

A Decade of Social Issues Research in India: Thematic Analysis of Doctoral Studies Through Topic Modeling (2011-2020)

Manorama Tripathi[#], Ankit Yadav^{\$}, Narasaiah Inugurthi[#] and Vinit Kumar^{~*}

[#]*Dr BR Ambedkar Central Library, Jawaharlal Nehru University, New Delhi - 110 067, India*

^{\$}*Central Library, National Law University, New Delhi - 110 078, India*

[~]*Department of Library and Information Science, Babasaheb Bhimrao Ambedkar University, Lucknow - 226 025, India*

^{*}*Email : mailvinitkumar@gmail.com*

ABSTRACT

Universities and research centres serve as epicentres of knowledge in any society, aiding policymakers in the formulation and implementation of policies based on research findings. This study examines doctoral theses on social issues awarded at Indian higher education institutions between 2011 and 2020, focusing on dominant themes and their temporal evolution. Topic modeling, specifically Latent Dirichlet Allocation (LDA) implemented in Python, was applied to analyse 448 theses downloaded from Shodhganga, a national repository. The analysis identified nine dominant themes, with the largest proportion (15.40 %) addressing child and gender issues, followed by 13.84 % on corporate, political, and legal challenges. Theses on social development and education accounted for 12.72 %, while 10.94 % focused on political issues in social development. Other topics included social identities, government policies, rural development, corporate social responsibility, and adolescent stress. The findings highlight the central role of child and gender-related issues in doctoral research, with shifting focuses over time in response to socio-political changes. Themes such as rural development and women's economic roles fluctuated, while social development and gender dynamics showed consistent growth. This study underscores the evolving academic priorities and potential gaps between academic research and societal needs, offering insights for future research directions in social issues.

Keywords: Computational text analysis; Social science research; Doctoral studies; Indian higher education; Content analysis; Topic modeling; Research trends

1. INTRODUCTION

Despite its vital role in societal development and policymaking, research on social issues has historically received less attention and investment compared to other scientific disciplines. This disparity persists even as Social Science Research (SSR) continues to provide crucial insights into the complex human behavior and social phenomena that shape our world. Gordon Allport, in his 1953 work "The Nature of Prejudice," highlighted this paradox by examining the fundamental differences between natural and social sciences¹. In most cases, research in exact sciences results in providing solutions to immediate challenges through technological innovations, whereas research in social sciences helps us to understand the complexities of the society and gives us valuable solutions for the social, economic and political challenges that were faced by the society by providing necessary inputs to transform the society through policy reforms and social interventions. As per a search conducted on 18 April 2025, there are 1623,230 journal articles published

during 2020-2024 and indexed in the Social Science Citation Index, Web of Science, a product of Clarivate Analytics. The USA has 505,515 and India has 30,149 articles². A study by Clarivate Analytics and the Chinese Academy of Sciences has evaluated and ranked countries in eleven broad research areas across all streams. The USA leads with a score of 193.69 while India is at the 15th rank with a score of 19.26. The broad 11 research areas also include economics, psychology and other social science areas³. The rapid strides in education, policy research, and the ethical and responsible use of AI propel the global sciences and humanities research output⁴. To some extent, the reasons for undervaluation of social science research can be attributed to the inherent complexities, including the intricacies of human behavior, the extended timeframes required for comprehensive understanding of social processes, and the unpredictability of outcomes in social science research. However, these challenges also highlighted the significance of social sciences for the human progress. The social sciences explore diverse aspects of human society, ranging from economic and political systems to cultural dynamics and psychological behaviors, offering crucial insights that help

policy decisions and societal progress. The relationship between social science research and policymaking is vital in contemporary India, where the prevalence of social issues like inequality, discrimination, and restricted access to resources continue to affect millions of people. A thorough academic investigation is necessary to develop effective solutions.

The role of higher education institutions in addressing these emerging challenges has become increasingly important. Universities and research centers serve as epicenters of knowledge in any society. They help the policy makers in formulation of policies and their implementation in light of the findings and recommendations of the research studies carried out in the higher education institutions. The findings of doctoral research in higher educational institutions are considered as the most substantial and original contributions to knowledge generation. Doctoral research serves as an indicator of broader research trends and emerging scholarly interest. A systematic analysis of the research conducted in these academic institutions can reveal both the evolving nature of social issues in the higher education institutions and the efficacy of academic responses to the challenges faced by society.

The present study examines the doctoral studies on social issues conducted at Indian higher education institutions between 2010 and 2020 and evaluates how academic research interacts with and shapes policy decisions aimed at tackling the complex societal problems and helps for the equitable growth. By analysing the doctoral dissertations, the study aims to identify dominant themes, trace evolving patterns, and explain how academic research in the country at the highest level has responded to emerging social issues in the country. The importance of this study is also lies in its application of topic modeling method. As a computational method, topic modeling method allows us to process and analyse large volumes of textual data from dissertations in short span of time, revealing the underlying patterns and dominant themes that might not be immediately apparent through traditional review methods. By examining a decade of such work, the study attempts to trace how research priorities have shifted and evolved over a period of ten years in response to changing social, economic and political situations, and attempts to identify any gaps between academic focus and pressing societal needs.

2. LITERATURE REVIEW

Several attempts have been made to apply topic modeling to social science data and literature. These attempts have uncovered latent themes and patterns within large volumes of textual data. Researchers have utilised various topic modeling techniques, such as Latent Dirichlet Allocation (LDA) and Structural Topic Models (STM), to analyse the content of social media, political speeches, and scholarly literature. By employing these methods, social scientists gain valuable insights into public opinion, policy discourse and emerging trends in the various domains of social sciences.

Hagen⁵ presented a framework for training and validating LDA models using e-petition data, highlighting the advantages of LDA for efficiently extracting multiple themes and reducing human bias in content analysis. Baranowski⁶ applied LDA to large corpora in the social sciences and discussed its advantages and disadvantages in automated text analysis. In the context of social welfare and other social science fields, topic modeling has been applied to review large volumes of research papers by identifying key topics and trends⁷. Egger & Yu⁸ compared topic modeling algorithms in analysing Twitter data for social science research.

Lee⁹, *et al.* recommended the use of computational topic modeling for analysing a large corpus but they noted that it generally produces moderate-quality results compared to the more detailed results obtained through manual content analysis. Manual content analysis provides a detailed and in-depth understanding that may not be detected by computational methods but is time intensive, costly, and requires significant intellectual labor to identify hidden topics, particularly when dealing with large datasets¹⁰. Cui¹¹, *et al.* conducted a study on social media data to understand human activities in Urban Green Spaces (UGSs). The complexity and diversity of UGS use are not captured by traditional methods, such as word frequency analysis, which necessitates the use of advanced techniques such as the Structural Topic Model (STM), as it facilitates the analysis of temporal and spatial patterns. Aminpour and Saidi¹² implemented a topic modeling algorithm called Latent Dirichlet Allocation (LDA) to determine trip purposes and identify travel behavioral changes caused by the COVID-19 disruption by analysing smart card transit data indicating reduced need for expensive and time-consuming manual surveys. Mennig¹³ analysed German parliamentary debates from 2017 to 2022, addressing agricultural issues using Structural Topic Modeling (STM). The study discovered 24 distinct topics, including agricultural-environment interactions, rural development, animal welfare, and farming economics. Pang¹⁴, *et al.* analysed 186,180 news reports on Chinese migrant workers between 2001 and 2021 by applying topic modeling and Critical Discourse Analysis (CDA). It categorised 17 topics into five main groups: work and labor, support and care, rights and interests, homecoming, and social emergencies. Hu¹⁵, *et al.* has applied Latent Dirichlet Allocation (LDA) technique for topic modeling to uncover the key topics that highlight how social media shapes and disseminates the Malaysian Chinese folk beliefs. 4012 text posts on Facebook were analysed, and four main themes were identified: practitioner worship, temple activities, deity legends, and merchandise about deity statues. Özmantar¹⁶, *et al.* identified key research themes in the field of teacher educators to reveal thematic trends overtime. They analysed 754 publications from the SCOPUS database and identified 10 key themes using the Latent Dirichlet Allocation technique. The study discussed how thematic trends are influenced by education journals, policy reforms, and COVID-19 disruption. It

also highlights the implications of this study for teacher education research and policymaking. Similarly, Kumar & Thakur¹⁷ applied LDA to study doctoral-level library and information science research trends in India and identified emerging research themes and the temporal evolution of research themes over the last ten years. Ismail¹⁸, *et al.* have analysed 18399 publications from sociology, anthropology and digital cultural studies and identified social identity, cultural identity and globalisation as key themes in identity studies, while identity in the COVID-19 crisis and racial disparities have been shown as emerging themes. Hamdan¹⁹, *et al.* analysed 65,742 records to highlight trends, emerging themes, patterns of productivity, and visibility within Social Science research on COVID-19. They have shown that the topics of telemedicine and well-being are the leading topics, while the emerging topics are related to interventions and mental well-being.

Despite the existence of numerous studies implementing topic modeling in social sciences, no research has been identified that attempted to analyse doctoral research output in the field of social issues. Building on the gaps identified in the literature, this study establishes its objectives.

3. OBJECTIVES

The primary aim of this study is to analyse the thematic landscape and temporal evolution of doctoral research on social issues in India from 2011 to 2020. Specifically, this study seeks to:

- Identify and analyze the dominant research themes in doctoral dissertations addressing social issues across Indian higher education institutions through topic modeling analysis.
- Track and examine the temporal evolution of research focuses over the decade (2011-2020), identifying significant shifts in research priorities and emerging areas of scholarly interest.
- Investigate emerging patterns within identified themes to understand:
- The distribution of research attention across various social issues.
- The emergence of new research areas within established themes.
- The relationship between contemporary social challenges and academic focus.

4. METHODOLOGY

The present study aims to provide an overview of doctoral-level research conducted in the domain of social issues. The culmination of doctoral-level research output is a thesis adjudicated before the candidate is awarded a doctorate degree. In India, the University Grants Commission (UGC), New Delhi, the national regulatory body for universities, mandates that universities upload an electronic version of their awarded thesis to the Shodhganga repository. For the present study,

electronic records of theses related to social issues were retrieved from the repository in October 2023 by using the keyword “social issues.” This term was selected as it broadly encompasses key themes within the domain, including societal change, social problems, and social challenges. Moreover, as the Shodhganga repository’s search functionality indexes both metadata and the full text of theses-unlike many bibliographic databases which typically index only the title and abstract-we believe this approach enabled a comprehensive retrieval of relevant theses. Given the centrality of the term “social issues” to the subject matter, it is unlikely that a thesis addressing such themes would omit this terminology altogether in the full text.

Initially, 451 theses were retrieved from the Shodhganga repository. After applying the inclusion and exclusion criteria, 448 theses were selected for this study. The inclusion and exclusion criteria were as follows:

- Timeframe: Theses submitted between 2011 and 2020
- Language: Only English-language theses were considered
- Keyword: Social Issues
- Subject area: Theses from social sciences and humanities were prioritised
- Readability: Theses in image form or those that were unreadable were excluded

In the second phase, comprehensive data collection was conducted from the 448 selected theses. Essential information, including title, author, supervisor, university name, objectives, methodology, findings, results, scope for future research, recommendations, and URLs were systematically extracted and recorded in a word file. This compiled information formed the corpus for the subsequent topic modeling analysis.

The text preprocessing phase involved multiple steps to prepare the corpus for topic modeling. The text was cleaned by removing special characters, numbers, and punctuation marks, followed by conversion to lowercase text. Commonly occurring stop words were eliminated from the corpus to reduce the noise in the analysis. A snowball stemmer was applied for word stemming to standardize the word forms. To capture meaningful phrases and maintain contextual relationships in the text, bigrams and trigrams were created using Python sklearn library.

The primary objective of this study was to identify the dominant themes explored by doctoral researchers at Indian universities in the context of social issues. To achieve this, we adopted a computational text analysis approach, employing quantitative methods to examine thematic patterns within a corpus of awarded doctoral theses. Given the scale of the corpus, manually reviewing the content of each thesis would have been time-consuming and impractical; hence, an automated approach, in this case topic modeling was deemed appropriate. Topic modeling was applied to this preprocessed corpus in order to uncover abstract themes within the text of theses. As a statistical technique widely used in natural language processing and machine learning, topic modeling analyses patterns of word co-occurrence across documents to

identify latent semantic structures that characterise the content¹⁷. Using algorithms such as Latent Dirichlet Allocation (LDA), topic modeling can automatically extract meaningful topics from large volumes of unstructured textual data^{20,21}. These topics are represented as clusters of frequently co-occurring words, enabling researchers to gain insight into the primary themes and content of document collections without manually examining each text. LDA has been implemented to identify relevant studies for systematic reviews in social sciences²². Hagen⁵ applied LDA models on e-petition data to efficiently extract multiple themes aligning them with social events. LDA topic models provide reliable coverage and depth of themes in a textual collection comparable to human coding but require manual interpretation of topics²³.

The model parameters, particularly the number of topics, were determined through coherence score optimisation and visualisation using the PyLDAvis library. The optimal number of topics yielding distinct thematic clusters was selected through iterative refinement of the number of topics and PyLDAvis visualisation. The validation process also included a manual review of the top words associated with each topic as well as verification of the top five topic assignments within each identified topic. One of the authors, with expertise in social sciences, evaluated topic interpretation and labelling to ensure scholarly relevance and accuracy.

5. RESULTS

This section presents the findings derived from the analysis of doctoral theses on social issues awarded by Indian universities between 2011 and 2020. The results are organised into four subsections. First, we examine the overall trends in the awarding of theses during the study period. This is followed by a discussion of the dominant themes identified through topic modelling. The subsequent subsection presents the distribution of dominant topics across the corpus, while the final part explores the temporal evolution of individual topics, shedding light on shifts in research focus over time.

Figure 1 shows the number of theses awarded annually from 2011 to 2020, with a total of 448 theses

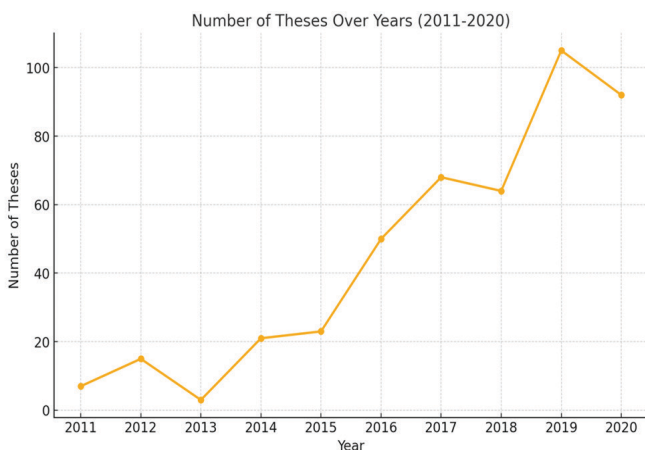


Figure 1. Number of theses on social issues awarded over the years (2011-2020).

during this period. The data reflect a general upward trend in the number of theses on social issues, with some fluctuations. The number of theses was relatively low from 2011 to 2013, peaking at just 15 in 2012 and dropping to three in 2013. From 2014 onward, there was a steady and substantial increase, reaching a maximum of 105 in 2019. However, by 2020, the number of theses had declined slightly to 92. The data indicate strong growth in academic output, particularly from 2016 onward, suggesting either an increase in student enrollment, greater emphasis on research, or improved academic infrastructure and support over time. The slight drop in 2020 might be attributed to external factors such as the COVID-19 pandemic, which may have impacted academic schedules.

5.1 Dominant Themes

Topic modeling on the corpus of 438 theses yielded nine dominant topics. A more detailed breakdown of the topics, along with their theme labels, is provided in Table 1. We compared the coherence scores to determine the number of topics and applied the LDA algorithm in Python. The steps involved preprocessing, tokenisation normalisation using the snowball stemmer, and topic modeling.

Table 1 provides an outline of the nine themes related to social, economic, and political issues, each accompanied by a brief description of its scope and a list of the top 20 keywords associated with each theme. These themes cover a range of topics, including rural development and women's roles in Corporate Social Responsibility (CSR), gender dynamics, and educational factors. The keywords provide insight into the focus areas of each theme, highlighting the central issues, such as "woman," "social," "education," "development," and "government." It shows that predominantly topics are related to societal development, gender issues and enabling impact of government policies.

We identified the heterogeneity of the topics using the PyLDAvis package in Python. Fig. 2 presents a multidimensional scaling visualisation of the topic distributions generated using the PyLDAvis library. Each bubble on the map in fig. 2 represents a distinct theme, and its spatial arrangement illustrates the semantic relationships between topics. The distance between bubbles indicates thematic distinctiveness; greater separation suggests more heterogeneous topics, whereas overlap indicates shared semantic content. This visualisation helps identify both distinct research themes and areas in which social issues share common theoretical or methodological approaches. The marginal topic distributions in Fig. 2 are represented by the size of each bubble, which corresponds to the prevalence of the associated topic within the corpus. Larger bubbles indicate themes that appear more frequently in the doctoral theses, while smaller bubbles represent less frequent themes.

The Table1 also shows that the role of women in society and gender dynamics has been covered in multiple

Table 1. Identified themes and their scopes

Topic number	Theme labels	Scope	Top 20 keywords
T-1	Rural development and women's economic roles	focuses on the role of women in rural development, including economic and social aspects.	woman, family, income, social, farmer, agricultural, among, rural, development, education, like, household, higher, issue, economic, employment, need, livelihood, high, government
T-2	Corporate social responsibility and public welfare	Explores the impact of CSR activities on rural and public welfare, particularly in education and health.	medium, social, csr, issue, research, development, year, various, rural, education, public, woman, health, awareness, among, like, need, program, analysis, people
T-3	Gender, caste, and political issues in social development	Examines the intersection of gender, caste, and political issues within social development.	social, woman, political, community, education, support, issue, government, caste, group, gender, economic, like, tribal, role, development, household, research, child, due
T-4	Cultural and social identities in the workforce	Discusses cultural, social, and religious identities within the workforce, particularly concerning women and caste.	social, issue, worker, woman, work, cultural, like, economic, community, india, often, religious, caste, within, family, security, among, role, identity, societal
T-5	Government policies and rural development programs	Analyzes government initiatives and policies aimed at rural development, with a focus on public services.	development, education, awareness, government, service, rural, policy, among, like, state, health, economic, social, issue, group, growth, across, program, institution, scheme
T-6	Social development and gender dynamics in education and work	Addresses social development with an emphasis on education, gender dynamics, and the workplace.	social, development, education, research, employee, community, factor, woman, relationship, among, india, management, higher, gender, system, issue, like, need, analysis, economic
T-7	Adolescent stress and educational factors	Focuses on the factors affecting adolescent stress, mental health, and educational outcomes.	student, factor, adolescent, stress, research, social, higher, school, among, relationship, need, like, support, analysis, issue, teacher, environment, family, problem, mental health
T-8	Child and gender issues in family and education	Explores the intersection of child welfare, gender issues, and education within the family context.	child, woman, education, family, health, school, social, role, issue, need, like, work, gender, traditional, due, support, parent, among, often, factor
T-9	Corporate, political, and legal challenges in indian development	Discusses the interaction between corporate interests, political issues, and legal challenges in Indian development.	india, csr, social, medium, issue, political, company, like, state, conflict, public, development, policy, economic, role, indian, community, challenge, law, government

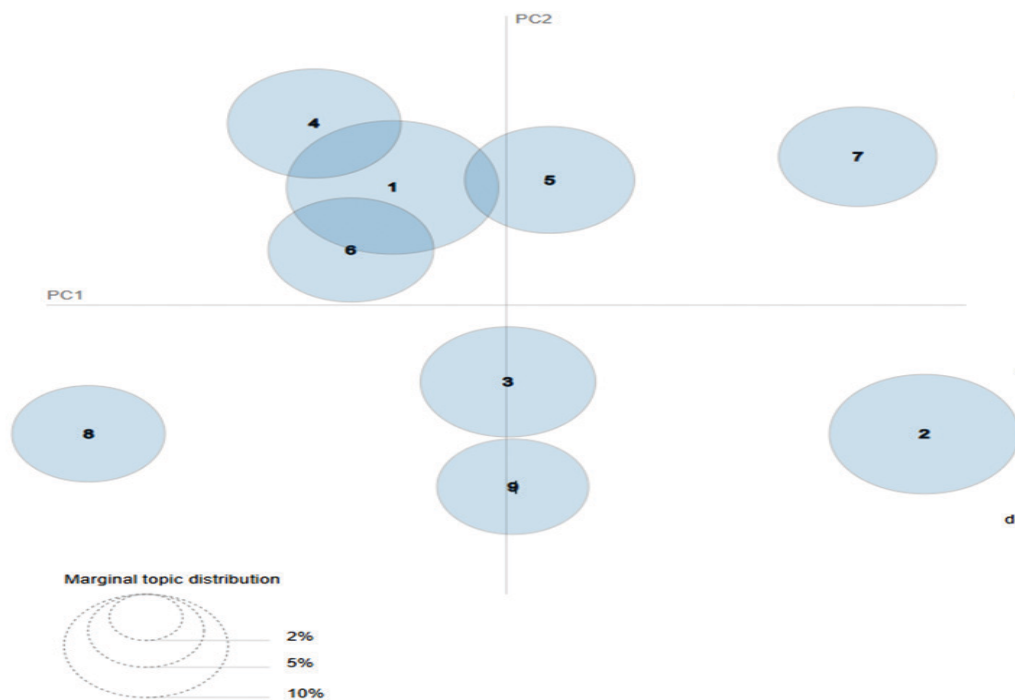


Figure 2. Intertopic distance map of topics.

themes. For instance, themes like T-1 (Rural Development and Women's Economic Roles), T-3 (Gender, Caste, and Political Issues in Social Development), T-4 (Cultural and Social Identities in the Workforce), T-6 (Social Development and Gender Dynamics in Education and Work), and T-8 (Child and Gender Issues in Family and Education) all focus on various aspects of gender. This indicates that there is a strong focus on understanding and improving the position of women in different contexts, such as rural development, workforce, education, or within family structures. The recurring presence of keywords like "woman," "gender," "education," and "social" underscores the importance of addressing gender inequalities and promoting women's empowerment in these areas.

It is also evident that there is increased focus on rural development and public welfare, particularly in the context of government policies and corporate social responsibility. Themes such as T-1, T-2 (Corporate Social Responsibility and Public Welfare), and T-5 (Government Policies and Rural Development Programs) highlight efforts to improve the conditions in rural areas through various initiatives. The occurrence of keywords like "rural," "development," "government," "education," and "health" in these themes suggests a concentrated effort to address socio-economic challenges in less urbanised regions. Moreover, T-2 and T-9 (Corporate, Political, and Legal Challenges in Indian Development) highlight the role of corporations and legal frameworks in shaping development outcomes, emphasising the need for responsible corporate behavior and effective policymaking to support sustainable development.

All themes have covered social and economic issues with concentration on gender roles, rural development

and impact of education and government policies. The intersectionality of gender, caste, and cultural identities (as seen in themes T-3 and T-4) has highlighted the complexity of societal development in diverse contexts. Further, the focus on adolescent stress and educational factors (T-7) points to an awareness of the challenges faced by younger populations, suggesting the need for supportive educational environments and mental health interventions. These themes have indicated a multidisciplinary approach to understanding and addressing the key challenges in societal development.

5.2 Dominant Topic Distributions

We calculated the count of theses and percentage share of the total theses for each dominant topic.

Table 2 presents the most-researched topics by showing the number of theses focused on each theme and their percentage share of the total. T-8 (Child and Gender Issues in Family and Education) leads with 69 theses, making up 15.40% of the total, showing that issues related to child welfare, gender, and education within the family context are a primary focus of researchers. Close behind is T-9 (Corporate, Political, and Legal Challenges in India), with 62 theses (13.84%) indicating a significant interest in how corporate and legal dynamics interact with political factors in shaping India's development. Themes such as T-6 (Social Development and Gender Dynamics in Education and Work) and T-3 (Gender, Caste, and Political Issues in Social Development) also feature prominently, reflecting ongoing concerns about gender, caste, and their influence on social and economic progress.

It is obvious that issues of gender dominate academic research, with a strong emphasis on how they intersect

Table 2. Distribution of Theses across the dominant topics

Dominant topic	Theme labels	Thesis counts	% Share in total theses
T-8	Child and gender issues in family and education	69	15.40
T-9	Corporate, political and legal challenges in india	62	13.84
T-6	Social development and gender dynamics in education and work	57	12.72
T-3	Gender, caste, and political issues in social development	49	10.94
T-4	Cultural and social identities in the workforce	47	10.49
T-5	Government policies and rural development programs	44	9.82
T-1	Rural development and women's economic roles	43	9.60
T-2	Corporate social responsibility and public welfare	39	8.71
T-7	Adolescent stress and educational factors	38	8.48

with education, work, and political structures. Themes focusing on rural development and government policies, such as T-5 (Government Policies and Rural Development) and T-1 (Rural Development and Women's Economic Roles), while having slightly lower shares (around 9-10%), still account for a substantial portion of the research output. This finding suggests that rural development, particularly concerning women and policy implementation, remains a critical area of study, although it is slightly overshadowed by broader social and gender concerns. Furthermore, T-2 (Corporate Social Responsibility) and T-7 (Adolescent Stress and Education) have the lowest number of theses but still represent important research areas. Child welfare, gender, and social development have emerged as dominant themes in academic research, and the research landscape is shaped by both social justice issues and the impact of corporate and political structures on development.

5.3 Individual Topic Trends: Temporal Evolution of Research Focuses

To understand the evolving trends of individual topics over the duration of the study, year-wise trends were plotted for each dominant topic.

Figure 3 shows the trends in the number of theses across the nine different themes (T-1 to T-9) over the years. Several themes, such as T-1 (Rural Development and Women's Economic Roles) and T-5 (Government Policies and Rural Development Programs), exhibit fluctuating research interests, with notable peaks in earlier years such as 2014 and 2012, respectively, followed by more stable numbers later on. Themes like T-7 (Adolescent

Stress and Educational Factors) also showed an early spike, particularly in 2012, but overall, these themes appear to have had declining or stabilising interest in recent years. This pattern suggests that, while these areas were once prominent, research activity has either plateaued or shifted towards other priorities.

By contrast, themes such as T-6 (Social Development and Gender Dynamics in Education and Work), T-3 (Gender, Caste, and Political Issues in Social Development), and T-9 (Corporate, Political, and Legal Challenges in India) showed a gradual increase in research activity, especially after 2015. These themes have attracted more attention, reflecting a growing academic focus on gender dynamics, social development, and corporate and political challenges. The steady growth in these areas suggests that they are becoming increasingly relevant in contemporary research, likely driven by ongoing societal changes and the importance of these issues in shaping narratives of modern development. The graphs also highlight the following:

Themes such as T-1 (Rural Development and Women's Economic Roles) and T-5 (Government Policies and Rural Development Programs) had strong early interest, with peaks in 2014 and 2012, but experienced less research activity after that. This suggests that, while rural development and government policy were significant topics earlier, academic interest may have shifted to other areas, or these issues may now be considered more saturated fields of study.

5.3.1 Growing Interest in Social and Corporate Issues:

Themes such as T-6 (Social Development and Gender Dynamics in Education and Work), T-3 (Gender, Caste, and Political Issues in Social Development), and T-9

(Corporate, Political, and Legal Challenges in India) show a noticeable upward trend in thesis counts over time, particularly after 2015. This suggests an increasing academic focus on the intersection of gender, social justice, corporate influence, and legal challenges. These topics are becoming increasingly important as society deals with modern complexities, such as gender equality, corporate responsibility, and political dynamics.

5.3.2 Intermittent Focus on Certain Topics

Themes such as T-8 (Child and Gender Issues in Family and Education) and T-2 (Corporate Social Responsibility and Public Welfare) show intermittent spikes in research output. This indicates that, while these areas remain important, they may attract attention periodically, possibly driven by specific events, policies, or societal shifts that renew interest in these subjects for brief periods.

5.3.3 Declining or Stabilising Themes

Topics such as T-7 (Adolescent Stress and Educational Factors) had early spikes in interest, particularly in 2012, but they have since seen less research focus. This suggests that while these issues were once highly researched, academic attention may have shifted to emerging topics or there may be reduced novelty in these fields.

These insights reflect evolving academic interests, with some traditional areas of research stabilising, whereas others, particularly those related to social, gender, and corporate challenges, have gained momentum in recent years. This comprehensive analysis of research trends not only illuminates the evolving landscape of societal concerns but also sets the stage for a deeper exploration of the factors driving these shifts in academic focus.

6. DISCUSSION

Social issues are problems that affect individuals or groups within society and are perceived as undesirable. These issues arise due to continuous societal changes and impact social structures, values, and institutions. Social issues revolve around principles of equality and dignity, encompassing social protection, gender equality, and access to resources.

In India, social reform movements have sought to address societal concerns since ancient times. Large-scale efforts to eradicate social issues gained momentum in the 18th and 19th centuries, and the Indian national movement in the 20th century further focused on social upliftment. Despite independence, major social issues persisted, prompting governments to introduce laws and policies for their eradication. Research on social issues before 2000 primarily covered untouchability, the caste system, and poverty. Over time, studies expanded to discrimination in medical access, gender issues, and psychological concerns. This study analyses trends in social issue research, identifying key factors influencing these trends. It examines research between 2011 and 2020 and finds a general upward trend, with minimal fluctuations. Research output was low between 2011 and 2013 but increased in later years due to improved academic infrastructure and government support.

The study categorises social issue research into nine dominant themes using a topic modeling method, based on 20 keywords per theme. These themes predominantly cover social and economic concerns, focusing on gender roles, rural development, and the impact of education and government policies. The research highlights women's evolving role in society and the challenges they face. Additionally, studies examine rural development, public welfare schemes, and corporate social responsibility

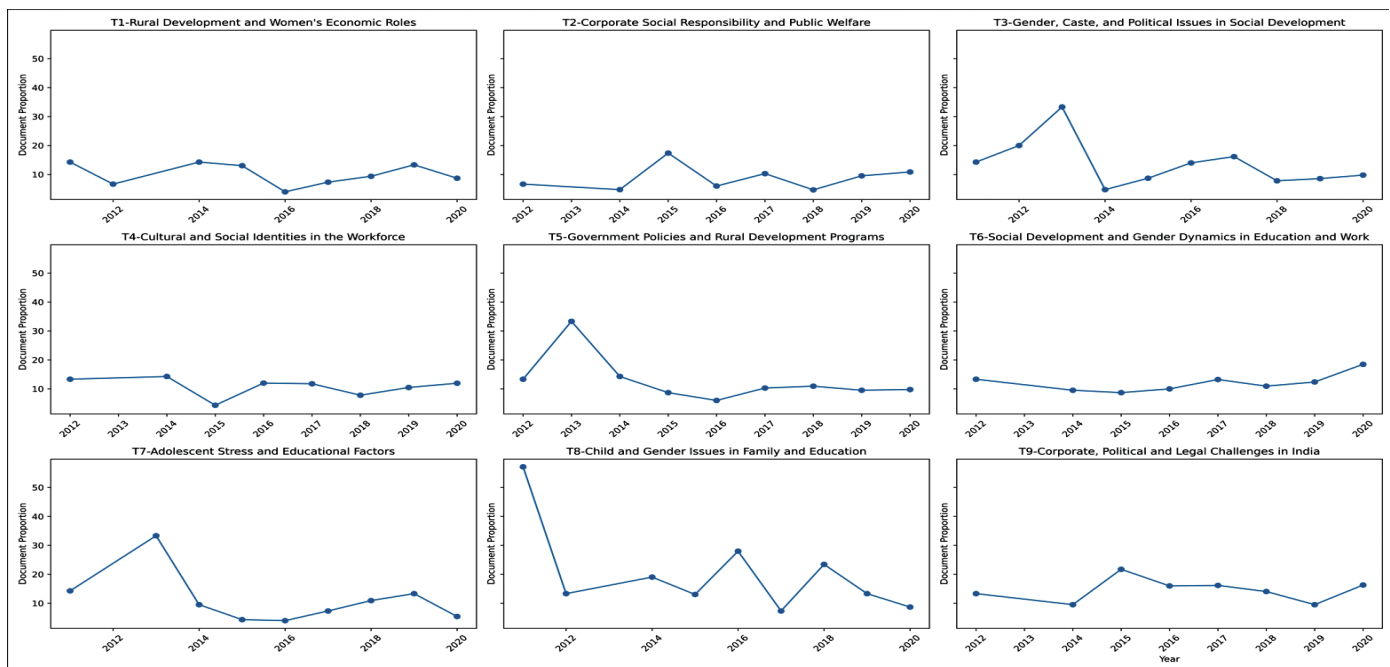


Figure 3. Individual topical trends for each dominant topic.

(CSR), underscoring the corporate sector's role in social development. Effective policy-making and implementation are essential for overall progress. Mental health issues, particularly stress and anxiety among adolescents, also emerge as significant themes. Findings suggest that school and home environments should be supportive to promote mental well-being. The research underscores the need for greater attention to mental health concerns.

Further analysis of subthemes reveals that topics related to women, children, families and education are the most extensively researched. This indicates growing scholarly interest in child welfare, gender equality and education within the family context, reflecting the influence of government policies promoting the welfare of women and children.

Trends within the nine themes show that certain topics, such as rural development, women's economic roles, and adolescent stress, fluctuate over time, with peak interest in the initial years followed by a decline or stabilisation. These variations reflect shifts in societal priorities. However, themes like social development and gender dynamics show a steady increase in research output, indicating growing academic focus on these areas.

The study demonstrates that research on social issues has evolved in response to societal changes, policy initiatives, and shifting academic interests. It highlights the importance of continued research in gender studies, mental health, rural development, and corporate responsibility to address emerging social challenges effectively.

7. CONCLUSION

Many challenges arising from social issues go beyond individual control. If society is resolute, social problems can be addressed. These social problems impede the progress of society. Therefore, we must collectively confront and strive to eliminate them for the greater good. Research on social issues plays a crucial role in policymaking. The suggestions and recommendations provided by the researchers serve as a guiding force in policy formulation. The research studies on social issues have highlighted the problems and challenges faced by the people in their day-to-day life. By conducting an in-depth analysis, these studies provided valuable feedback on existing government programs and offered necessary recommendations for better implementation of schemes and suggested the necessity of new initiatives and policy changes. A clear idea about social issues helps the policymakers in developing policies to benefit a large number of people in the society.

Social issues including inequality, gender discrimination, access to education, healthcare, and environmental sustainability remained as major challenges in the governance. These challenges continued to affect millions of people, especially marginalised communities. Therefore, the studies on social issues help the policymakers, civil society, and academic researchers to get clear idea on the existing social issues and formulate the policies accordingly. The study shows that during the period from 2011 to 2020,

the doctoral research in social sciences in India has been concentrating more on the contemporary social concerns and provided valuable insights and potential solutions to longstanding social problems. By identifying the trends in the research, the present study uncovers shifts in research priorities and also reflects on the growing interdisciplinary approaches in studying social issues. It is found that research scholars are increasingly combining the fields such as law, economics, and cultural studies to tackle multifaceted challenges faced by the society. In so doing, it provides a foundation for future research and policy initiatives aimed at promoting social equity and sustainable development.

The study concludes that there is an increasing focus with an upward trend in research on social issues, particularly on themes most relevant to contemporary society that require immediate policy interventions. Some themes exhibited sporadic increases and decreases, indicating that, while these issues are important, they attracted researchers' attention only during specific periods. Another broad trend observed was a stable or gradual decline in research on themes related to stress and educational factors.

This study offers valuable insights into the dominant themes of doctoral research on social issues in India. However, there are several limitations that must be acknowledged. First, the dataset used in this study comprised only those theses available in the Shodhganga repository. As a result, doctoral research from universities whose theses are not included in this repository are not represented which potentially limits the comprehensiveness of the findings. Furthermore, this study included theses written in English, which may have introduced bias. Including theses written in other Indian languages might have yielded different outcomes and a more nuanced understanding of the research landscape. Importantly, the study relied on topic modeling to identify key themes and trends across the corpus of doctoral theses. While these techniques are valuable and well recognised for highlighting prominent themes, future research should consider integrating manual analysis alongside quantitative methods to provide a more holistic view of social issues research in India. Future research could also expand the scope of this study by conducting comparative analyses with doctoral research from universities in other countries.

REFERENCES

1. Cherry F. The nature of The Nature of Prejudice. *J of the History of the Behavioral Sci.* 2000;36(4):489-98. doi: 10.1002/1520-6696(200023)36:4<489::AID-JHBS13>3.0.CO;2-N
2. Web of Science. <https://www.webofscience.com/wos/woscc/summary/b9f316d7-2898-43cb-99fe-77afdc040ea1-015be5c2f7/relevance/1>. (accessed 18 Apr 2025).
3. Clarivate. Clarivate and the Chinese academy of sciences release annual collaborative report to identify 125 research fronts. 2024. Available from: <https://img06.en25.com/Web/ClarivateAnalytics/%7B357fb5b8->

- a40d-466e-8712-3aaa0aa1778b%7D_2024_Research_Fronts-ActiveFields_LeadingCountriesRegions-EN.pdf. (accessed 18 Apr 2025).
4. Business research company. Social sciences and humanities services market to hit \$116.16 Billion by 2029: Decoding growth drivers, trends, and key players. 2025. Available from: <https://www.openpr.com/news/3847395/social-sciences-and-humanities-services-market-to-hit-116-16>. (accessed 18 Apr 2025).
 5. Hagen L. Content analysis of e-petitions with topic modeling: How to train and evaluate LDA models? *Inf Processing and Managm*. 2018;54:1292-307. doi: 10.1016/j.ipm.2018.05.006
 6. Baranowski M. Epistemological aspect of topic modelling in the social sciences: Latent dirichlet allocation. *Przegląd Krytyczny*. 2022;4(1):7-16. doi: 10.14746/pk.2022.4.1.1
 7. Jelodar H, Wang Y, Yuan C, Feng X. Latent dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*. 2017;78:15169-11. doi: 10.1007/s11042-018-6894-4
 8. Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*. 2022;7:1-16. doi: 10.3389/fsoc.2022.886498
 9. Lee H, Jang SJ, Sun FF, Broadwell P, Yoon S. Comparing themes extracted via topic modeling and manual content analysis: Korean-language discussions of dementia on Twitter. *Advances in informatics, management and technology in healthcare*. 2022;295:230-3. doi: 10.3233/SHTI220704
 10. Lamba M, Madhusudhan M. *Text Mining for Information Professionals: An Uncharted Territory*. Cham: Springer; 2022.
 11. Cui N, Malleson N, Houlden V, Yan Y, Comber A. Using Twitter to understand spatial-temporal changes in urban green space topics based on structural topic modelling. *Cities*. 2025;157:105601. doi: 10.1016/j.cities.2024.105601
 12. Aminpour N, Saidi S. Unveiling mobility patterns beyond home/work activities: A topic modeling approach using transit smart card and land-use data. *Travel Behaviour and Society*. 2025;38:100905. doi: 10.1016/j.tbs.2024.100905
 13. Mennig P. Who cares about agriculture? Analysing German parliamentary debates on agriculture and food with structural topic modeling. *Food Policy*. 2025;130:102788. doi: 10.1016/j.foodpol.2024.102788
 14. Pang N, Sun M, Zhu H. Representation of migrant workers in Chinese newspapers: Combining topic modeling and critical discourse analysis. *Digital Scholarship in the Humanities*. 2024;39(4):1092-111. doi: 10.1093/llc/fqae062
 15. Hu N, Ho KC, Fan PS. Malaysian Chinese folk beliefs on facebook based on LDA topic modelling. *Humanities and Soc Sci Commun*. 2024;11(1):1-14. doi: 10.1057/s41599-024-03066-6
 16. Özmantar MF, Gökdağ K, Hangül T, Agaç G. Research themes and trends in the field of teacher educators: A topic modelling study. *Teaching and Teacher Education*. 2024;148:104696. doi: 10.1016/j.tate.2024.104696
 17. Kumar V, Thakur K. Using text analysis to study doctoral-level library and information science research trends in India. *Annals of Lib and Inf Studies*. 2022;69(3):191-202. doi: 10.56042/alis.v69i3.58719
 18. Ismail A, Munsil H, Yusuf AM, Hijjang P. Mapping one decade of identity studies: A comprehensive bibliometric analysis of global trends and scholarly impact. *Social Sciences*. 2025;14(2):92. doi: 10.3390/socsci14020092
 19. Hamdan W, Alsuqaih H. Research output, key topics, and trends in productivity, visibility, and collaboration in social sciences research on COVID-19: A scientometric analysis and visualisation. *SAGE Open*. 2024;14(4). doi: 10.1177/21582440241286217
 20. Thakur K, Kumar V. Application of text mining techniques on scholarly research articles: Methods and tools. *New Review of Academic Librarianship*. 2021;28(3):279-302. doi: 10.1080/13614533.2021.1918190
 21. Sukhija N, Tatineni M, Brown N, Moer M, Rodríguez P, Callicott. Topic modeling and visualisation for big data in social sciences. In: El Baz D, Bourgeois J, editors. 2016. *International IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCCom/IoP/SmartWorld)*. France: IEEE; 2016. p. 1198–205.
 22. Mo Y, Kontonatsios G, Ananiadou S. Supporting systematic reviews using LDA-based document representations. *Systematic Reviews*. 2015;4. doi: 10.1186/s13643-015-0117-0
 23. Nanda G, Jaiswal A, Castellanos H, Zhou Y, Choi A, Magana A. Evaluating the coverage and depth of latent dirichlet allocation topic model in comparison with human coding of qualitative data: The case of education research. *Machine Learning and Knowledge Extraction*. 2023;5(2):473-90. doi: 10.3390/make5020029

ACKNOWLEDGEMENTS

The study was funded as a major project by Indian Council of Social Science Research, New Delhi vide File No. 02/199/GN/2021-22/ICSSR/RP/MJ.

CONTRIBUTORS

Dr. Manorama Tripathi holds a PhD in Library and Information Science from Banaras Hindu University, Varanasi. She is currently serving as the Acting Librarian at the Central Library, Jawaharlal Nehru University, New Delhi. Her research interests include: the Information-seeking behavior of researchers, Innovative library services, and Scientometrics.

Her contributions to the study are conceptualising the idea, Literature review, Manuscript draft preparation and overall supervision.

Dr. Narasaiah Inugurthi is Deputy Director at Archives on Contemporary History, JNU. He had completed his Masters in Modern History from Sri Venkateswara University, Tirupati and completed M.Phil. and PhD from University of Hyderabad. He has also worked as Post Doctoral Fellow at Tata Institute of Social Sciences, Hyderabad.

His contributions to the study are assistance in deciding the labels of dominant themes, manuscript writing, proof reading and editing.

Mr. Ankit Yadav is Semi-Professional Assistant at National Law University Delhi, where he contributes to various academic and administrative functions. With a keen interest in mathematics and computer science.

His contributions to the study are data collection and manuscript formatting.

Dr. Vinit Kumar is an Assistant Professor at the Department of Library and Information Science, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow, India. He has more than 13 years of experience in teaching and research in Library and Information Science. His research interests are Text mining, Research methods, Cultural heritage information, and Social media analytics. His contributions to the present study are data processing, analysis, tabulation, graphical presentation of data, manuscript draft preparation, proof reading and editing.