

ГЕНДЕРНИЙ АНАЛІЗ ЗАСОБІВ МАСОВОЇ КОМУНІКАЦІЇ

Філатьєв Данило

Науковий керівник: д. іст. наук, професор Пасічник Н. О.

*Центральноукраїнський державний педагогічний університет
імені Володимира Винниченка, м. Кропивницький, Україна*

У статті досліджуються проблеми гендерної чутливості популярних та авторитетних українських засобів масової інформації. Показано, що проблема гендерної нерівності відображена в тематиці публікацій, експертних публікаціях, складі редакційних колегій, змістовних характеристиках тексту тощо. Також аналізується редакційний та експертний склад популярних медійних ресурсів та виокремлюється значущість гендерних ознак у поданих новинах. Дані дослідження є важливими для кластеризації засобів масової інформації та подальшого поетапного дослідження кластерів.

Ключові слова: гендерний склад, кластерний аналіз, змістовна значущість, гендерний паритет.

GENDER ANALYSIS OF MASS MEDIA

Danylo Filatiev

Scientific supervisor: Doctor of History, Professor N. Pasichnyk

*Volodymyr Vynnychenko Central Ukrainian State Pedagogical University,
Kropyvnytsky, Ukraine*

The article examines the problems of gender sensitivity of the popular and authoritative Ukrainian mass media. It is shown that the problem of gender inequality is reflected in the topics of the publications, expert publications, composition of editorial boards, content characteristics of the text, etc. The article also analyses the editorial and expert composition of the popular media resources and highlights the importance of gender features in the submitted news. These studies are important for the clustering of media and further gradual cluster research.

Key words: gender composition, correlation analysis, cluster analysis, semantic significance, gender parity.

Сучасне суспільство в процесі свого розвитку породжує величезний обсяг текстової інформації, що відображає різні аспекти його діяльності. Цілий пласт цієї інформації відноситься до так званої новинної інформації – повідомлення

про поточні події, що відбуваються в реальному світі. Зазвичай, новинну інформацію формують засоби масової інформації (ЗМІ), оскільки їхня основна діяльність – це інформування суспільства про всі події, що відбуваються у світі.

Згідно досліджень Державної служби статистики, в Україні 53,6% населення становлять жінки, проте незважаючи на такі кількісні показники, українські жінки обмежено відображені в медіа. Інститут масової інформації провів моніторингове дослідження гендерного балансу популярних українських онлайн-медіа, яке засвідчує, що жінки як експертки коментують майже кожен п'ятий матеріал в онлайн-медіа (19% загальної кількості експертних коментарів, 81% експертних коментарів у медіа припадає на чоловіків). Жінки виступають героїнями матеріалів ЗМІ у 29% від загальної кількості цієї тематики, проте вони згадуються переважно в матеріалах про шоу-біз, кримінал або в «жовтих» новинах.

Сфера мас-медіа як один із найважливіших соціальних інститутів, що виконують функції передачі та поширення інформації, формування цінностей і норм від одного покоління до іншого через інтеграцію гендерних підходів, спроможна закласти основи для становлення суспільної гендерної культури. Зокрема, цікавим є дослідження інформації, що подають новинні агрегатори, які інтегрують новини з тисяч джерел в «новинні кластери», надаючи користувачам доступ до ранжованих за значенням і тематикою сюжетів, і є одними з найбільш популярних ресурсів. У цих інформаційних ресурсах реалізуються різні стратегії уявлення новинної інформації користувачам. Ключовою проблемою новинних агрегаторів є вирішення задачі кластеризації – формування груп близьких документів, що моделюють новинні сюжети. На практиці результат автоматичної кластеризації може бути потім змінений в залежності від корпоративної політики (в тому числі і вручну редакторами) чи даних про запити користувачів або обговорень в блогах тощо.

Проблемі гендерної проблематики в ЗМІ присвячені праці вітчизняних вчених Т.Бурейчак, Ю.Галустяна, О.Кісь, а також зарубіжних дослідників – У. Е. Кілборн, Л. Дж. Басбі, Г. Лайчті, Г. Саллівана та ін. Різноманітні

лінгвістичні аспекти текстів з метою виявлення гендерної складової, використання фемінітивів відображено у працях В. Зірки, В. Слінчук, О. Сидоренко та ін. Використовувані на практиці методи кластеризації, що залежать від великого числа параметрів, налаштованих на реальні новинні потоки, досліджуються в працях [1–4], а у роботі [5] наведені результати дослідження мір змістовної значущості різних понять, що є необхідним для побудови більш точних кластеризацій в галузі ЗМІ. В [6] досліджуються гендерні ресурси українських мас-медіа тощо. Проте, майже відсутні публікації з проблеми дослідження гендерних стереотипів, гендерної нерівності в доступних колекціях новинних повідомлень та досліджень у відповідних кластерах ЗМІ.

Метою статті є здійснення аналізу редакційного та експертного складів популярних медійних ресурсів на предмет гендерної чутливості, проаналізувати з точки зору популярності видання, його авторитетності та дослідження характеристик смислового значення елементів простору гендерних ознак, що є необхідним для побудови кластеризації ЗМІ та подальшого дослідження кластерів.

Загальне вирішення задачі кластеризації текстів може бути зведене до наступних етапів: 1. Визначення об'єктів простору ознак, під яким будемо розуміти безліч значимих слів і словосполучень, що визначають зміст текстів в масиві. 2. Обчислення заходи смислової значущості для кожного поняття документа. 3. Обчислення значень міри схожості між текстами документа. 4. Застосування методів кластерного аналізу для створення груп схожих текстів. 5. Перевірка істинності результатів кластерного рішення.

Відповідно до цих етапами в першу чергу потрібно визначити множину слів і словосполучень, за якими будуть оцінюватися тексти. Потім необхідно визначити, за якими критеріями потрібно оцінювати смислове значення кожного елемента простору ознак, і відповідно до цих критеріїв призначити кожному слову або словосполученню ваговий коефіцієнт їх смислової значущості в просторі ознак [5].

Дуже поширеними є метричні алгоритми кластеризації [4], що враховують особливості новинних повнотекстових документів, що досягнуто векторизацією завдання – представимо документи у вигляді векторів в метричному просторі і введемо міру близькості. Міру близькості між документами часто розглядають як косинус між нормалізованими векторами ознак документів:

$$\mu(d_i, d_j) = \sum_m \frac{d_{im}}{\|d_i\|} \cdot \frac{d_{jm}}{\|d_j\|},$$

де матрицю $M = \{\mu_{ij}\} = \{\mu(d_i, d_j)\}$ називають матрицею близькості між документами.

Міру близькості між різними кластерами, як окремий випадок, близькість між окремим документом і кластером, можна визначити різними способами

1. Відстань ближнього сусіда: $\mu_{\max}(C_k, C_l) = \max_{\substack{d_i \in C_k \\ d_j \in C_l}} \mu(d_i, d_j)$.
2. Відстань далекого сусіда: $\mu_{\min}(C_k, C_l) = \min_{\substack{d_i \in C_k \\ d_j \in C_l}} \mu(d_i, d_j)$.
3. Групова середня відстань: $\mu_{\text{average}}(C_k, C_l) = \frac{1}{|C_k|} \cdot \frac{1}{|C_l|} \cdot \sum_{\substack{d_i \in C_k \\ d_j \in C_l}} \mu(d_i, d_j)$.
4. Групова відстань між нормалізованими центрами кластерів:

$$\mu_{\text{center}}(C_k, C_l) = \mu(d_*(C_k), d_*(C_l)), \text{ де, наприклад, } d_*(C_k) = \sum_{d_i \in C_k} \frac{d_i}{\left\| \sum_{d_j \in C_k} d_j \right\|},$$

В якості запобіжної смислової значущості слів і словосполучень часто використовується так звана статистична міра $TF - IDF$ [5]:

TF – відношення числа входжень найменування поняття до загальної кількості найменувань понять документа. Таким чином оцінюється важливість найменування поняття t_i в межах окремого документа: $tf(t, d) = \frac{n_t}{\sum_k n_k}$, де n_k – число входжень найменування поняття t в документ; $\sum_k n_k$ – загальне число найменувань понять в даному документі; IDF – інверсія частоти, з якою деяке

найменування поняття зустрічається в документах колекції. Для кожного унікального слова в межах конкретної колекції документів існує тільки одне значення IDF: $idf(t, D) = \log \frac{|D|}{\sum_k n_k}$, де $|D|$ – число документів у колекції;

$|\{d_i \in |D| t \in d_i\}|$ – число документів з колекції D , в яких зустрічається t (якщо $n \neq 0$).

Міра $TF - IDF$ є добутком двох множників:

$$tf - idf(t, d, D) = tf(t, d) \cdot idf(t, D).$$

Для побудова різних статистичних моделей потрібно провести певні дослідження, зокрема, проаналізовано склад багатьох популярних українських медіаресурсів на предмет представлення жінок в редакції, досить часто ця інформація є відсутньою.

Таблиця 1. Гендерний склад редакцій українських ЗМІ

Медіаресурс	Редакційний склад	
	з них жінок	з них чоловіків
Фокус	33%	67%
delo.ua/	41%	59%
Політека	43%	57%
dengi.ua	50%	50%
Ліга-нет	56%	44%
comments.ua	56%	44%
Кореспондент	56%	44%
УНІАН	59%	41%
РБК Україна	59%	41%
Главред	59%	41%
Новий час	59%	41%
24 канал	60%	40%
obozrevatel.com	60%	40%
СВН – Центральноукраїнське бюро новин	60%	40%
Форбс Україна	67%	33%
Мінфін	67%	33%

vesti.ua	68%	32%
Дзеркало тижня	69%	31%
Лівий берег	71%	29%
mind.ua	77%	23%

За проведеним дослідженням можна навести такі факти [7-8]

- Чоловіки домінують в засобах масової інформації на всіх платформах – в більшості випадків вони визначають напрям розвитку видання, його спеціалізацію, займають високі посади в редакції.
- Поступово журналістика стає «жіночою» професією, в більшості редакцій переважає жіночий склад працівників.
- В редакціях чоловіки займаються такими напрямками спорт, міжнародна політика, економіка, технології тощо.
- Абсолютне панування чоловіків на даний час в таких посадах як фотокореспондент тощо.

Більш детально про тематику публікацій

Таблиця 2. Гендерний аналіз авторів по основним рубрикам

Теми	Чоловіки	Жінки
Світова політика	86%	14%
Технології	83%	17%
Спорт	78%	22%
Економіка і бізнес	72%	28%
Правосуддя і злочинність	46%	54%
Внутрішня політика	45%	55%
Соціальні проблеми	44%	56%
Освіта	34%	66%
Наука	33%	67%
Культура	32%	68%
Розваги	28%	72%
Релігія	24%	76%
Лайфстайл	24%	76%
Погода	21%	79%
Здоров'я	16%	84%

Дослідження гендерних пропорції експертів в різних сферах в публікаціях українських ЗМІ є наступними. За твердження українського Інституту масової

інформації присутність жінок-експерток в онлайн-ЗМІ є меншою ніж чоловіків, а саме 19% жінок-експертів проти 81% чоловіків. Коефіцієнт присутності в українських інтернет-ЗМІ експерток залишається відносно стабільним, якщо порівнювати від першого кварталу 2020 року, але в порівнянні з 2018-2019 роками, то рівень присутності жінок-експерток дещо впав, а саме він становив 24%. Найбільш часто в своїх публікаціях експерток згадують такі онлайн ресурси як ТСН (31%), 24 канал (27%), УНІАН та НВ (по 23%).

Якщо в матеріалах про міжнародну політику жінок-експертів залучали в кожній п'ятій публікації, то внутрішню політику коментували жінки-експертки вже у 40% всіх матеріалів. Але якщо говорити про такі галузі, як здоров'я, освіта, психологія чи діти, то тут рівень представленості жінок становить більше половини публікацій. Напрями, де жінки традиційно низько представлені – це кримінальна хроніка, бізнес та економіка, спорт та технології.

Також в матеріалах ЗМІ історій про жінок-героїнь є втричі меншим, ніж про чоловіків, причому ці показники приблизно становлять 28% героїнь проти 72% героїв. В цьому питанні гендерного паритету дотримуються наступні видання "Політека" (46%), ТСН (44%) та 24 канал (37%). Проте на жаль в цьому напрямку до гендерного балансу ще далеко, крім того, в багатьох матеріалах, зокрема, в рубриках шоубізнес, жовта преса жінки згадуються в сексистських та зневажливих патернах, причому близько третини публікацій стосувалися шоубізнесу, де найбільш часто згадували іноземних жінок-популярностей, причому більшість текстів мають сексистське спрямування текстів.

Якщо говорити про зовнішню політику, то в матеріалах ЗМІ найчастіше згадували канцлера ФРН Ангелу Меркель, переможницю білоруських президентських виборів Марію Тихановську. У публікаціях, що висвітлюють внутрішню політику – героїнями матеріалів найчастіше є лідерка політичної партії «Батьківщина» Юлія Тимошенко, народні депутатки від "Європейської солідарності" Ірина Геращенко, від "Слуги народу" Ірину Верещук тощо. Такі матеріали складала близько 15%.

Проте в матеріалах, що стосуються кримінальної хроніки, надзвичайних ситуацій чи дорожньо-транспортних пригод, коефіцієнт згадування жінок є меншим – близько 10%, і в більшості таких матеріалах жінки виступають в ролі жертв чи потерпілих. В матеріалах з таких питань як, суспільство, наука, права людини, здоров'я чи культура жінки представлені лише в 15% публікацій.

Якщо проаналізувати динаміку наповнення новин на сайті Укр.нет для дослідження смислових значень елементів простору гендерних ознак для подальших досліджень.

Таблиця 3. Частота гендерних ознак на прикладі агрегатора Укр.нет

	Жінки	Чоловіки	Всього понять
Частота поняття в заголовках	28	59	149

Оцінимо важливість найменувань гендерних понять, маємо

$$TF(female, d) = \frac{28}{149} = 0.19 \text{ та } TF(male, d) = \frac{59}{149} = 0.4,$$

$$IDF(female, D) = \ln \frac{|D|}{|\{d_i \in |D| t \in d_i\}|} = \ln \frac{115}{28} = 0.61,$$

$$IDF(male, D) = \ln \frac{|D|}{|\{d_i \in |D| t \in d_i\}|} = \ln \frac{115}{59} = 0.29,$$

тоді

$$TF - IDF(female, d, D) = TF(t, d) \cdot IDF(t, D) = 0.12 \text{ та}$$

$$TF - IDF(male, d, D) = TF(t, d) \cdot IDF(t, D) = 0.12.$$

Таким чином, дані дослідження показали, що велику вагу в $TF - IDF$ отримують найменування понять з високою частотою зустрічальності в конкретному документі і з відносно низькою частотою в межах всієї колекції текстів. Як бачимо, в цілому дані коефіцієнти показують однакову важливість гендерних ознак у новинах ЗМІ. Дані дослідження є основою для подальшої побудови різних моделей кластеризації в галузі ЗМІ та подальшого гендерного аналізу отриманих кластерів.

Список літератури

1. Андреев А.М., Березкин Д.В., Морозов В.В., Симаков К.В. Метод кластеризации документов текстовых коллекций и синтеза аннотаций кластеров // Труды 10-й Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» RCDL'2008. – Дубна, 2008. – С. 220-229.

2. Антонов А.В., Баглей С.Г., Мешков В.С. Подход к выявлению подмножеств похожих документов // Труды 10-й Всерос. науч. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» RCDL'2008. – Дубна, 2008. – С. 197-199.

3. Васильев В.Г., Кривенко М.П. Методы автоматизированной обработки текстов. – М.: ИПИ РАН, 2008. – 302 с.

4. Воронцов К.В. Лекции по алгоритмам кластеризации и многомерного шкалирования. – www.ccas.ru/voron/download/Clustering.pdf.

5. Захаров В. Н., Мусабаев Р. Р., Красовицкий А. М., Козловская Я. Д., Хорошилов Алдр А., Хорошилов Ал-ей А., Метод кластеризации новостных сообщений средств массовой информации на основе их концептуального анализа, Системы и средства информ., 2019, том 29, выпуск 3, 52–65

6. Гендерні ресурси українських мас-медіа: ціна і якість: Практичний посібник для журналістів та працівників ЗМІ./Н.М.Сидоренко, М.М.Скорик, та ін.; ред-упоряд. Н.М.Сидоренко. – К.:”К.І.С.”, 2004. – 76 с.

7. Жарова Л. Макроекономічний вимір гендерної рівності. Гендер і гроші. №31 (42). 2017. С.11–15

8. Гендерні медійні практики: Навчальний посібник із гендерної рівності та недискримінації для студентів вищих навчальних закладів /Колектив авторів. – Київ, 2014. – 206 с.