все в целом дает основание считать современный сентимент-анализ надежным методом даже при межязыковых сопоставлениях. При анализе факторов, влияющих на эмоциональный тон внутри языков, он не вызывает сомнения. Все же непосредственные сопоставления эмоционального тона в разных языках требуют осторожности.

Литература

- Афанасьев, Д. О., Федорова, Е. А., Рогов, О. Ю. (2019). О влиянии тональности новостей в международных СМИ на рыночный курс российского рубля: текстовый анализ. *Экономический журнал Высшей школы экономики*, 23(2), 264–289. <https://doi.org/10.17323/1813-8691-2019-23-2-264-289>
- Богданов, А. Л., Дуля, И. С. (2019). Сентимент-анализ коротких русскоязычных текстов в социальных медиа. *Вестник Томского государственного университета. Экономика*, 47, 220–241. <https://doi.org/10.17223/19988648/47/17>
- Гималетдинова, Г. К., Довтаева, Э. Х. (2020). Сентимент-анализ читательского интернет-комментария к политическому дискурсу. *Политическая лингвистика*, 1, 42–51. <https://doi.org/10.26170/pl20-01-05>
- Перминов, Г. А., Ахмедова, С. С. (2014). Сентиментный анализ влияния новостей на фондовый рынок. *Аудит и финансовый анализ*, 4, 408–413.
- Федорова, Е. А., Бечвая, К. З., Рогов, О. Ю. (2018). Влияние тональности новостей на курс биткоина. *Финансы: теория и практика*, 22(4), 104–113. <https://doi.org/10.26794/2587-5671-2018-22-4-104-113>

Ссылки на зарубежные источники см. в разделе *References*.

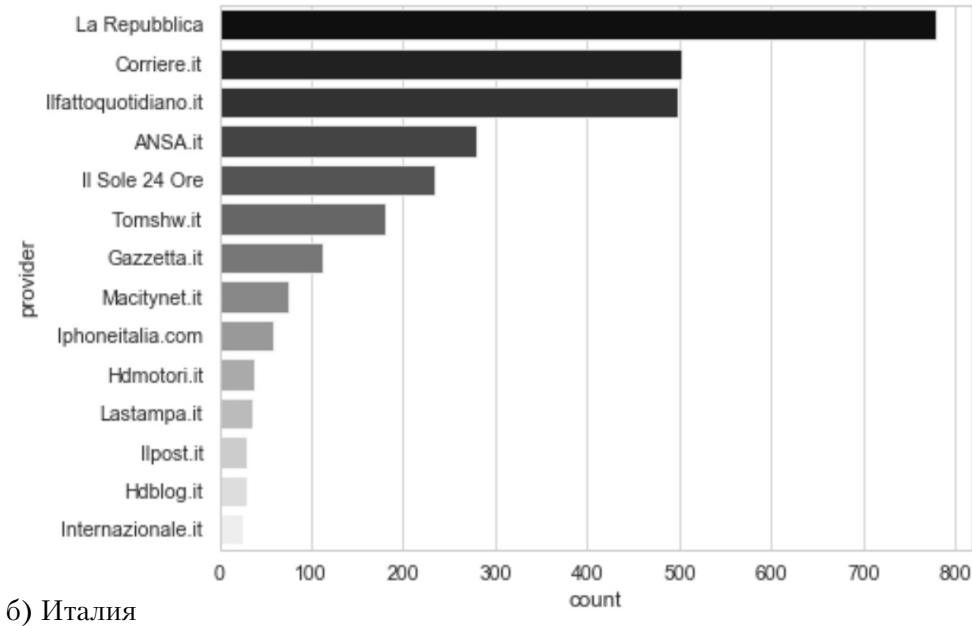
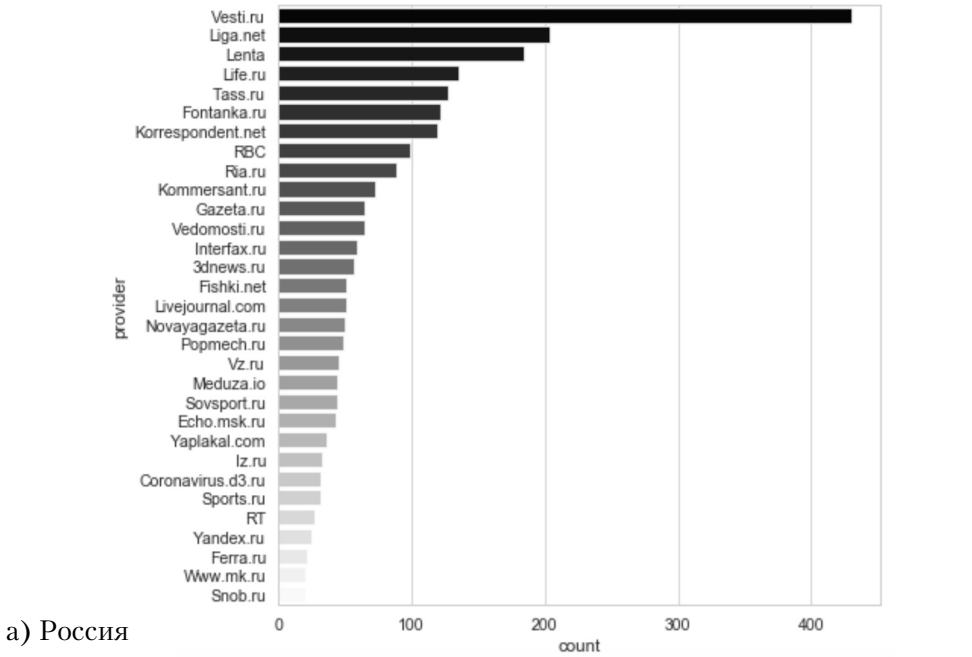
References

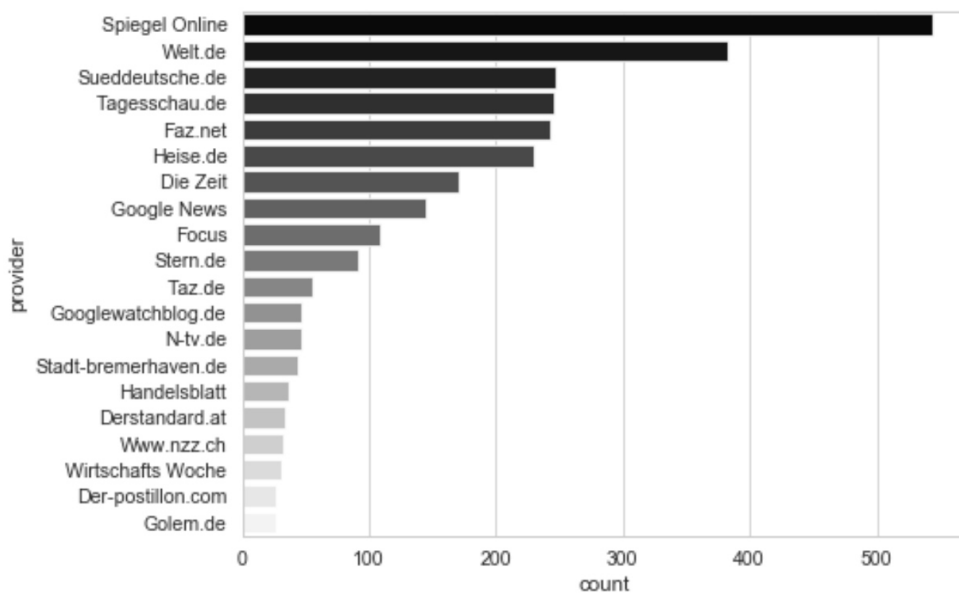
- Afanasyev, D. O., Fedorova, E. A., & Rogov, O. Yu. (2019). On the impact of news tonality in international media on the Russian ruble exchange rate: textual analysis. *Ekonomicheskii Zhurnal Vyshei Shkoly Ekonomiki [HSE Economic Journal]*, 23(2), 264–289. <https://doi.org/10.17323/1813-8691-2019-23-2-264-289> (in Russian)
- Alamoodi, A. H., Zaidan, B. B., Zaidan, A. A., Albahri, O. S., Mohammed, K., Malik, R. Q., Almahdi, E. M., Chyad, M. A., Tareq, Z., Albahri, A. S., Hameed, H., & Alaa, M. (2020). Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Systems with Applications*, 167, Article 114155. <https://doi.org/10.1016/j.eswa.2020.114155>
- Baoqing, G., Alexeev, V., Bird, R., & Yeung, D. (2020). Sensitivity to sentiment: News vs social media. *International Review of Financial Analysis*, 67, Article 101390. <https://doi.org/10.1016/j.irfa.2019.101390>
- Bautin, M., Vijayarenu, L., & Skiena, S. (2008). International sentiment analysis for news and blogs. In *ICWSM 2008 – International Conference on Weblogs and Social Media* (pp. 19–26), Seattle, Washington, USA. Menlo Park, CA: The AAAI Press.

- Bogdanov, A. L., & Dulya, I. S. (2019). Sentiment analysis of short Russian texts in social media. *Vestnik Tomskogo Gosudarstvennogo Universiteta. Ekonomika [Tomsk State University Journal of Economics]*, 47, 220–241. <https://doi.org/10.17223/19988648/47/17> (in Russian)
- Chalkiadakis, I., Yan, H., Peters, G. W., & Shevchenko, P. V. (2021). Infection rate models for COVID-19: Model risk and public health news sentiment exposure adjustments. *PLoS ONE*, 16, Article e0253381. <https://doi.org/10.1371/journal.pone.0253381>
- Chen, M., & Chen, T. (2019). Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. *Future Generation Computer Systems*, 96, 692–699. <https://doi.org/10.1016/j.future.2017.10.028>
- Desai, P. S. (2021). News Sentiment Informed Time-series Analyzing AI (SITALA) to curb the spread of COVID-19 in Houston. *Expert Systems with Applications*, 180, Article 115104. <https://doi.org/10.1016/j.eswa.2021.115104>
- Ecker, U. K., Lewandowsky, S., Chang, E. P., & Pillai, R. (2014). The effects of subtle misinformation in news headlines. *Journal of Experimental Psychology. Applied*, 20(4), 323–335. <https://doi.org/10.1037/xap0000028>
- Fedorova, E. A., Bechvaya, K. Z., & Rogov, O. Yu. (2018). The influence of the tonality of news on the exchange rate of bitcoin. *Finansy: Teoriya i Praktika [Finance: Theory and Practice]*, 22(4), 104–113. <https://doi.org/10.26794/2587-5671-2018-22-4-104-113> (in Russian)
- Garcia, K., & Berton, L. (2021). Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Applied Soft Computing*, 101, Article 107057. <https://doi.org/10.1016/j.asoc.2020.107057>
- Gavilanes, M. F., Juncal-Martínez, J., García-Méndez, S., Costa-Montenegro, E., & González-Castaco, F. (2019). Differentiating users by language and location estimation in sentiment analysis of informal text during major public events. *Expert Systems with Applications*, 117, 15–28. <https://doi.org/10.1016/j.eswa.2018.09.007>
- Ghasiya, P., & Okamura, K. (2021). Investigating COVID-19 news across four nations: A topic modeling and sentiment analysis approach. *IEEE Access*, 9, 36645–36656. <https://doi.org/10.1109/ACCESS.2021.3062875>
- Gimaltdinova, G. K., & Dovtaeva, E. Kh. (2020). Sentiment analysis of the reader's internet commentary on a political text. *Politicheskaya Lingvistika [Political Linguistics]*, 1, 42–51. <https://doi.org/10.26170/pl20-01-05>
- Godbole, N., Srinivasaiah, M., & Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM), USA*. Palo Alto, CA: The AAAI Press. <http://www.icwsml.org/papers/5-Godbole-Srinivasaiah-Skienna-demo.pdf>
- Hui, J., Hoon, G. K., & Zainon, W. N. M. W. (2017). Effects of word class and text position in sentiment-based news classification. In *Procedia Computer Science* (Vol. 124, pp. 77–85). Amsterdam: Elsevier. <https://doi.org/10.1016/j.procs.2017.12.132>
- Kartick, G., & Banerjee, R. (2019). Does OPEC news sentiment influence stock returns of energy firms in the United States? *Energy Economics*, 77, 34–45. <https://doi.org/10.1016/j.eneco.2018.03.017>
- Katayama, D., Yasunobu, K., & Kazuhiko, T. (2019). A method of sentiment polarity identification in financial news using deep learning. *Procedia Computer Science*, 159, 1287–1294. <https://doi.org/10.1016/j.procs.2019.09.298>
- Kumar, N., Nagalla, R., Marwah, T., & Singh, M. (2018). Sentiment dynamics in social media news channels. *Online Social Networks and Media*, 8, 42–54. <https://doi.org/10.1016/j.osnem.2018.10.004>

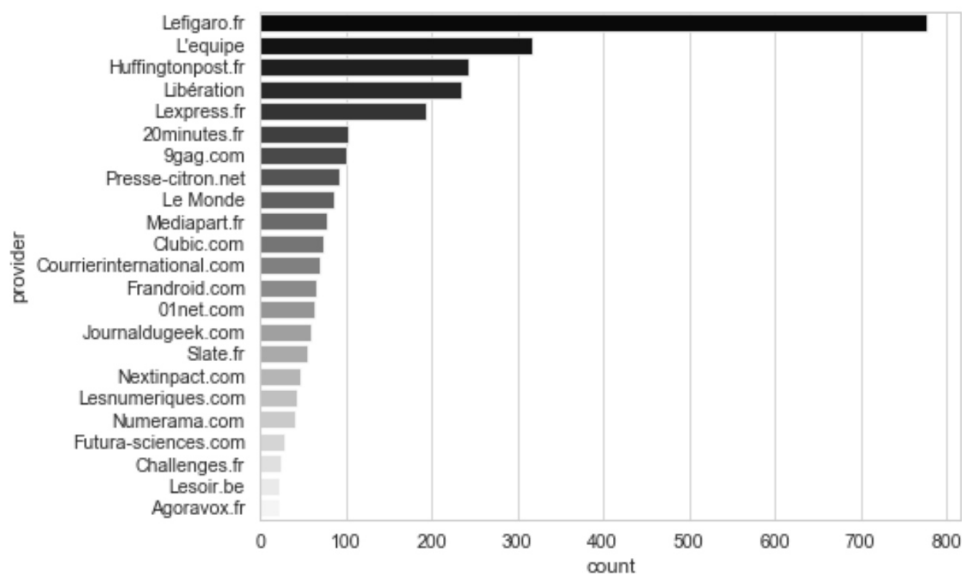
- Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing & Management*, 57, Article 102212. <https://doi.org/10.1016/j.ipm.2020.102212>
- Matheus, A., Pereira, A., & Benevenuto, F. (2020). A comparative study of machine translation for multilingual sentence-level sentiment analysis. *Information Sciences*, 512, 1078–1102. <https://doi.org/10.1016/j.ins.2019.10.031>
- Melo, T., & Figueiredo, C. M. (2021). Comparing news articles and tweets about COVID-19 in Brazil: Sentiment analysis and topic modeling approach. *JMIR Public Health and Surveillance*, 7(2), Article e24585. <https://doi.org/10.2196/24585>
- Perminov, G. A., & Akhmedova, S. S. (2014). Sentimentnyi analiz vliyaniya novostei na fondovyi rynek [Sentiment analysis of the news impacting the stock market]. *Audit i Finansovyi Analiz*, 4, 408–413.
- Pessutto, L. R. C., Vargas, D. S., & Moreira, V. P. (2020). Multilingual aspect clustering for sentiment analysis. *Knowledge-Based Systems*, 192, Article 105339. <https://doi.org/10.1016/j.knosys.2019.105339>
- Rahab, H., Zitouni, A., & Djoudi, M. (2020). SANA: Sentiment analysis on newspapers comments in Algeria. *ArXiv*, Article abs/2006.00459. <https://arxiv.org/pdf/2006.00459.pdf>
- Rognone, L., Hyde, S., & Zhang, S. (2020). News sentiment in the cryptocurrency market: An empirical comparison with Forex. *International Review of Financial Analysis*, 69, Article 101462. <https://doi.org/10.1016/j.irfa.2020.101462>
- Ruz, G. A., Henríquez, P. A., & Mascareco, A. (2020). Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. *Future Generation Computer Systems*, 106, 92–104. <https://doi.org/10.1016/j.future.2020.01.005>
- Smales, L. A. (2016). News sentiment and bank credit risk. *Journal of Empirical Finance*, 38(A), 37–61. <https://doi.org/10.1016/j.jempfin.2016.05.002>
- Soroka, S., & McAdams, S. (2015). News, politics, and negativity. *Political Communication*, 32, 1–22.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62(3), 1139–1168.
- Yadav, R., Kumar, A. V., & Kumar, A. (2019). News-based supervised sentiment analysis for prediction of futures buying behavior. *IIMB Management Review*, 31(2), 157–166. <https://doi.org/10.1016/j.iimb.2019.03.006>

Диаграмма числа новостей по тематике коронавирусной инфекции используемых источников из разных стран с не менее чем 20 записями





в) Германия



г) Франция

Приложение 2

Независимые характеристики регрессионных моделей оценки тональности новостей по тематике коронавирусной инфекции локальных сетевых изданий

| | Россия | Италия | Германия | Франция |
|---|---|---|---|---|
| Показатели распространения эпидемии внутри страны | <i>Заражения:</i> текущий день, -3, -4 дня <i>Смертность:</i> текущий день | <i>Заражения:</i> текущий день, -1, -3 дня <i>Смертность:</i> -1, -3 дня | <i>Заражения:</i> текущий день, -4 дня <i>Смертность:</i> -1, -3 дня | <i>Смертность:</i> текущий день -2, -3 дня |
| Глобальные параметры распространения эпидемии | <i>Заражения:</i> текущий день, -1, -3 дня <i>Смертность:</i> текущий день -1, -3 дня | <i>Заражения:</i> текущий день, -1, -3 дня <i>Смертность:</i> текущий день -2, -3 дня | <i>Заражения:</i> -1, -2, -4 дня | <i>Заражения:</i> текущий день, -1, -2, -3 дня <i>Смертность:</i> текущий день, -1, -2, -3 дня |
| Оценка эмоциональной окраски новостей страны | -1, -2, -4, -5 дней | -2, -3 дня | -1, -2, -3, -4 дня | -1, -2, -3 дня |
| Взаимосвязь с оценкой эмоциональной окраски новостей других стран | | | Италия: -1, -2 дня | |