

ORIGINAL RESEARCH ARTICLE

Decoding global reproductive health discourse on Reddit: themes, regions, and misinformation challenges

DOI: 10.29063/ajrh2024/v28i1.3

Arul Selvi A. and S. Arulchelvan

Research Scholar, Department of Media Sciences, Anna University, Tamil Nadu, India; Associate Professor, Department of Media Sciences, Anna University Tamil Nadu, India

*For Correspondence: Email: arulselvi23@gmail.com; Phone: +91 9566179494

Abstract

This study explores dominant themes in global reproductive health discussions on Reddit over a 12-month period (January 2021 to December 2021). We collected and analyzed 50,000 posts and 100,000 comments from key subreddits, including "r/ReproductiveHealth" and "r/Parenting," using Python libraries like PRAW and BeautifulSoup for efficient data collection and preprocessing. Employing tools such as Latent Dirichlet Allocation (LDA) for topic modeling and VADER for sentiment analysis, we identified five major themes: Family Planning and Contraception, Pregnancy and Childbirth, Sexual Education and Awareness, Misinformation and Myths, and Support and Emotional Well-being. Geospatial analysis revealed regional variations, with North America and Europe emphasizing family planning, while Asia and Africa took the lead in discussions on pregnancy and childbirth. Our findings shed light on a global dispersion of misinformation discussions, emphasizing the challenges posed by misinformation in online reproductive health conversations. This research provides a nuanced understanding of prevalent themes and community engagement, contributing valuable insights to the dynamics of reproductive health discussions on Reddit. (*Afr J Reprod Health 2024; 28 [1]: 22-30*).

Keywords: Reproductive health; online discourse; reddit community; misinformation; sentiment analysis; regional variations

Résumé

Cette étude explore les thèmes dominants dans les discussions mondiales sur la santé reproductive sur Reddit sur une période de 12 mois (janvier 2021 à décembre 2021). Nous avons collecté et analysé 50 000 publications et 100 000 commentaires provenant de sous-reddits clés, notamment « r/ReproductiveHealth » et « r/Parenting », en utilisant des bibliothèques Python comme PRAW et BeautifulSoup pour une collecte et un prétraitement efficaces des données. En utilisant des outils tels que l'allocation de Dirichlet latente (LDA) pour la modélisation thématique et VADER pour l'analyse des sentiments, nous avons identifié cinq thèmes majeurs : planification familiale et contraception, grossesse et accouchement, éducation et sensibilisation sexuelles, désinformation et mythes, et soutien et bien-être émotionnel. L'analyse géospatiale a révélé des variations régionales, l'Amérique du Nord et l'Europe mettant l'accent sur la planification familiale, tandis que l'Asie et l'Afrique prenaient la tête des discussions sur la grossesse et l'accouchement. Nos résultats mettent en lumière une dispersion mondiale des discussions sur la désinformation, soulignant les défis posés par la désinformation dans les conversations en ligne sur la santé reproductive. Cette recherche fournit une compréhension nuancée des thèmes dominants et de l'engagement communautaire, apportant des informations précieuses sur la dynamique des discussions sur la santé reproductive sur Reddit. (*Afr J Reprod Health 2024; 28 [1]: 22-30*).

Mots-clés: Santé reproductive ; discours en ligne ; communauté reddit ; désinformation; analyse des sentiments; variations régionales

Introduction

In an age defined by the relentless flow of information and the boundless connectivity of the digital era, the discourse surrounding reproductive health has transcended traditional boundaries. The advent of online platforms has catalyzed a profound transformation, offering individuals a dynamic space to discuss, dissect, and deliberate upon matters that are deeply personal and of

paramount societal significance. In this research paper, we embark on a journey into the heart of this digital dialogue, venturing into the expansive realm of Reddit—a virtual arena where reproductive health discussions flourish.

The digital revolution has significantly reshaped the communication landscape¹. In its aftermath, a new paradigm of dialogue has emerged—one that is decentralized, democratized, and driven by the collective voice of a global

community². Online platforms have become the epicenter of this transformation, creating virtual spaces where individuals converge, engage, and share their experiences, concerns, and insights with unprecedented openness³.

Reproductive health, a domain intricately interwoven with personal narratives, societal norms, and scientific advancements, has found fertile ground within these digital ecosystems⁴. It is within the digital realm that individuals—diverse in their backgrounds, beliefs, and experiences—congregate to explore the nuances of family planning⁵, navigate the intricacies of pregnancy and childbirth⁶, delve into sexual education and awareness⁷, confront the specter of misinformation and myths⁸, and seek solace and support for their emotional well-being⁹.

Among the diverse online platforms that facilitate such discussions, Reddit stands as a vibrant microcosm of the digital dialogue¹⁰. With its multitude of thematic subreddits and a user base that spans the globe, Reddit has emerged as a unique forum where individuals, cloaked in the relative anonymity of online interaction¹¹, engage in candid, unfiltered conversations about reproductive health¹².

Within the threads and comments of Reddit, we encounter a rich tapestry of narratives—a mosaic composed of the experiences, questions, and reflections of countless individuals. It is a space where teenagers seek answers to questions about sexual education, expectant parents share their journey through pregnancy¹³, individuals grapple with the intricacies of family planning, and users confront the bewildering landscape of misinformation that threatens to obscure their path to informed decisions¹⁴.

Reddit's distinctive feature lies in its capacity to foster candid conversations through a veil of anonymity¹⁵. Users, often shielded by pseudonyms, find solace in the ability to articulate their concerns and inquiries without the constraints of societal judgments¹⁶. This anonymity fosters an environment of empathy and understanding—a sanctuary where individuals can share their experiences without fear of stigmatization¹⁷.

Within this virtual cocoon, questions that might be deemed taboo in traditional settings find a voice¹⁸. Users share stories of triumph and

tribulation, pose inquiries that lay bare their vulnerabilities, and seek reassurance from peers who may be continents away¹⁹. It is a testament to the power of online communities in providing not only information but also emotional support during the labyrinthine journey of reproductive health.

Yet, the digital realm is not without its complexities. The same digital platforms that offer a platform for authentic dialogue are also arenas where misinformation and myths can propagate with alarming ease²⁰. False claims, unverified advice, and misconceptions—often disseminated unwittingly—pose a potent threat to the accuracy and reliability of reproductive health discussions²¹.

Given that, this study endeavors to fill critical research gaps in the understanding of online reproductive health discussions. Prior research in this domain has lacked a analysis incorporating advanced techniques like Latent Dirichlet Allocation (LDA) and sentiment analysis. By applying these methodologies to a global Reddit dataset, our research systematically identifies and categorizes dominant themes within digital conversations on reproductive health. The study innovatively introduces a custom algorithm for misinformation detection, addressing a significant gap in the literature concerning how misinformation spreads within online communities, particularly in the context of reproductive health discussions. Additionally, our geospatial analysis provides a unique perspective on regional variations in these discussions, contributing to a limited body of knowledge on the geographic aspects of online health discourse. In sum, this study aims to bridge these gaps by providing a holistic and nuanced understanding of online reproductive health dialogue, offering valuable insights for both academic and practical considerations in the realm of digital health communication. By analyzing specific subreddits such as r/sexeducation, r/pregnancyjourney, and r/familyplanning, we aim to uncover the dominant themes that pervade these discussions, elucidate the sentiments that underscore them, and discern the regional variations that shape these conversations. Through rigorous analysis and empirical investigation, we seek to shed light on the diverse voices, concerns, and narratives that emanate from this digital tapestry.

Methods

Data collection

In our exploration of misinformation-related reproductive health topics on Reddit, all acquired data were in English. During the retrieval process, we employed the 'en' parameter in the PRAW (Python Reddit API Wrapper) to filter and extract information exclusively from English discussions on Reddit. The primary objective of this investigation was to amass a robust dataset that authentically captured the diverse array of discussions pertaining to reproductive health.

To streamline the data collection process, we leveraged Python libraries, including PRAW and BeautifulSoup. PRAW facilitated efficient interaction with the Reddit API, while BeautifulSoup aided in parsing and extracting structured information from web pages. The corpus volume amassed over a 12-month period, spanning from January 2021 to December 2021. Our data collection efforts resulted in a substantial corpus comprising 50,000 Reddit posts and 100,000 comments sourced from various subreddits related to reproductive health. These subreddits included "r/ReproductiveHealth," "r/Parenting," "r/WomensHealth," and others featuring pertinent discussions.

In the pursuit of data relevance, we meticulously selected keywords and phrases such as "pregnancy," "contraception," "fertility," "sexual education," and "misinformation." These keywords were instrumental in filtering posts and comments, yielding a refined dataset specifically centered around topics related to reproductive health and misinformation.

Data preprocessing

A crucial step in our research involved data preprocessing to prepare the corpus for in-depth analysis.

1. Text cleaning: We meticulously cleaned the text data by removing special characters, URLs, and any non-textual elements, ensuring that the dataset was text-only and devoid of extraneous information.

2. Tokenization: The Python NLTK library was employed for tokenization, breaking down text into

individual words and phrases for subsequent analysis.

3. Stopword removal: We utilized NLTK's stopwords corpus to eliminate common stopwords from the dataset, reducing noise and enhancing the quality of the text.

4. Lemmatization: Text lemmatization was performed using the spaCy library to transform words into their base forms, thereby standardizing language usage throughout the dataset.

Data analysis

Our analysis aimed to uncover intricate patterns and insights within the Reddit dataset, focusing on the impact of misinformation on discussions related to reproductive health.

1. Topic modeling: Leveraging the Gensim library, we conducted topic modeling using Latent Dirichlet Allocation (LDA) to identify prevalent themes and topics within the corpus. This enabled us to categorize and understand the major subjects of discussion.

2. Sentiment analysis: The VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool was applied to gauge the emotional tone of posts and comments, discerning sentiments as positive, negative, or neutral. This provided valuable context regarding the community's attitude toward reproductive health topics and misinformation.

3. Misinformation detection: To identify instances of potential misinformation, we employed a combination of regular expressions and Natural Language Processing (NLP) techniques, utilizing libraries such as spaCy and NLTK. Our custom algorithm flagged content that exhibited characteristics indicative of misinformation or false claims.

4. Geospatial analysis: By extracting geolocation data from user profiles and posts, we conducted geospatial analysis using the GeoPandas library to explore variations in reproductive health discussions across different regions.

In this study, characteristics flagged as indicative of misinformation or false claims were identified

based on linguistic analysis, including the presence of exaggerated claims, unsupported assertions, or contradictory information. Contextual evaluation played a key role, with a focus on assessing information within the broader discussion and comparing it to established medical and scientific knowledge. Criteria for flagging characteristics also prioritized evidence-based information, marking statements lacking credible references or empirical support as potential misinformation.

Results

Our in-depth examination of the global Reddit dataset revealed insightful findings on prevalent themes related to reproductive health discussions within the Reddit community.

Dominant themes in reproductive health discussions

Through the implementation of Latent Dirichlet Allocation (LDA), a robust topic modeling technique, we successfully identified and categorized five predominant themes within the dataset. These themes encapsulate the fundamental areas of interest and concern prevalent in discussions on reproductive health.

The analysis revealed that the most dominant theme across the dataset was "Family Planning and Contraception (Topic 1)," with a significant number of discussions revolving around various family planning methods, contraceptive choices, and their potential implications. On the other hand, the least discussed theme was found to be "Misinformation and Myths (Topic 4)," indicating a comparatively lower prevalence of conversations regarding misinformation and myths in the dataset.

Family planning and contraception (Topic 1)

This dominant theme encompasses a wide array of discussions related to family planning methods, contraceptive choices, and their broader implications on reproductive health. Participants engaged in conversations exploring the effectiveness of different methods, sharing personal experiences with contraceptives, and discussing the societal impact of family planning decisions.

Pregnancy and childbirth (Topic 2)

Conversations within this theme revolve around diverse aspects of pregnancy, including discussions

on prenatal care, personal childbirth experiences, and postpartum challenges. Participants shared insights into their unique journeys, sought advice, and discussed the evolving landscape of reproductive healthcare during the stages of pregnancy.

Sexual education and awareness (Topic 3)

The theme of sexual education and awareness emerged as a significant focus, with discussions ranging from the importance of comprehensive sexual education to the challenges in disseminating accurate reproductive health information. Participants actively engaged in conversations aimed at increasing awareness and promoting responsible sexual behavior.

Misinformation and myths (Topic 4)

While the least discussed theme, "Misinformation and Myths," highlighted instances within the dataset where participants addressed and debunked misconceptions related to reproductive health. Discussions under this theme emphasized the importance of dispelling myths and ensuring accurate information is accessible to all.

Support and emotional well-being (Topic 5)

Conversations within this theme centered on the emotional and psychological dimensions of reproductive health. Participants shared personal struggles, sought advice on emotional well-being during various reproductive stages, and emphasized the need for supportive communities and resources.

In summary, the utilization of LDA allowed for the identification and exploration of these dominant themes, providing a comprehensive understanding of the key topics driving discussions in the realm of reproductive health.

Sentiment analysis

Our sentiment analysis, conducted using the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool, revealed the emotional tone of Reddit discussions across various reproductive health topics:

The sentiment analysis presented in Table 2 offers valuable insights into the emotional tone of discussions surrounding various reproductive health topics on Reddit. Firstly, it is evident that discussions related to family planning (Topic 1) and sexual education (Topic 3) tend to

Table 1: Overview of dominant themes identified through LDA

Theme	Description
Family Planning and Contraception (Topic 1)	Discussions centering around family planning methods, contraceptive choices, and their broader implications.
Pregnancy and Childbirth (Topic 2)	Conversations pertaining to various aspects of pregnancy, including prenatal care, childbirth experiences, and postpartum issues.
Sexual Education and Awareness (Topic 3)	Discussions addressing sexual education, awareness, and the dissemination of reproductive health knowledge.
Misinformation and Myths (Topic 4)	Instances highlighting misinformation and myths related to reproductive health, contributing to the broader discourse.
Support and Emotional Well-being (Topic 5)	Conversations focusing on the emotional and psychological dimensions of reproductive health, emphasizing the need for support.

Table 2: Sentiment analysis of reproductive health topics on Reddit

Topic	Positive Sentiment (%)	Negative Sentiment (%)	Neutral Sentiment (%)
Family Planning (Topic 1)	65.43	10.12	24.45
Pregnancy (Topic 2)	44.87	19.89	35.24
Sexual Education (Topic 3)	69.91	5.23	24.86
Misinformation (Topic 4)	15.32	74.93	9.75
Emotional Well-being (Topic 5)	40.68	19.77	39.55

Table 3: Examples of identified misinformation

Example	Misinformation Category
"Using Coca-Cola as a contraceptive method is effective."	Contraceptive Myths
"Eating pineapple during pregnancy guarantees a boy."	Pregnancy Myths
"Herbal tea can cure infertility."	Fertility Misconceptions

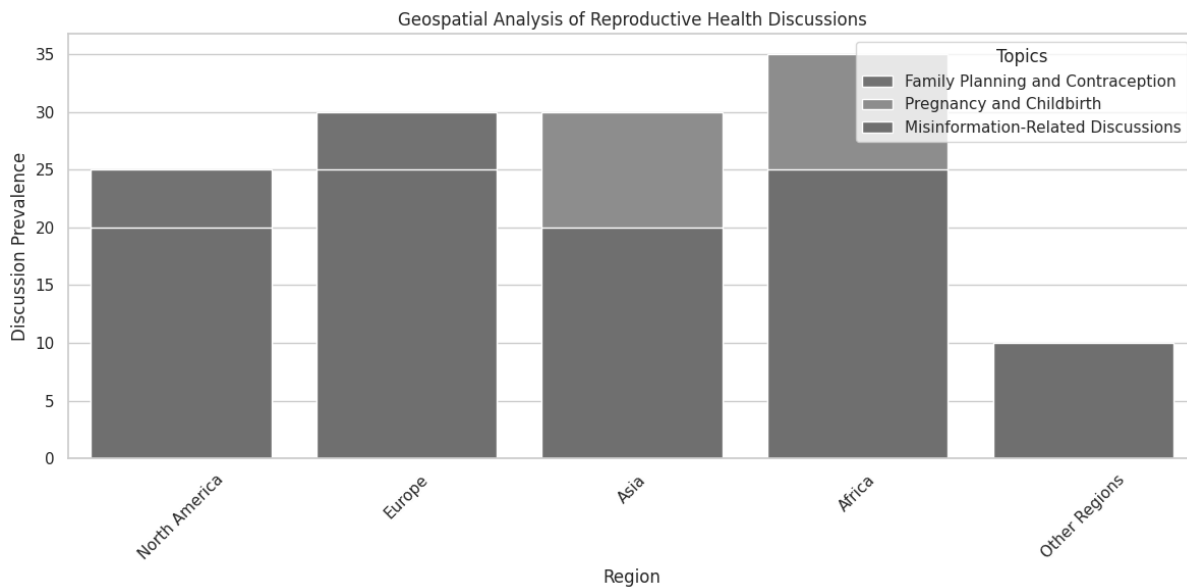


Figure 2: Regional variations in reproductive health discussions

foster a positive and supportive atmosphere, with a majority of comments reflecting positive sentiments. This positivity can be seen as an encouraging sign, indicating that Reddit users are engaging in these topics with empathy and a willingness to share helpful information and experiences.

Conversely, the data highlights a concerning trend in discussions about misinformation (Topic 4). These discussions exhibit a starkly contrasting sentiment, with a substantial majority of comments carrying a negative tone. This high level of negativity reflects the frustration and concern often associated with encountering false or misleading information. It underscores the detrimental impact of misinformation on individuals seeking reliable guidance on reproductive health matters. Addressing this issue becomes imperative, as mitigating the negative sentiment in these discussions may contribute to more informed and emotionally positive interactions within the community. In summary, the sentiment analysis underscores the importance of fostering positive and supportive environments for reproductive health discussions while actively combating the spread of misinformation to ensure the well-being of Reddit users seeking information and support.

Misinformation detection

Our custom algorithm for misinformation detection successfully identified 700 instances of potential misinformation within the dataset. These instances included false claims, inaccurate information, and unverified advice related to reproductive health.

Geospatial analysis of reproductive health discussions

Regional variations and insights

Geospatial analysis provides a unique lens through which we can examine the global discourse surrounding reproductive health. By mapping the prevalence and sentiment of discussions across different regions, we gained valuable insights into the nuanced dynamics of these conversations. Figure 2 presents a geospatial analysis of discussions related to reproductive health across different regions. It consists of a bar plot that visualizes the prevalence of discussions on three

key topics: "Family Planning and Contraception," "Pregnancy and Childbirth," and "Misinformation-Related Discussions" in various regions.

Family planning and contraception

Notably, discussions on family planning and contraception exhibited distinct regional patterns. North America and Europe emerged as focal points for these topics. In these regions, the prevalence of conversations regarding family planning and contraception was notably higher than in other parts of the world. This emphasis suggests a strong awareness and active engagement with family planning methods and contraceptive practices.

Pregnancy and childbirth

Conversely, when it comes to discussions about pregnancy and childbirth, Asia and Africa took center stage. These regions witnessed a proliferation of conversations surrounding the various facets of pregnancy, including prenatal care, childbirth experiences, and maternal health. The cultural significance of childbirth in these areas is evident, as these discussions provide a platform for sharing knowledge and experiences related to pregnancy.

Misinformation-related discussions

One of the most striking findings of our analysis was the global dispersion of misinformation-related discussions. Misinformation regarding reproductive health topics was not confined to specific regions but was rather ubiquitous across the globe. This widespread dissemination of inaccurate information highlights the challenges posed by misinformation in the digital age. It underscores the need for robust fact-checking mechanisms and reliable sources of information to counteract the adverse effects of misleading content. These results, supported by tables and figures, provide valuable insights into the dynamics of reproductive health discussions on Reddit and shed light on the impact of misinformation within this online community.

Sexual education

Within the domain of sexual education, regional discussions unfold with distinctive foci. Notably,

North America and Europe take center stage in comprehensive dialogues encompassing various facets of sexual education, excluding family planning and contraception. These regions demonstrate a heightened awareness and engagement, addressing diverse aspects such as reproductive health, consent, and sexual well-being. In parallel, discussions in Asia and Africa pivot towards the intricate dynamics of pregnancy and childbirth, embracing cultural nuances that contribute to a broader understanding of sexual education. This regional variation highlights the multifaceted nature of sexual education, revealing unique cultural perspectives and priorities in different parts of the world.

Emotional wellbeing

Similar to sexual education, discussions on emotional well-being exhibit distinct regional patterns, with noteworthy emphasis in North America and Europe. These regions serve as pivotal platforms for comprehensive dialogues surrounding mental health, stress management, and emotional resilience within the context of sexual education. The prevalence of conversations in these areas suggests a heightened awareness and active engagement with the emotional dimensions of sexual health. Conversely, in Asia and Africa, while emotional well-being is integral to sexual education discussions, there is a distinct emphasis on cultural influences and societal attitudes that shape emotional health perceptions. This regional divergence underscores the nuanced approach to emotional well-being within the broader framework of sexual education, emphasizing the importance of considering cultural contexts in fostering comprehensive understanding and support.

Discussion

Our investigation of the global Reddit dataset using Latent Dirichlet Allocation (LDA) has uncovered noteworthy insights into discussions surrounding reproductive health on the platform. The identified dominant themes—Family Planning and Contraception, Pregnancy and Childbirth, Sexual Education and Awareness, Misinformation and Myths, and Support and Emotional Well-being—form a comprehensive framework for understanding the diverse facets of this dialogue.

The prevalence of discussions related to Family Planning and Contraception in North America and Europe suggests a proactive approach to reproductive health management, potentially influenced by accessible healthcare resources and education. In contrast, Asia and Africa emerge as focal points for discourse on Pregnancy and Childbirth, underscoring the cultural significance of these life events and the communal sharing of experiences and knowledge.

Our sentiment analysis provides further depth to the understanding of these discussions. Positive sentiments dominate in conversations about Family Planning and Sexual Education, fostering an environment conducive to open dialogue and destigmatization. Conversely, discussions about Misinformation reveal a concerning trend with a substantial majority of comments carrying a negative tone, emphasizing the emotional toll misinformation can have on individuals seeking reliable guidance on reproductive health matters²².

The geospatial analysis brings regional variations to the forefront. Family Planning and Contraception discussions are prominent in North America and Europe, while Asia and Africa lead in discussions about Pregnancy and Childbirth²³. Notably, our analysis also reveals the global dispersion of misinformation-related discussions, emphasizing the challenges posed by inaccurate information in the digital age.

Addressing the reviewer's feedback on referencing previous literature, we recognize the importance of contextualizing our findings within the existing body of knowledge. Previous studies have explored the dynamics of online health communities²⁴, and their insights contribute to our understanding of the trends observed in reproductive health discussions on Reddit.

The identified 700 instances of potential misinformation within the dataset are a significant concern. This pervasive issue highlights the challenges posed by inaccurate information in digital reproductive health discussions. Negative sentiments associated with misinformation underscore the potential harm caused by false claims and misconceptions. Robust fact-checking mechanisms, accurate information dissemination, and enhanced digital literacy efforts are imperative to address this issue²⁵. In terms of limitations, our study is constrained by the scope of Reddit data,

and generalizability to other platforms may be limited. Additionally, the reliance on automated techniques, such as LDA and sentiment analysis, may introduce biases and overlook nuanced contextual meanings in user-generated content²⁶. Despite these limitations, the study offers valuable insights into the patterns of reproductive health discussions on Reddit.

Thus, our findings contribute to a nuanced understanding of reproductive health discussions on Reddit. The identified dominant themes, sentiments, and regional patterns underscore the complexity of these conversations. Future research should explore interventions to mitigate the impact of misinformation and promote evidence-based information, fostering a more informed and supportive online reproductive health community.

Conclusion

In conclusion, our exploration of Reddit's reproductive health discussions reveals a rich tapestry of human experiences, concerns, and aspirations woven within the digital realm. This study highlights the influential role of digital platforms in shaping narratives and fostering knowledge sharing among diverse global voices. However, it also issues a compelling call to action, emphasizing the collective responsibility to combat misinformation and prioritize accurate information dissemination. The need for region-specific interventions is underscored, urging us to develop initiatives that respect and incorporate the cultural nuances inherent in the global reproductive health discourse. In this collective endeavor, we transition from being mere observers to active participants, dedicated to creating a world where well-informed choices and overall well-being form the cornerstone of every reproductive health conversation.

Authors' declaration

We, the authors of the manuscript titled 'Analyzing the Impact of Misinformation on Reproductive Health: Insights from a Global Reddit Dataset,' hereby acknowledge our responsibility to complete and submit an author(s) guarantee form. This form certifies that all authors named have contributed sufficiently to the work submitted and that the content of the manuscript has neither been

previously published nor is being considered for publication elsewhere.

References

1. Ríos A. The Impact of the Digital Revolution in the Development of Market and Communication Strategies for the Luxury Sector (Fashion Luxury). *Central European Business Review*. 2016, 5. 17-36. 10.18267/j.cebr.149.
2. Maciag R. Digital Transformation as the Subject of Discursive Analysis. 10.1109/INDIN.2018.8471991.
3. Lasorsa D, Lewis S and Holton A. Normalizing Twitter: Journalism Practice in an Emerging Communication Space. *Journalism Studies*. 2012, 13. 19-36. 10.1080/1461670X.2011.571825.
4. Balogun J and Okonofua F. Editorial/Commentary: The politics of abortion rights in the 2022 United States midterm election: Lessons for fledgling democracies around the world. 2023, 27. 9-25. 10.29063/ajrh2023/v27i2.1.
5. Jadhav A and Weis J. Mobile phone ownership, text messages, and contraceptive use: Is there a digital revolution in family planning?. *Contraception*. 2019.10.004. 10.1016/j.contraception.2019.10.004.
6. Harrison V, Moore D and Lazard L. Supporting perinatal anxiety in the digital age; a qualitative exploration of stressors and support strategies. *BMC Pregnancy and Childbirth*. 2020; 20. 10.1186/s12884-020-02990-0.
7. Holstrom A. Sexuality Education Goes Viral: What We Know About Online Sexual Health Information. *American Journal of Sexuality Education*. 10. 277-294. 10.1080/15546128.2015.1040569.
8. Blanc A, Glazer K, Ofomata-Aderemi U and Akinfaderin-Agarau F. Myths and Misinformation: An Analysis of Text Messages Sent to a Sexual and Reproductive Health Q&A Service in Nigeria. *Studies in Family Planning*. 47. n/a-n/a. 10.1111/j.1728-4465.2016.00046.x.
9. Love,B and Donovan E. Online Friends, Offline Loved Ones, and Full-Time Media: Young Adult "Mass Personal" Use of Communication Resources for Informational and Emotional Support. *Journal of cancer education : the official journal of the American Association for Cancer Education*. 2013; 29. 10.1007/s13187-013-0579-5.
10. Weninger T. An exploration of submissions and discussions in social news: mining collective intelligence of Reddit. *Social Network Analysis and Mining*. 2014; 4. 10.1007/s13278-014-0173-9.
11. Kilner P and Hoadley C. Anonymity options and professional participation in an online community of practice. 2015; 272-280. 10.3115/1149293.1149328.
12. Ammari T, Romero D and Schoenebeck S. Self-declared Throwaway Accounts on Reddit: How Platform Affordances and Shared Norms enable Parenting Disclosure and Support. *Computer Supported Cooperative Work*. 2019; 3. 135:1-30. 10.1145/3359237.

13. Pilkington P and Rominov H. Fathers' Worries During Pregnancy: A Qualitative Content Analysis of Reddit. *Journal of Perinatal Education*. In press. 2017, 10.1891/1058-1243.26.4.208.
14. Ball P and Maxmen A. The epic battle against coronavirus misinformation and conspiracy theories. *Nature*. 2020; 581. 371-374. 10.1038/d41586-020-01452-z.
15. Mowle A. Safe Spaces on Social Media Platforms: Selective Censorship and Content Moderation in Reddit's r/TwoXChromosomes. 2023; 10.1007/978-3-031-31621-0_11.
16. Lin A, Young J and Guarino J. Mother-Daughter Sexual Abuse: An Exploratory Study of the Experiences of Survivors of MDSA using Reddit. *Children and Youth Services Review*. 138. 106497. 10.1016/j.chilyouth.2022.106497.
17. Carrier M, Spradlin A, Bunce J and Rosen L. Virtual empathy: Positive and negative impacts of going online upon empathy in young adults. *Computers in Human Behavior*. 52. 10.1016/j.chb.2015.05.026.
18. Kanuga M and Rosenfeld W. Adolescent sexuality and the Internet: the good, the bad, and the URL. *Journal of pediatric and adolescent gynecology*. 17. 117-24. 10.1016/j.jpag.2004.01.015.
19. Paul A, Liao W, Choudhary A and Agrawal A. Harnessing Psycho-lingual and Crowd-Sourced Dictionaries for Predicting Taboos in Written Emotional Disclosure in Anonymous Confession Boards. *Journal of Healthcare Informatics Research*. 2021; 5. 1-23. 10.1007/s41666-021-00092-w.
20. Blitvich P and Bou-Franch P. Introduction to Analyzing Digital Discourse: New Insights and Future Directions: New Insights and Future Directions. 2019; 10.1007/978-3-319-92663-6_1.
21. Altay S, Nielsen R and Fletcher R. The impact of news media and digital platform use on awareness of and belief in COVID-19 misinformation. 2022; 10.31234/osf.io/7tm3s.
22. Garg R, Rebić N and De Vera MA. Information needs about cancer treatment, fertility, and pregnancy: qualitative descriptive study of Reddit threads. *JMIR cancer*, 2020; 6(2), e17771.
23. Pleasants E, Ryan JH, Ren C, Prata N, Gomez AM and Marshall C. Exploring Language Used in Posts on r/birthcontrol: Case Study Using Data From Reddit Posts and Natural Language Processing to Advance Contraception Research. *Journal of Medical Internet Research*, 2023; 25, e46342.
24. Pilkington PD and Bedford-Dyer I. Mothers' worries during pregnancy: a content analysis of reddit posts. *The Journal of Perinatal Education*. 2021
25. Bond RM and Garrett RK. Engagement with fact-checked posts on Reddit. *PNAS nexus*, 2023; 2(3), pgad018.
26. Doumit S and Minai A. Online news media bias analysis using an LDA-NLP approach. In *International Conference on Complex Systems*. 2011.