

## РАЗБИРАНЕ НА УЕБ ПУБЛИКАЦИИТЕ В СЪВРЕМЕННАТА ДИГИТАЛНА СРЕДА ЧРЕЗ ПРИЛОЖЕНИЕ НА КОНЦЕПЦИЯТА ЗА ИНТЕЛИГЕНТНОСТ НА МЕДИИТЕ

Пламен Милев<sup>1</sup>

e-mail: [pmilev@unwe.bg](mailto:pmilev@unwe.bg)

### Резюме

*Статията изследва подходите за анализ на онлайн публикации от водещи телевизионни медии в България, използвайки съвременни методи за обработка на големи данни. Чрез комбиниране на количествени и качествени аналитични техники се демонстрират възможностите за извличане на смисъл и идентифициране на тенденции в дигиталното медийно пространство. Изследването подчертава значението на концепцията за интелигентност на медиите в анализа на онлайн съдържание и допринася за по-доброто разбиране на трансформациите в медийната екосистема. Тези подходи също така предлагат практически насоки за усъвършенстване на анализа и управлението на медийното съдържание.*

**Ключови думи:** анализ на данни, онлайн публикации, интелигентност на медиите, големи данни, дигитална медийна екосистема

**JEL:** C88, L86

### Увод

В епохата на дигитализацията медиите играят централна роля в оформянето на обществените нагласи и поведение, като тяхната динамична и непрекъснато променяща се природа създава предизвикателства за разбирането и анализа на съдържанието им. Съществуващите подходи за анализ на медийно съдържание често се ограничават до специфични аспекти, което не осигурява цялостно разбиране на сложните взаимодействия между аудиторията и медийния контекст. В този смисъл концепцията за интелигентност на медиите, която интегрира различни аналитични методи, включително обработка на естествен език (NLP) и количествени анализи, представлява обещаващ подход за преодоляване на тези ограничения. Медийното съдържание играе ключова роля в съвременното общество, като формира общественото мнение и предоставя информация по значими теми. В контекста на динамично развиващата се дигитална среда, изследването на онлайн медии-

<sup>1</sup> Доцент, доктор, катедра „Информационни технологии и комуникации“, факултет „Приложна информатика и статистика“, УНСС, ORCID: 0000-0002-4867-0586

те става все по-актуално, особено когато става въпрос за идентифициране на тематичните приоритети и модели на публикуване. С развитието на интернет технологиите и разширяващата се роля на медиите, количественият и качественият анализ на публикуваното съдържание предлага възможности за разбиране на социалните, икономическите и политическите тенденции.

Настоящото изследване се стреми да допринесе за по-задълбочено разбиране на медийната среда чрез представяне на методологична рамка за интелигентна обработка на данни от уеб публикации. Макар темата за медийната интелигентност да е частично проучена в академичната литература, интеграцията на различни видове анализи остава недостатъчно развита. Това изследване предлага синтезиран подход, който комбинира различни методи за идентифициране на тенденции, нагласи и връзки в дигиталната медийна среда. Очакваните резултати включват създаването на основа за по-задълбочени изследвания, както и идентифицирането на подходящи аналитични подходи, които могат да бъдат прилагани за разбиране на медийното съдържание в различни контексти. Изследването се фокусира върху онлайн публикациите на българските телевизионни медии за 2024 г., като използва данни, извлечени чрез web scraping. Анализират се публикации от седем значими медийни платформи: Bulgaria ON AIR, Bloomberg TV, BNT, bTV, Евроком, Euronews Bulgaria и Nova TV. Особено внимание се отделя на динамиката в публикуването на съдържание, сезонните колебания и тематичните приоритети. Целта на настоящото изследване е да анализира публикуваното съдържание в тези онлайн медии, за да се разкрият основните тематични акценти, редакционните приоритети и динамиката в отразяването на важни социални и политически въпроси. За постигането на тази цел се използват различни видове анализи, които осигуряват цялостен поглед върху медийното съдържание.

## Методи

Концепцията за интелигентност на медиите (media intelligence) еволюира като интегративен подход, който комбинира методи за анализ на данни и технологии за обработка на естествен език (NLP). Тази концепция е от особено значение в контекста на съвременната дигитална среда, където огромните обеми информация изискват автоматизирани и интелигентни решения за извличане на смисъл и идентифициране на тенденции. Съществуващата литература подчертава различни приложения на NLP в анализа на медийно съдържание. Например Bock и др. (2021) разглеждат как NLP може да бъде използван за динамичен анализ на тематичната значимост на съдържание, което е от ключово значение за разбирането на тенденциите в медиите. По подобен начин, Albanese и др. (2019) използват комбинация от

NLP и анализ на времеви редове, за да изследват влиянието на медийните кампании върху общественото мнение.

Дигиталната ера бележи радикални промени в медийната екосистема, характеризирани с растяща персонализация на съдържанието и алгоритмично управлявани информационни потоци. В този контекст, Тодоров (2013) изтъква, че „алгоритмите, които стоят зад съвременните платформи, не само разширяват обхвата на медийното съдържание, но и утвърждават нови форми на контрол върху аудиториите“. Тази перспектива подчертава как цифровите технологии трансформират ролята на медиите, като променят както достъпа до информация, така и механизмите на въздействие върху обществото. В друго свое изследване Тодоров (2010) акцентира, че медиите играят ключова роля в представянето на социални, икономически и политически проблеми пред аудиторията. Това изисква високо ниво на медийна интелигентност, тъй като редакторите и журналистите трябва да идентифицират важните аспекти на кризите и да ги представят обективно.

Същевременно, чрез анализи, базирани на данни и технологии, медийните организации могат да предвиждат обществените реакции и да адаптират съдържанието си, за да останат релевантни и доверени източници на информация. Съвременната медийна среда предоставя нови възможности за изследване на взаимоотношението между властта и медийните послания. В свое изследване Кулевски (2013) разкрива, че властта на медийните послания надхвърля рамките на журналистическата практика и влияе върху обществените структури и динамика.

Резултатите на Катранджиев (2008) допринасят за по-доброто разбиране на рекламните стилове и тяхната роля в медийната интелигентност. Те предлагат аналитична рамка за изследване на медийното съдържание, която може да бъде адаптирана и разширена с помощта на съвременни технологии като машинно обучение и обработка на данни. Тази интеграция е от ключово значение за развитието на медийните стратегии в дигиталната епоха.

Ангова и Вълчанов (2018) предоставят изчерпателен преглед на новите бизнес модели в онлайн медиите, като акцентират върху тяхната роля за адаптация към дигиталната среда. Техният анализ подчертава важността на иновациите, ориентирани към потребителите, и технологичните решения за устойчивостта на медийния сектор в условията на динамични промени. Тази рамка е съвместима с концепцията за медийната интелигентност, която предлага нови възможности за управление и развитие на медийните организации.

Едно от важните направления в литературата е анализът на тоналността, който дава възможност за измерване на обществените нагласи към различни теми. Shu и др. (2020) предоставят обширен преглед на методите за анализ на тоналност, подчертавайки тяхната значимост в контекста на фалшивите

новини и дезинформацията. В допълнение, Spinde (2021) разработва интердисциплинарен подход за автоматизирано откриване и визуализация на пристрастия в медийното съдържание.

Друго значимо направление в литературата е изследването на социалните мрежи като инструмент за създаване на връзки между потребители, които иначе не биха се осъществили. Тази способност на социалните мрежи да създават „мрежи от знания“ (Велев, 2012) подчертава тяхната важност за разширяване на обхвата на интелигентността на медиите. Според Велев (2012), социалните медии не само улесняват връзките между потребителите, но и позволяват гъвкавост и адаптивност в управлението на информацията. Тази уникалност прави социалните медии мощен инструмент за анализ на съдържание. В дискусиата за социалните медии като „Пета власт“ (Михайлова, 2021) се подчертава значението на новите медийни форми в трансформацията на медийната екосистема. В този контекст, медийната интелигентност трябва да обхване методологии за анализ на социалните мрежи, управление на дигитални потоци от данни и разкриване на зависимости между традиционни и нововъзникващи медийни практики. Тази перспектива подкрепя нуждата от иновативни подходи за наблюдение и адаптиране към променящите се условия, които включват изкуствен интелект и автоматизирани инструменти за анализ на съдържанието. Освен това, новите метафори, като „Петата власт“, подчертават значението на разширената роля на медиите в контекста на медийната интелигентност.

Друг значим аспект на интелигентността на медиите е анализът на визуализации, които предлагат интуитивен начин за разкриване на ключови теми и концепции. Тези подходи позволяват бързо идентифициране на доминиращи тенденции и теми в медийните публикации.

В контекста на четвъртата индустриална революция, анализът на Проданов (2018) подчертава системните промени, които икономиката на големите данни носи на обществото. Големите данни не само трансформират традиционните политикономически системи, но и стават основен ресурс за иновации и персонализация в производството и потреблението. Както отбелязва авторът, „петролът“ на XXI век са данните, което налага нови форми на регулация и управление (Проданов, 2018). Тази трансформация е от особено значение за разбирането на съвременните медийни системи, където алгоритмите, базирани на големи данни, предефинират начините на комуникация и разпространение на съдържание. Медийната интелигентност изисква не само обработка на големи обеми информация, но и разпознаване на модели и зависимости в тези данни, които често остават скрити за традиционните методи за анализ. Според Йорданова и Стефанова (2019), голяма част от данните, с които работят организациите, са неструктурирани – текстови документи, видео записи и други, което затруднява тяхното обработване с

традиционни технологии. В контекста на медиите, това означава, че платформите за анализ на големи данни трябва да се адаптират към разнообразните формати на медийното съдържание, като например анализ на текстове, изображения и видеа. Боянов (2021) акцентира върху ролята на анализа на големи данни и изкуствения интелект (AI) като основни инструменти за подобряване на процесите в IIoT (The Industrial Internet of Things). В контекста на медийната интелигентност, тези технологии могат да се използват за обработка на огромни масиви от данни от различни източници, включително социални медии, уеб съдържание и потребителски взаимодействия. AI и машинното обучение (ML) предоставят възможност за прогнозиране на тенденции и автоматизиране на решенията в реално време.

Етичните въпроси, разгледани от Христова (2023) в контекста на научните публикации, имат пряко отношение и към сферата на медийните публикации. Технологии като ChatGPT и други инструменти, подпомагани от изкуствен интелект, вече се използват за създаване на съдържание както в академични, така и в медийни среди. Основната разлика е в това, че докато научните публикации изискват строга проверка и отговорност от страна на автора, медийните публикации често се създават с акцент върху скоростта и масовото разпространение на информация. Подобно на научните текстове, медийното съдържание, създадено с помощта на изкуствен интелект, може да предизвика въпроси за автентичността и отговорността. Христова (2023) подчертава, че „чатботът не може да изпълнява критерия за авторство“, тъй като липсва способността му да носи отговорност за съдържанието (Христова, 2023). Тази теза е особено релевантна за медийните публикации, където прозрачността относно използването на ИИ е от решаващо значение за доверието на аудиторията.

Изследване на Кирилов (2019) подчертава важноста на информационните технологии за ефективно управление на сложни процеси, като тези в практическото обучение. Аналогично, медиите се изправят пред сходни предизвикателства, свързани с управлението на данни и персонализацията на съдържанието. Интеграцията на тези принципи в медийната интелигентност подчертава значимостта на дигиталните решения за адаптиране към динамичната среда на съвременната информация. Развитието на дигиталното образование, както е анализирано от Казакова (2023), има значителна връзка с концепцията за медийна интелигентност и трансформацията на медийната екосистема в съвременната цифрова среда. Според Казакова (2023), дигитализацията на образованието е не само необходима адаптация към научно-техническия прогрес, но и средство за изграждане на ключови дигитални умения и компетенции, които са основополагащи за разбирането и критичното възприемане на медийното съдържание.

Независимо от прогреса в литературата, съществуват значителни предизвикателства, включително технически ограничения, свързани с обработката на

големи обеми данни, и необходимостта от по-ефективни алгоритми за контекстуален анализ. Настоящото изследване се опира на тези основи, за да предложи рамка за интегриран анализ, която съчетава различни подходи за разбиране на медийното съдържание. Изследването използва концепцията за интелигентност на медиите като интегративен подход, комбиниращ съвременни технологии и методи за анализ на медийно съдържание. Подходът е базиран на обработка на големи данни и технологии за обработка на естествен език (NLP), които позволяват извличане на смисъл, идентифициране на тенденции и визуализация на ключови теми в дигиталното медийно пространство.

Методологията на изследването включва три ключови етапа: събиране на данни, обработка на данни и анализ на данни (фигура 1). Тази методология представлява общоприета рамка, която обхваща етапите на работа с данни в различни научни и приложни дисциплини. Използвани са данни от публикации в онлайн платформи на водещи телевизионни медии в България, като подходът е насочен към постигане на по-дълбоко разбиране на динамиките в дигиталната медийна среда.



*Източник:* Авторска разработка, базирана на общоприети подходи

**Фигура 1:** Методология на изследването

На първия етап (събиране на данни) се използват техники за web scraping, което се постига чрез имплементирането на Java базирани инструменти. Използва се платформата JavaFX за създаване на приложение за извличане на HTML съдържание. За допълнителна обработка (парсване) на HTML съдържанието се използва разработената от Hedley и др. (2023) библиотека jsoup. Събрани са публично достъпните публикации за 2024 г. от официалните новинарски уебсайтове на седем водещи български телевизии. Данните включват заглавия, дати и текстове на публикациите, като общият брой събрани публикации е 195330.

На втория етап (обработка на данни) събраните данни се почистват и структурират в удобен за анализ формат. Премахват се дублирани записи, както и публикации с липсваща или непълна информация. Извършено е и

нормализиране на текстовете, което включва преобразуване в малки букви, отстраняване на пунктуации и др. На този етап присъстват и задачи, свързани с моделирането на данните за целите на аналитичната обработка.

На третия етап (анализ на данни) се прилагат разнообразни аналитични техники, включващи подходи за изготвянето на различни количествени анализи, както и визуализации на най-често срещаните думи за идентифициране на тенденции в данните. За целите на визуализациите на количествените анализи се използват наличните в Microsoft Excel възможности за създаване на графики.

Ограниченията на изследването включват технически и контекстуални ограничения при обработката на данни, както и специфични особености на разглежданите медийни източници. Анализирани са уеб публикации от избрани за целите на изследването телевизионни медии, което означава, че резултатите може да не са представителни за цялостната медийна среда в България. Също така, количествените методи подчертават предимно честота и наличие на думи, без да отчитат контекста, което в някои случаи може да доведе до повърхностни интерпретации. Въпреки посочените ограничения, методологията предлага основа за по-задълбочени, съпроводени от качествен анализ, бъдещи изследвания, които да разширят разбирането за взаимодействието между медийното съдържание, потребителите и технологиите.

## Резултати

Избраните за целите на изследването телевизии и техните уебсайтове представляват ключови медии в България, които покриват разнообразни теми и категории новини. Те включват национални телевизии като BNT, bTV, Nova TV и специализирани канали като Bloomberg TV и Euronews Bulgaria, което гарантира широка палитра от гледни точки. Сайтовете обхващат телевизии с национално покритие, които информират за събития в цялата страна и са признати източници за новинарско съдържание. Избраните сайтове предлагат новини в различни категории, включително политика, икономика, общество, спорт и култура, което осигурява разнообразие за количествени и качествени анализи. Всички избрани сайтове поддържат активни платформи за новинарски публикации, които са леснодостъпни за аудиторията. Това ги прави подходящи за анализ на онлайн съдържание. Bloomberg TV предоставя икономически и финансови новини, Euronews Bulgaria е фокусирана върху европейската перспектива, а Nova и bTV са водещи в националните новини. Това позволява сравнение на различни журналистически подходи и теми. Телевизиите, чиито сайтове са избрани, са утвърдени институции в медийната среда на България. Те следват етични журналистически стандарти, което гарантира качество и достоверност на публикуваното съдържание. Избраните уебсайтове са сред най-посещава-

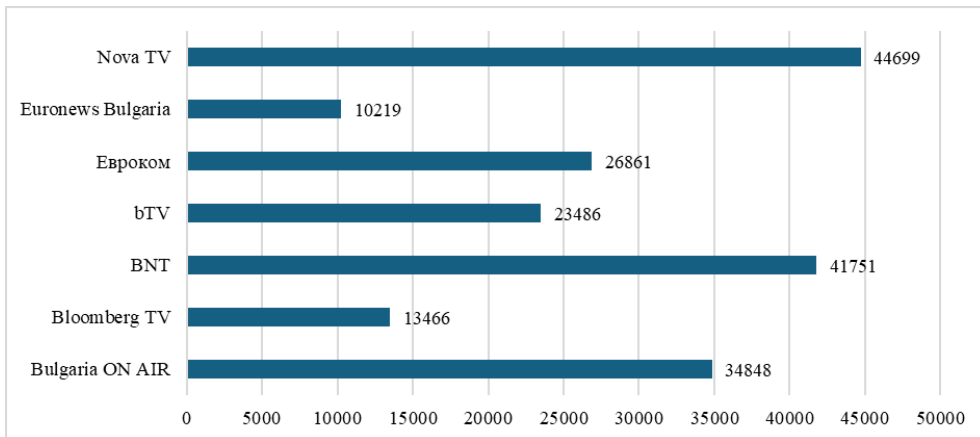
ните източници на новини в България. Техният онлайн трафик, популярност и въздействие върху общественото мнение правят анализа на съдържанието им релевантен за медийното изследване. Анализираните медии са представени в таблица 1, като периодът, който обхваща изследването, е 2024 г.

**Таблица 1:** Изследвани онлайн медии, подредени по азбучен ред на техния линк

| Наименование      | Линк                                                                    | Брой публикации |
|-------------------|-------------------------------------------------------------------------|-----------------|
| Bulgaria ON AIR   | <a href="https://bgonair.bg">https://bgonair.bg</a>                     | 34848           |
| Bloomberg TV      | <a href="https://bloombergtv.bg">https://bloombergtv.bg</a>             | 13466           |
| BNT               | <a href="https://bntnews.bg">https://bntnews.bg</a>                     | 41751           |
| bTV               | <a href="https://btvnovinite.bg">https://btvnovinite.bg</a>             | 23486           |
| Евроком           | <a href="https://eurocom.bg">https://eurocom.bg</a>                     | 26861           |
| Euronews Bulgaria | <a href="https://euronewsbulgaria.com">https://euronewsbulgaria.com</a> | 10219           |
| Nova TV           | <a href="https://nova.bg/news">https://nova.bg/news</a>                 | 44699           |

*Източник:* Авторско проучване

На фигура 2 е представено количественото разпределение на уеб публикациите по медии.



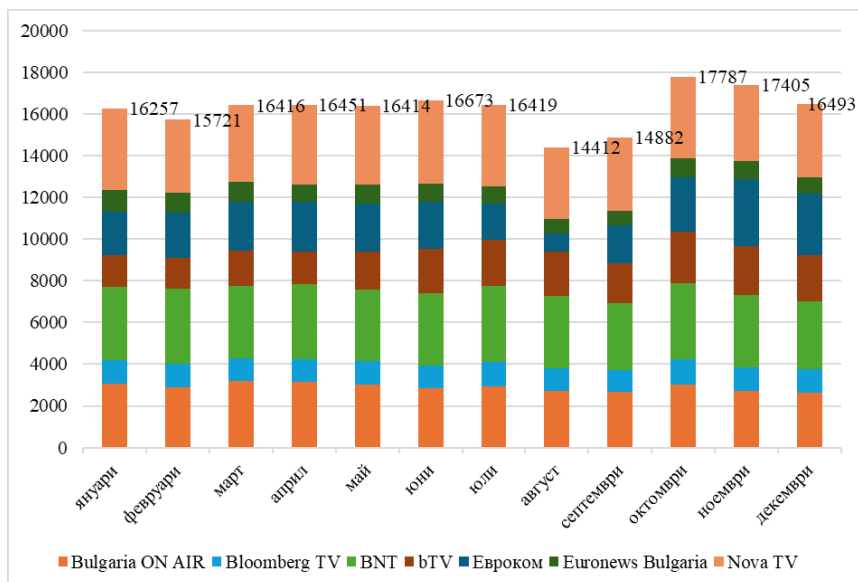
*Източник:* Авторско проучване

**Фигура 2:** Количествено разпределение на уеб публикациите по медии

Анализът на публикациите от избраните телевизионни медии в България за 2024 година разкрива значителни количествени и качествени характеристики на медийното съдържание. Общият брой публикации варира значи-

телно между медиите, като най-голям е обемът при Nova TV и BNT, което подчертава водещата им роля в предоставянето на новинарско съдържание. Bulgaria ON AIR, bTV и Евроком също демонстрират значителна активност, докато Euronews Bulgaria и Bloomberg TV са с по-нисък общ брой публикации, което е отражение на тяхната по-специализирана тематика.

На фигура 3 е представено количественото разпределение на уеб публикациите по месеци.

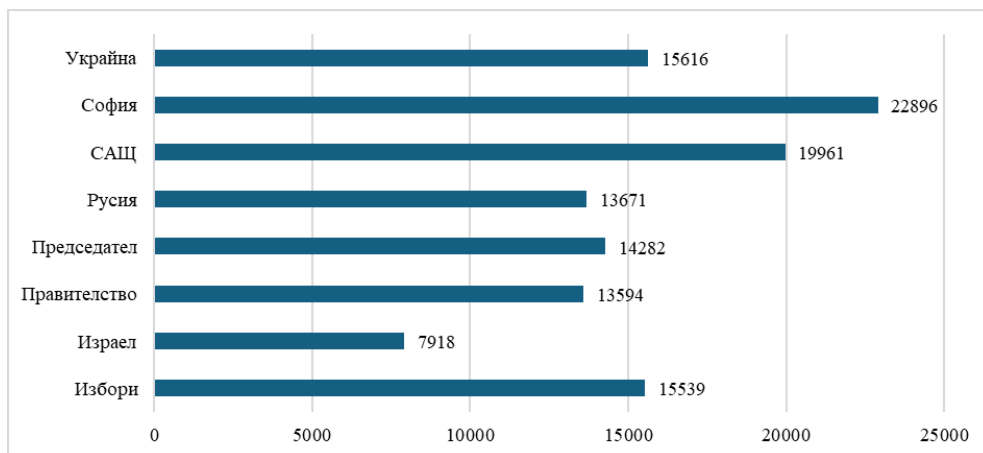


Източник: Авторско проучване

**Фигура 3:** Количествено разпределение на уеб публикациите по месеци

Динамиката в броя на публикациите през годината показва сезонни колебания. През летните месеци август и септември се наблюдава намаление на публикуваното съдържание в някои медии, докато през октомври и ноември обемът на публикациите се увеличава, вероятно в резултат на политически или социално значими събития. Тази тенденция е характерна за всички медии, като при някои е по-слабо, а при други е по-силно изразена.

На фигура 4 е представено количественото разпределение на най-често срещаните думи в уеб публикациите. При анализа на най-често срещаните думи от изследваните медийни публикации, политическите партии са изключени от резултатите. Тази стъпка е предприета с цел да се избегне политически фокус и да се запази основната цел на изследването – изследване на медийната екосистема чрез аналитични подходи, насочени към разбиране на общите теми и тенденции в дигиталното пространство.



Източник: Авторско проучване

**Фигура 4:** Количествено разпределение на най-често срещаните думи в уеб публикациите

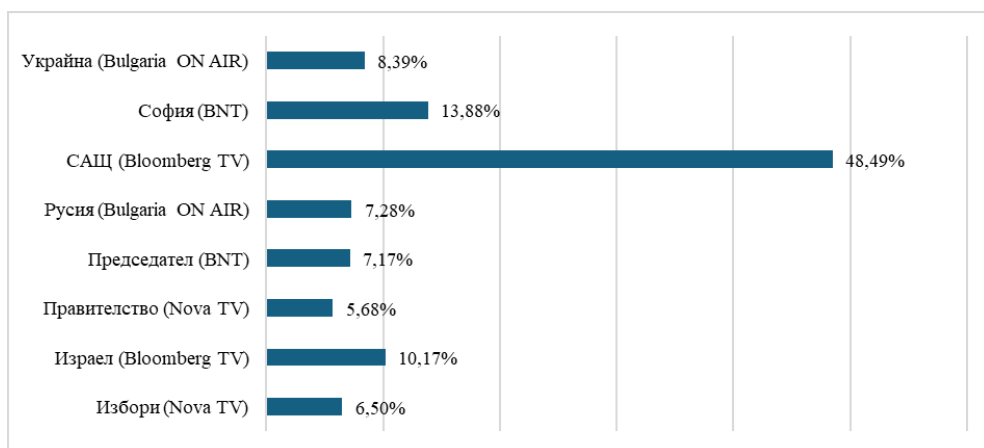
Най-често срещаните думи в медийното съдържание разкриват ключовите теми, които са доминирали в новинарския дневен ред. Думата „София“ е най-често използвана, което потвърждава ролята на столицата като основен център на политически, социални и икономически събития. Международните теми също заемат водещо място, като думите „САЩ“, „Русия“ и „Украйна“ отразяват глобалното влияние на ситуацията в Украйна, дипломатическите отношения и международната политика. В следващата таблица е представена честотата на най-често срещаните думи в съответните медии (таблица 2).

**Таблица 2:** Най-често срещани думи в съответните медии, подредени по азбучен ред

| Дума         | Медия           | Брой публикации с тази дума в медията |
|--------------|-----------------|---------------------------------------|
| Избори       | Nova TV         | 2907                                  |
| Израел       | Bloomberg TV    | 1369                                  |
| Правителство | Nova TV         | 2537                                  |
| Председател  | BNT             | 2993                                  |
| Русия        | Bulgaria ON AIR | 2537                                  |
| САЩ          | Bloomberg TV    | 6529                                  |
| София        | BNT             | 5796                                  |
| Украйна      | Bulgaria ON AIR | 2925                                  |

Източник: Авторско проучване

На фигура 5 е представено процентното разпределение на най-често срещаните думи в уеб публикациите спрямо общия брой публикации в съответната медия.



*Източник:* Авторско проучване

**Фигура 5:** Процентно разпределение на най-често срещаните думи в уеб публикациите спрямо общия брой публикации в съответната медия

Процентното разпределение на най-често срещаните думи спрямо общия брой публикации в съответните медии показва специфични акценти в съдържанието на всяка медия. Например Bloomberg TV акцентира върху „САЩ“, докато Bulgaria ON AIR и BNT поставят акцент върху думи като „Украйна“ и „София“.

## Дискусия

Резултатите от анализа подчертават важноста на количествения и качествения подход при изследването на медийното съдържание. Големият обем публикации в медиите, като Nova TV и BNT, отразява активната им роля в информирането на обществото, докато по-ниският обем на публикации в Bloomberg TV и Euronews Bulgaria може да бъде обяснен с тяхната специализация в определени тематични области. Тези различия могат да се използват за идентифициране на медийни стратегии и предпочитания на аудиторията.

Сезонните колебания в броя на публикациите отразяват динамиката на общественения дневен ред. Летният спад, през месеците август и септември, в медиите и увеличението през есента, в месеците октомври и ноември, съвпадат с ключови политически и социални събития, като избори и обществе-

ни дебати. Това демонстрира как медийното съдържание е тясно свързано със сезонни и обществени фактори, които влияят върху обема и тематиката на публикуваните материали.

Анализът на най-често срещаните думи разкрива значими тематични приоритети. Използването на думите „София“, „САЩ“, „Русия“ и „Украйна“ показва фокуса на медиите както върху вътрешнополитически, така и върху международни теми. Възможно е честотата на използване на тези термини да бъде индикатор за политическите, икономическите и социалните процеси, които привличат общественото внимание.

Процентното разпределение на ключовите думи спрямо общия брой публикации допълнително подчертава уникалния характер на всяка медия. Bloomberg TV, със значително внимание върху „САЩ“, демонстрира специфичен международен акцент, докато Bulgaria ON AIR и BNT акцентират върху регионални и национални теми. Тази специфика може да бъде използвана за по-задълбочено разбиране на редакционните политики и аудиторията на различните медии.

## **Заключение**

Настоящото изследване представя анализ на публикациите в онлайн медиите на водещи телевизионни канали в България за 2024 г. Използвайки иновативни методи за извличане и анализ на данни, проучването разкрива ключовите тематични акценти, динамиката в публикуването и редакционните приоритети на медиите. Резултатите подчертават значението на теми като „София“, „САЩ“, „Украйна“, „Русия“ и „Избори“, които се появяват с висока честота и обхващат широк спектър от политически, икономически и социални въпроси. Анализът на динамиката в публикуването показва, че броят на публикациите варира значително между медиите и сезоните, като пиковите моменти съвпадат с важни събития или политически кампании. Използването на различни методи за анализ демонстрира потенциала на автоматизирания подход за разбиране на големи обеми от данни. Това не само разкрива съдържателните акценти в медийното пространство, но и предоставя възможности за по-задълбочено изследване на социалните и политическите измерения на медиите. Настоящото изследване има няколко важни приноса. То предлага структурирана методология за изследване на онлайн медийното съдържание, която може да се приложи в различни контексти. Освен това резултатите подчертават значението на онлайн медиите като източник на информация и платформа за обществени дебати. Това е особено важно в съвременната дигитална ера, където медиите играят ключова роля за формиране на общественото мнение.

В заключение, изследването допринася за по-доброто разбиране на динамиката в онлайн медиите и техния дневен ред, като същевременно подчертава необходимостта от по-нататъшни изследвания, които да разгледат въздействието на медийното съдържание върху аудиторията и общественото мнение.

### **Спонсориране на научното изследване**

Публикацията съдържа резултати от изследване, финансирано със средства от целева субсидия за НИД на УНСС по договор № НИД НИ-22/2024/А.

### **Използвана литература**

- Велев, Д. (2012). Изследване на възможностите на онлайн социалните мрежи при управление на извънредни ситуации, Научни трудове на УНСС, (3), с. 87-118. (Velev, D., 2012, Izsledvane na vazmozhnostite na onlayn sotsialnite mrezhi pri upravlenie na izvanredni situatsii, Nauchni trudove na UNSS, (3), s. 87-118).
- Йорданова, С. & Стефанова, К. (2019). Предизвикателствата на големите данни – същност, характеристики и технологии, Научни трудове на УНСС, (1), с. 13-31. (Yordanova, S. & Stefanova, K., 2019, Predizvikatelstvata na golemite dannī–sashtnost, harakteristiki i tehnologii. Nauchni trudove na UNSS, (1), s. 13-31).
- Казакова, М. (2023). Възможности и тенденции за развитие на дигиталното образование, Научни трудове на УНСС, (5), с. 153-166. (Kazakova, M., 2023, Vazmozhnosti i tendentsii za razvitie na digitalното obrazovanie, Nauchni trudove na UNSS, (5), s. 153-166).
- Катранджиев, Х. (2008). Класификация на рекламните стилове у нас – резултати от контент анализа на рекламната в българските медии, Годишник на УНСС, (2), с. 343-382. (Katrandzhiev, H., 2008, Klasifikatsia na reklamnite stilove u nas – rezultati ot kontent analiza na reklamata v balgarskite medii, Godishnik na UNSS, (2), s. 343-382).
- Кирилов, Р. (2019). Информационни проблеми на практическото обучение, Научни трудове на УНСС, (2), с. 185-196. (Kirilov, R., 2019, Informatsionni problemi na prakticheskoto obuchenie, Nauchni trudove na UNSS, (2), s. 185-196).
- Кулевски, Ц. (2013). Властта на медийното послание, Научни трудове на УНСС, (3), с. 79-110. (Kulevski, Ts., 2013, Vlastta na mediynoto poslanie, Nauchni trudove na UNSS, (3), s. 79-110).
- Михайлова, К. (2021). Социални факти и смислови полета на метафората „Четвъртата власт“, Научни трудове на УНСС, (4), с. 43-72. (Mihaylova,

- К., 2021, *Sotsialni fakti i smislovi poleta na metaforata „Chetvartata vlast“*, Nauchni trudove na UNSS, (4), s. 43-72).
- Проданов, Х. (2018). Четвъртата индустриална революция и политическата икономика на големите данни, Научни трудове на УНСС, (3), с. 223-251. (Prodanov, H., 2018, *Chetvartata industrialna revolyutsia i politicheskata ikonomia na golemite dannii*, Nauchni trudove na UNSS, (3), s. 223-251).
- Тодоров, П. (2013). Власт на медиите и власт над медиите, Научни трудове на УНСС, (1), с. 11-53. (Todorov, P., 2013, *Vlast na mediite i vlast nad mediite*, Nauchni trudove na UNSS, (1), s. 11-53).
- Тодоров, П. (2010). Медиите в кризата и кризата в медиите, Годишник на УНСС, (1), с. 129-172. (Todorov, P., 2010, *Mediite v krizata i krizata v mediite*, Godishnik na UNSS, (1), s. 129-172).
- Христова, М. (2023). Преглед на публикационните политики на водещи издателства в Европа и Америка във връзка с използването на технологии, подпомагани от изкуствен интелект, Научни трудове на УНСС, (4), с. 131-146. (Hristova, M., 2023, *Pregled na publikatsionnite politiki na vodeshti izdatelstva v Evropa i Amerika vav vrazka s izpolzvaneto na tehnologii, podpomagani ot izkustven intelekt*, Nauchni trudove na UNSS, (4), s. 131-146).
- Albanese, F., Pinto, S., Semeshenko, V., & Balenzuela, P. (2020). Analyzing mass media influence using natural language processing and time series analysis, *Journal of Physics: Complexity*, 1(2), 025005.
- Angova, S. & Valchanov, I. (2018). New online media business models, *Yearbook of UNWE*, (1), pp. 287-306.
- Bock, A., Palladino, A., Smith-Heisters, S., Boardman, I., Pellegrini, E., Bienenstock, E. J. & Valenti, A. (2021). An NLP approach to quantify dynamic salience of predefined topics in a text corpus, arXiv preprint arXiv:2108.07345.
- Boyanov, L. (2021). A Conceptual Approach for Industrial Internet of Things Assessment, *Yearbook of UNWE*, (2), pp. 97-107.
- Hedley, J., & jsoup contributors. (2023). jsoup: Java HTML Parser (Version 1.17.2), <https://jsoup.org>
- Shu, K., Mahudeswaran, D., Wang, S., Lee, D. & Liu, H. (2020). Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media, *Big data*, 8(3), pp. 171-188.
- Spinde, T. (2021). An interdisciplinary approach for the automated detection and visualization of media bias in news articles, In 2021 International Conference on Data Mining Workshops (ICDMW), IEEE, pp. 1096-1103.

\*\*\*

## UNDERSTANDING WEB PUBLICATIONS IN THE DIGITAL AGE BY APPLYING MEDIA INTELLIGENCE

Assoc. Prof. Plamen Milev, PhD

Department of Information Technologies and Communications

Faculty of Applied Informatics and Statistics

University of National and World Economy

*e-mail: pmilev@unwe.bg*

### Abstract

*The article explores approaches for analyzing online publications from leading television media in Bulgaria, utilizing modern big data processing methods. By combining quantitative and qualitative analytical techniques, it demonstrates the potential for extracting meaning and identifying trends in the digital media landscape. The study highlights the importance of the concept of media intelligence in analyzing online content and contributes to a better understanding of transformations within the media ecosystem. These approaches also offer practical guidelines for improving the analysis and management of media content.*

**Keywords:** data analysis, online publications, media intelligence, big data, digital media ecosystem

**JEL:** C88, L86