Swachh Bharat Abhiyan Mission
Tamil Nadu

Executive Summary

The goal of this study is to evaluate specific aspects of the Swachh Bharat Mission. To understand this, we developed a sentiment polarity indicator based on the media data extracted from a variety of sources such as newspapers and magazines both on the state and national levels. Geographically, the study will focus specifically on rural areas in the state of Tamil Nadu. The data we have gathered consists of 6797 pieces of written textual data in more than 260 newspapers and other media types covering the time period from August 2014 until April 2016.

Media data is distributed between five main aspects, referred to as themes, of the mission: behavioral change; construction quality; information, education and communication activities; gender, and third party involvement. Quantitative analysis of the data indicates that there is a clear difference in theme preferences between national and Tamil papers. Tamil publications tend to focus on such topics as behavioral change (51%) and construction quality (32%) with the rest of the categories representing only 17% of the overall coverage. Nation-wide publications, on the other hand, concentrated on issues of construction quality (35%); information, education and communication activities (21%) as well as the third party involvement (19%). It is also worth mentioning that Gender is the only theme which received the least amount of attention from both media types.

Sentiment analysis of the data provides further details on how the SBM themes have been covered since the inception of the mission. From a temporal perspective, for instance, Tamil publications expose an overall decrease in their positive content and a slight increase in the negative content. The Nation-wide publications, on the other hand, expose a decrease in negative content. Based on the interrelation of positive and negative sentiments, we further derived an indicator of sentiment polarity. Polarity indicator provides an insight into whether a given theme is perceived positively, negatively or neutrally by the media. While all five themes were covered positively by the media, Gender was covered more positively than any other theme with a positive score of 20. The theme of behavioral change, on the other hand, was ranked as the least positive with the score of 13.

Table of Contents

1. Background and Objectives

Swachh Bharat Mission (SBM) is a major initiative of the current government, launched by Prime Minister Modi on the second of October, 2014.[1] The objective of the mission was to accelerate the process of achieving universal sanitation coverage and to increase the focus on responsible hygiene and sanitation. One of the key planks of the program is to build nearly 12 crore toilets in rural India to ensure an Open Defecation Free India by October of 2019.

The response to the program has been overtly positive, being credited to bring the long-standing issues of sanitation to the forefront of public attention as well as a policy priority. However, while some progress has been achieved[2], there are still outstanding issues in the implementation of the program. Several reports have found that the toilets remain underused[3] or used for unrelated purposes (such as storehouses).[4] Thus, it is important to critically analyze the performance of the mission to understand some major flaws or remaining issues in the implementation of the program.

Conceptually, the study follows two steps. Firstly, the data will be analyzed with regards to the content of the main themes summary of which is presented in the Table 1. Secondly, the sentiment polarity has been created using tools of the Natural Language Processing (NLP).[5] The sentiment polarity score provides for a common denominator for analysis of mission's aspects. Namely, the indicator was useful in scoring and ranking the above-mentioned themes based on sentiment of the media coverage.

---

[1] http://www.firstpost.com/india/govet-to-launch-intensive-nationwide-swachh-bharat-campaign-from-25-sept-2443694.html

[2] http://economictimes.indiatimes.com/news/poltiics-and-nation/52-citizens-find-their-city-cleaner-after-600-days-of-swachh-bharat-mission/articleshow/52702533.cms

[3] http://www.dailymail.co.uk/indiahome/indianews/article-3196966/Modi-s-Swachh-Bharat-mission-built-80-lakh-new-toilets-people-using-them.html

[4] http://timesofindia.indiatimes.com/city/bareilly/UP-villagers-prefer-open-fields-raze-Swachh-loos/articleshow/50582495.cms

[5] See the methodology for further details.

2. Data and Methodology

Using methods of natural language processing, we aim to evaluate the media perception of the Swachh Bharat Mission in the region of Tamil Nadu. To account for a variety of perspectives, both Tamil and India-wide publications are included in the study. Articles for both types of publications are included in the study only when they discuss the mission for the state of Tamil Nadu. Data processing and analysis is divided into several steps. Figure 1 provides an overview of the main methodological steps used for the present analysis. Each of the steps is further described in more detail.

Figure 1: Methodology Overview
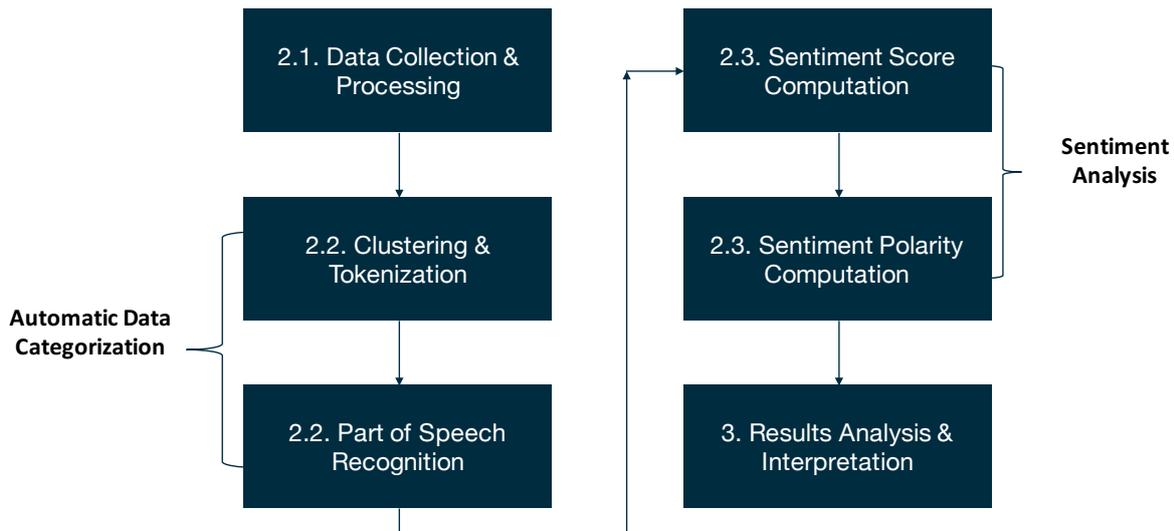
2.1. Data Collection and Processing

Data used in this report consists of a set of publications on Swachh Bharat Mission (SBM) collected from two sources: Google News Aggregator for English publications and a collection of digitally available publications in Tamil. All available data on the topic from September of 2014 and April 2016 has been collected, amounting to 6797 news pieces from more than 12 regional and 235 national publications. After the collection of textual data, both English and Tamil datasets were cleaned. Semi-automated cleaning of the data allowed for the elimination of irrelevant information, such as stop words, punctuation as well signs that do not carry meaning.[6] Due to interest in specific aspects[7] of the project, search, and analysis of the articles were narrowed down to five specific themes:

a) **TPI,** or Third Party Involvement**:** the category which accounted for the role of NGOs, contractors or other outside agencies;
b) **BC,** or Behavioral change: the category which accounts for issues related to change in behavior especially with regards to open defecation fields in rural areas;
c) **IEC**: Information, Education and Communication Activities of the project;
d) **Gender**: articles discussing involvement of women or gender-related aspects of the mission;
e) **Quality:** articles that dealt with the quality of construction on the ground.

The goal of this study is to evaluate specific aspects of the SBM mission. To understand how those aspects rank based on media perceptions, we developed a sentiment polarity based on the data extracted from a variety of sources such as newspapers and magazines both on state

---

[6] Signs such as &, various types of brackets and others.

[7] Aspects were identified by the PAC team.

and national level. Geographically, the study will focus specifically on rural areas in the states of Tamil Nadu. The dataset for the study on Tamil Nadu consists of 6797 pieces of written textual data from more than 260 newspapers covering the time period from August 2014 until April 2016. The following table provides a distribution of the dataset according to the main themes of the study.

Table 1: National and Tamil Data

Language	Behavioral Change	Construction Quality	IEC	Gender	Third Party Involvement
English	753 articles	1670 articles	1016 articles	420 articles	880 articles
Tamil	1052 articles	666 articles	104 articles	104 articles	132 articles

2.2. Data Categorization

Two steps are performed at the stage of data categorization. Initially, the collected data is categorized according to the five **clusters**, or themes, mentioned above. In the next step, the raw textual data is **tokenized**. That is, natural language processing algorithms – using pre-established dictionaries – break up the sequence of data into pieces of speech. Depending on the purpose of the assignment, the data can be tokenized into words, specifically defined keywords, phrases, and even symbols. Since the present analysis will be engaged in analysis of words within a given text, the data has been tokenized down into individual words.

The tokenized dataset is further analyzed for **parts of speech**. Specifically, we use part-of-speech (POS) tagging methods to classify words as nouns, verbs, adverbs and other based on their syntactic function. Since some parts of speech – as nouns or pronouns – do not carry any sentiment, POS tagging helps to create a subset of relevant sentiment-carrying elements such as adjectives and adverbs into a corpus of text ready for analysis. For the purpose of the present study, we updated the algorithm to account not only for adjectives but also for adverbs and a collection of SBM-specific nouns and phrases (for instance, "open defecation fields"). The algorithm was also trained to understand negation and emphasis in sentences. Words such as "not" or "no" which invert the meaning of words and consequently the direction of sentiment have been controlled for.

2.3. Sentiment Analysis

Once the creation of the sentiment-carrying corpus is completed, we **compute sentiment scores**. Namely, the data is scored through the Sentiwordnet dictionary for positive and negative sentiment.[8] After each sentence is assigned a positive and a negative score, the scores of all the sentences are obtained and averaged to represent the article score. Using positive and negative sentiments, we further obtain the **sentiment polarity** score for each of our five themes. The polarity score is created using the following formula:[9]

---

[8]Sentiwordnet is a lexical database that is useful for opinion mining. It assigns a score of positivity and negativity to each word derived from the Wordnet dictionary. Wordnet is a large database that groups words and acts as a thesaurus ad backend for natural language processing tools. It is maintained by Princeton University. For detailed information on Sentiwordnet see https://wordnet.princeton.edu. For more information on Sentiwordnet see Esuli, A., & Sebastiani, F. (2006, May). "Sentiwordnet: a publicly available lexical resources for opinion mining. In proceedings of LREC (Vol. 6, pp. 417-422).

[9] For a more detailed information of the formula and the method please see: William Lowe, Kenneth Benoit, Slava Mikhaylov, and Michael Laver (2011). "Scaling Policy Preferences From Coded Political Texts." Legislative Studies Quarterly 26(1, Feb): 123-155.

**Polarity Score = ln (PositiveSentiment + 0.5) – ln (NegativeSentiment + 0.5)**

The polarity score is useful in understanding whether the given theme is covered in positive, negative or neutral tone. The formula provides us with a score which will equal to zero, some positive value or some negative value. The value of zero will signify that the sentiment for the given category is neutral and does not expose any positive or negative sentiment. Positive score and negative score will signify positive and negative sentiment accordingly. Besides providing a sense of sentiment for a given theme, sentiment polarity is useful in comparing the sentiment of two or more themes.

3. Sentiment Analysis

3.1. Coverage Overview

Media reflects the current political agenda as well as the perception towards the discussed topics. But not all topics get the same attention, i.e. same coverage. Thus, looking at the extent to which some aspects of a policy are prioritized over others, gives a first idea of policy priorities. Figure 2 shows the media coverage of the themes described in section 2, by illustrating the number of articles that covered the respective theme. The numbers are then expressed as a share of the total national articles (light blue) and Tamil articles (dark blue) on SBA. For example, around every fifth article in national newspapers (21%) covered some aspects of information, education and communication of SBA.

Figure 2: Theme coverage by national and local papers
In % of all articles
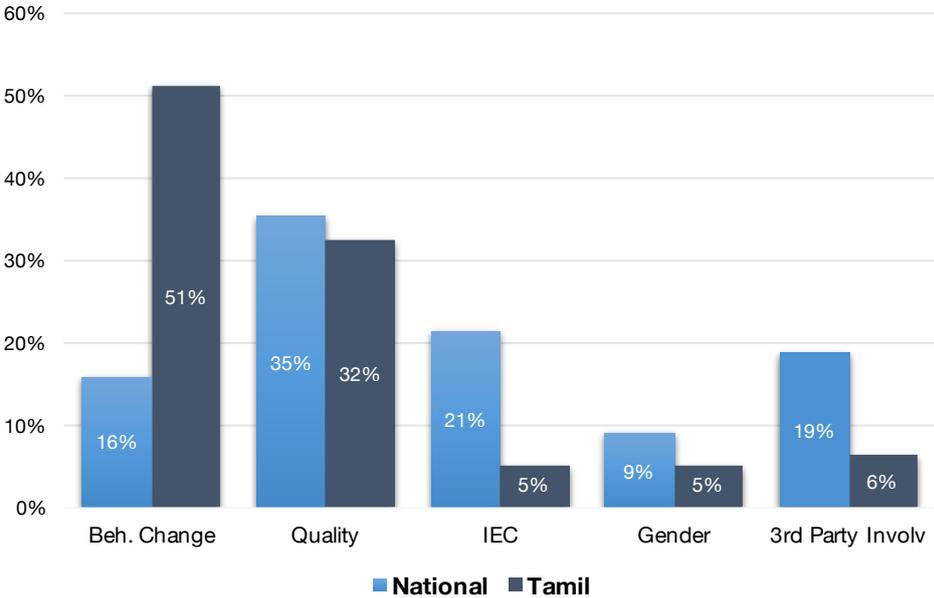
As can be noticed topic preferences differ widely between Tamil and English publications. Namely, more than half of Tamil publications (51%) focused on articles that discussed how the mission does or does not trigger behavioral change. These are the articles mostly focusing on such issues as the proper usage of newly built toilets (as opposed to usage of open defecation fields, for instance) and other topics focusing on how the mission affects behavior of individuals in Tamil Nadu. The second most popular topic in Tamil papers was the quality of construction. That is, 32% of all Tamil articles focused on construction quality of the toilets and other issues related to construction. The other three categories – gender, IEC and third party involvement – represented only 17% of the Tamil publications.

While national publications also focused on construction quality (35%), the papers also concentrated on the information, education and communication activities (IEC) aspect of the mission. To be more precise, with 21% of articles focusing on ICE, one fifth of the national papers discussed how the mission propagates information about sanitation and cleanliness. And approximately another fifth of articles (19%) focused on third party involvement in the mission. That is, national papers also discussed how NGOs, subcontractors and other organizations were involved in the mission. While focus of national and Tamil papers varies considerably, there is one theme on which both types of publications devoted comparatively equal amount of coverage. Namely, articles discussing how the mission affects gender and women, represented by the theme Gender in this study, received the the least amount attention

when compared to the rest of the categories in both datasets. National media covered gender 9% of the time and Tamil media only 5%.

Having an overview of the themes' importance is indicative of what is more discussed in the media. This quantitative overview of the coverage, however, does not indicate whether the sentiment in those article was negative or positive coverage for SBA. The next chapter of the study employs sentiment analysis methods to understand how themes were perceived within the public sphere.

3.2. Sentiment Change over time

This part of the study is preoccupied with analyzing the sentiment of media towards the mission. Taking a temporal perspective, we also analyse how negative and positive sentiments towards SBM have changed, or have not, since its inception. That is, we will track how positive and negative statements in the newspaper content changed over time relative to the third quarter of 2014 (base quarter). Thus, the goal of the chapter is to provide insights about temporal change of media sentiment within a given dataset. Analyzing the base quarter for national publications, one can observe that positive and negative newspaper sentiments were close to zero both for National and Tamil publications. That is, during the third quarter of 2014, media coverage of the mission was neutral and newspapers did not expose extremely positive or negative sentiments.

But how did the sentiment evolve over time? In order to answer this question, we looked at the change of the sentiment from quarter to quarter. Furthermore, we differentiated between national and Tamil newspapers. The results are summarized in Figure 3 below. Looking at the picture overall, one can conclude that positive sentiment fluctuated more or less around zero while the negative content of the coverage decreased. Taking Q4/2015 as an example, one can observe that there was almost 7% increase in positive sentiment in national publications when compared to Q3 of 2014. Similarly, we observe a 25% decrease of the negative sentiment in the fourth quarter of 2015 comparative to Q3 of 2014.

Figure 3: Sentiment change over time
In % change relative to the Q3 of 2014
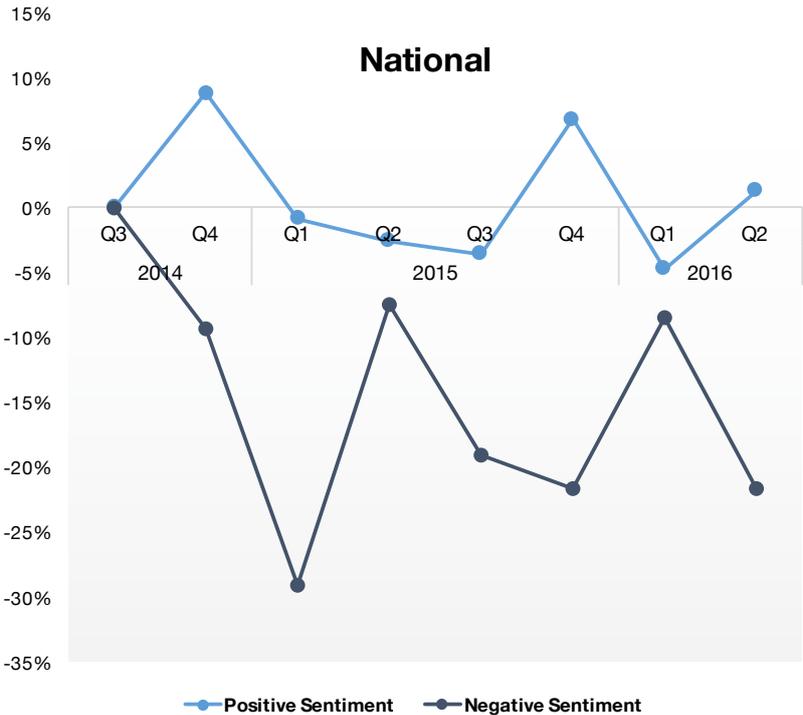
Sentiment change within Tamil publications provides a different picture. First of all, as opposed to the National publications, it is the positive content that has been decreasing when compared to the base quarter. Looking into the Q1, Q2 and Q3 of 2015 on Figure 4, for instance, one observes a steady decline in the positive sentiment of the SBM coverage. The decreasing positive trend is steady across the entire period with the exception of the last quarter of 2015. The detailed study of the articles for that period provides indicates that the sudden increase of the positive media coverage was caused by the Rotary club's announcement to contribute to the mission.

Figure 4: Sentiment change in Tamil papers
In % change relative to the Q3 of 2014
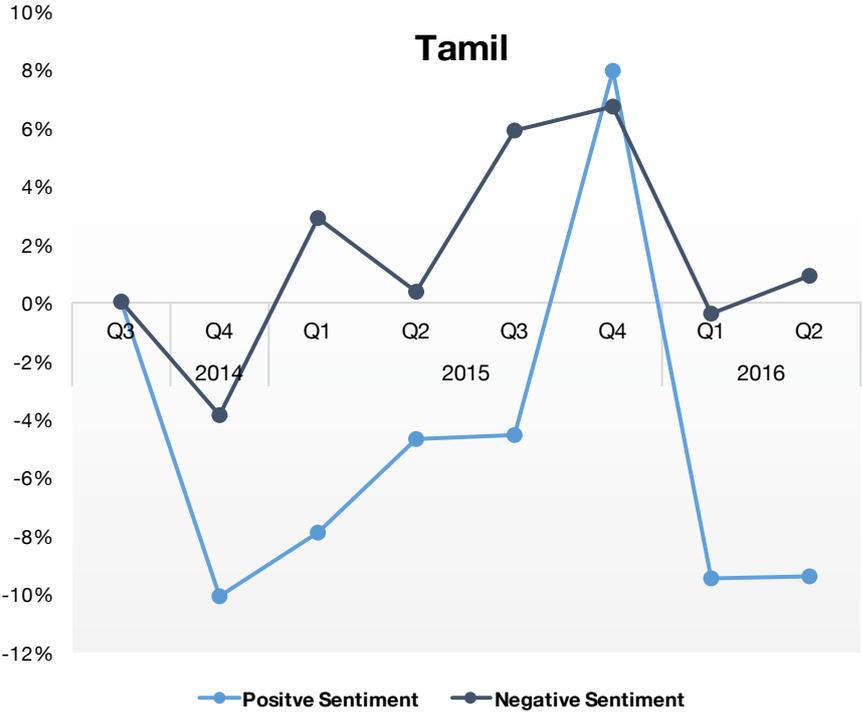
Secondly, Tamil data, unlike its national counterpart, has seen an overall increase of the negative sentiment for the entire period. Starting with a 3% increase in the negative sentiment in Q1 of 2015, the trend finishes with 1% increase in the second quarter of 2016.This finding is critical since it could be indicative of the fact that the media perception of the mission has become more negative.

3.3. Sentiment Ranking

While having an idea of the changes in sentiment, the analysis in the previous chapter lacks a focused approach to the five SBM themes that have been identified as important aspects of the study. It would be interesting, for instance, to understand how media coverage themes can be compared as well as the detailed context of each theme. To understand this, we developed a polarity indicator which enables to compare several themes of the mission. As suggested in the literature[10], using sentiment polarity provides a comparative aspect within the sphere of natural language processing. The results of implementing these methods for the present study indicate the following insights. First of all, taken together, media coverage of the SBM has been positive. While calculations of sentiment polarity do allow for negative values to occur, none of

---

[10] Will, Lowe, Benoit, Kenneth, Slava, Mikhaylov, & Laver, Michael. 2011, "Scaling Policy Preferences from Coded Political Texts." Legislative Studies Quarterly, vol. 36, no.1, :123–55.

the themes in the present study received an absolutely negative scoring. Nevertheless, most of the rankings are also close to zero, indicating that for two categories specifically – third party involvement and behavioral change – the sentiment polarity is positive only by a small margin.

Detailed ranking of the five themes are provided by the Figure 5. With the positive score of 20, Gender is the theme that has been covered most positively by the media. Gender is closely followed by the themes on third party involvement and IEC. And the last two categories are Quality and Behavioral Change.

Figure 5: Sentiment Ranking
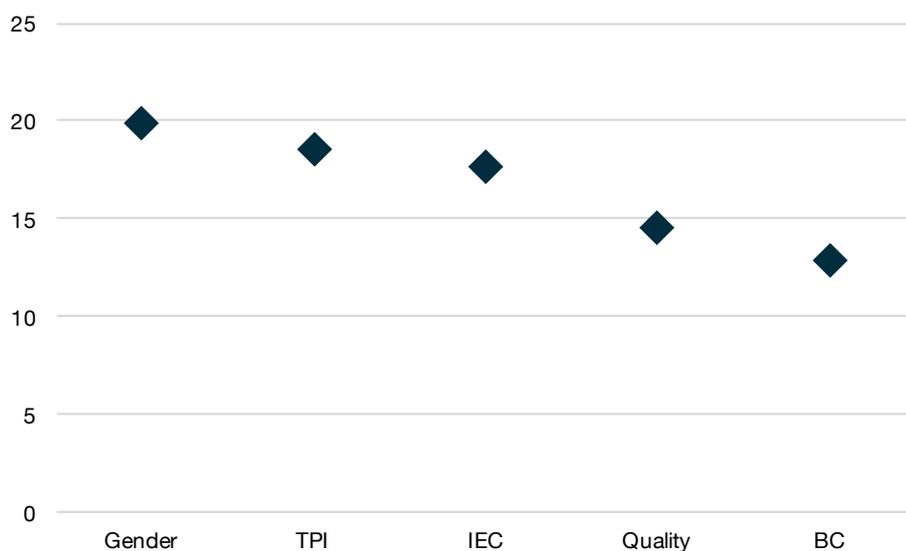
We further analyzed each of the categories in detail with the goal of understanding the context behind the sentiment ranking. One could wonder, for instance, about specific terms and topics that are discussed within the theme on Gender. What are the most frequent topics and can the overall positive ranking of Gender be attributed to those specific words and concepts? To understand this, we pulled five most frequent terms that have been used within each theme. Figure 6 on the following page visualizes those terms for each of our themes. Furthermore, the Figure provides a polarity score for each of the words within the theme. Seeing a negative polarity score of the word "work" within Gender, for example, we can infer that media has mostly exposed negative sentiment of that word when discussing gender within the SBM.

First of all, we looked into narratives within the category of **Gender** which ranked as the most positive theme. One of the two highest references are towards the National Rural Employment Guarantee Act (or Nrega). As it is known[11], Nrega is especially beneficial for the most vulnerable groups of rural population, including women. Other interesting concepts that occur within the dataset are work, devi[12] as well as training. Articles covering these terms speak particularly of such issues as provision of safe toilets for women, women's engagement within the project as well as their employment.[13] But there are also articles[14] taking a more critical approach with regards to gender and the mission. Some articles argue, for example, that the mission does not at all address gender-specific issues such as violence against women, sexual violence, problems related to reproductive health, discrimination faced by women due to social and cultural norms, feminization of poverty and many others.
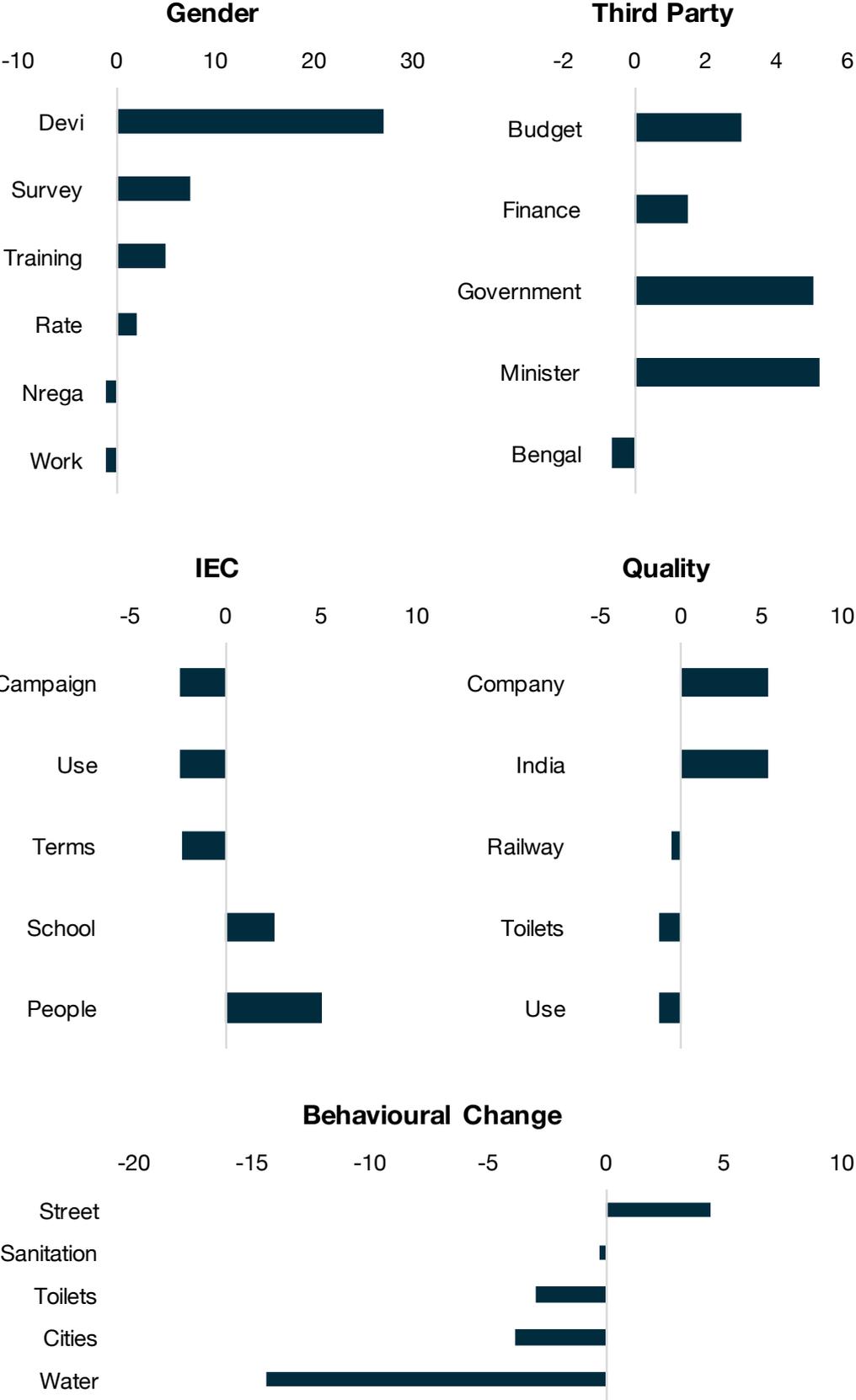
---

[11]http://www.indiaspend.com/cover-story/half-full-half-empty-10-years-of-nrega-80147
[12] Woman
[13]http://articles.economictimes.indiatimes.com/2016-04-09/news/72186451_1_manual-scavengers-swachh-bharat-abhiyan-dalit
[14] http://indianexpress.com/article/india/india-others/govt-team-insisted-not-to-use-the-word-caste-say-activists/

Figure 6: Sentiment polarity of the most frequent words within five themes

Gender

Devi, Survey, Training, Rate, Nrega, Work

Third Party

Budget, Finance, Government, Minister, Bengal

IEC

Campaign, Use, Terms, School, People

Quality

Company, India, Railway, Toilets, Use

Behavioural Change

Street, Sanitation, Toilets, Cities, Water

Coverage of **third-party involvement** within the project revolves around one single narrative: project funding and finance. Thus, most frequent terms within this theme are budget, finance, government, and other finance-related topics. One of the specific issues that has been covered in the media is, for instance, the financial support of the World Bank.[15,16] Some articles[17] make projections into the future specifically asking if the mission will meet its targets in 2015-2016 given the fact that the state government released only 49% of the overall sanctioned funding to the states. Yet some other articles talk about how institutions which provide essential services, such as airports, should be contributing to the financing of the programme.[18] The analysis indicates that media coverage of the third party involvement focuses heavily on the financial aspects. Other categories such as delivery and the quality of the construction, or the specific approach of Tamil Nadu to implementation of the project have not been widely covered.

The most frequent terms occurring within the **IEC** are terms such as campaign, use, school, terms, people, and others. This is the only theme where we do not observe clusters of words related to a single story or narrative. There are articles which discuss SBM activities within school settings. There are also publications which focus on the SBM activities in reaching out to different organizations for spreading and propagating information of the mission. One of them, for example, explores activities of religious organizations which volunteered to spread the word and tries to understand the main motivation of such an organization (which in this specific case was the organization's belief into how SBM is related to the ideal of Hinduism).[19]

When analyzing the category of **Quality**, we observe two predominant stories emerging. First of all, the prevailing majority of articles engaged with the quality of toilets and in particular the construction quality. Time-wise, these articles mostly concentrate on the past, meaning that media tends to focus on the quality of toilets and other facilities that have already been constructed. Some of the most recent publications focused especially on how the project meets or fails to meet its targets in rural areas.[20] The second and relatively recent narrative within the theme is the SBM evaluation of railway stations. Articles particularly cover surveying and ranking of the 400 busiest railway stations.[21]

Some of the most important topics covered by the articles on **behavioral change** are represented by such terms as "water", "sanitation", "toilets", "cities", and "drinking". The word "cities" refers mostly to the SBM rankings of cities that has been conducted in 2015, where the city of Thiruchirapalli (Trichy) of Tamil Nadu has been ranked on the second[22] among the 476 cities. Additionally, there are articles that preoccupy themselves with the topics on technology and how technologically advanced position of the Indian population – more Indians have mobile subscriptions than access to a toilet – can be used to deal with prevailing sanitation issues (such as open defecation fields).[23]

---

[15]http://www.thestatesman.com/news/business/world-bank-project-for-swachh-bharat-abhiyan-cleared/131859.html

[16]http://articles.economictimes.indiatimes.com/2014-10-27/news/55483085_1_12-crore-toilets-world-bank-finance-ministry

[17]http://www.thehindubusinessline.com/economy/swachh-bharat-fund-pipeline-to-states-clogged/article8429634.ece

[18]http://www.livemint.com/Companies/fsFFghgh2PQaPKsRLKhLEK/Fund-toilets-at-airports-via-fees-not-donations-suggests-p.html

[19]http://articles.economictimes.indiatimes.com/2014-10-02/news/54560353_1_organisations-swachh-bharat-patanjali-yogpeeth

[20]http://www.financialexpress.com/economy/swachh-bharat-abhiyan-urban-toilet-plan-76-pct-behind-schedule-gujarat-on-top/244195/

[21] For a more detailed reading on the ranking, please refer to the following article.

[22] http://www.economylead.com/government/swachh-bharat-ranking-mysuru-in-first-position-trichy-second-85225

[23] http://www.financialexpress.com/fe-columnist/editorial-apportunity/77943/

There are also pieces that take a more critical approach towards the mission's potential in triggering behavioral change. Specifically, the mission has been criticized for not bringing anything new to the problems related to the infrastructure. The mission is sometimes compared to the Total Sanitation Campaign[24] and it is argued that the government fails to see the actual problems related to the interdependence between infrastructure and behavioral change. Namely, some articles state that construction of toilets only does not lead to the needed change in behavior. That is, there is more to be done for behavioral change, especially in the areas of managing sewage systems and providing clean water.

---

[24] http://thewire.in/11318/only-a-change-in-government-behaviour-can-clean-up-india/

4. Annex

Table of National Publications

Number	National Publication
1	ABP Live
2	ABP Live (blog)
3	ABP News
4	afaqs
5	Ahmedabad Mirror
6	Amar Ujala Shimla
7	ANINEWS
8	AsiaNews.it
9	Assam Tribune
10	Autocar Professional
11	Avenue Mail
12	Bangalore Mirror
13	Bank Jobs in India (blog)
14	Best Media Info
16	Bollywood Helpline
17	Bollywood Hungama
18	Bollywood Life
19	Bollywood Spice
20	Bollywood.com
21	Boom Live
22	Bureaucracy Today
23	Business of Cinema
24	Business Standard
25	Business Today
26	Business Wire India
27	Businessinsider India
28	BW Businessworld
29	Calcutta Telegraph
30	Campaign India
31	CarTrade.com
32	Catch News
33	CauseBecause
34	Chandigarh Tribune
35	CineGoer.net
36	Citizen Matters Bangalore
37	Citizen Matters Bangalore (blog)
38	CounterCurrents.org
39	Cricket Country
40	Daijiworld.com
41	Daily Bhaskar
42	Daily Bhaskar (press release) (blog)

43	Daily Bhaskar (registration) (blog)	
44	Daily Excelsior	
45	Daily Mail	
46	Daily News & Analysis	
47	Daily Pioneer	
48	DailyO	
49	Dazeinfo	
50	Deccan Chronicle	
51	Deccan Herald	
52	DeshGujarat	
53	Diamond World Magazine	
54	E-Pao.net	
55	Eastern Mirror	
56	Economic Times	
57	Economic Times (blog)	
58	Economy Lead	
59	Equity Bulls	
60	ET Auto	
61	ET Retail	
62	ETAuto.com	
63	ETBrandEquity.com	
64	ETHealthWorld.com	
65	ETRealty.com	
66	ETRetail.com	
67	ETTelecom.com	
68	Exchange4Media	
69	exchange4media.com	
70	Express Computer	
71	FilmiBeat	
72	FilmiBeat Hindi	
73	Financial Express	
74	Firstpost	
75	Firstpost (satire)	
76	FitnHit.com	
77	Forbes India	
78	Free Press Journal	
79	Free Press Journal (blog)	
80	Frontline	
81	Governance Now	
82	Greatandhra.com	
83	Greentech Lead	
84	Gulf Times	
85	HERE. NOW	
86	Hindu Business Line	
87	Hindustan Times	

88	Huffington Post India
89	Human Rights Watch
90	I Government
91	IANS
92	Ibcworldnews
93	IBNLive
94	IBNLive (blog)
95	iDiva.com
96	India Education Diary
97	India PRwire
98	India Today
99	India Tribune
100	India West
101	India.com
102	indiablooms
103	Indiainfoline
104	Indian Cooperative
105	Indian Express
106	Indiantelevision.com
107	Indiatimes.com
108	indiatvnews on mobile (press release)
109	indiatvnews.com
110	indiatvnews.com (blog)
111	indiatvnews.com (satire) (press release) (registration)
112	Indileak
113	Inext Live
114	Infotech Lead (registration) (blog)
115	International Business Times India Edition
116	Jagran Josh
117	Jagran Post
118	Jansatta
119	KanglaOnline
120	Kashmir Images
121	Kaumudi Online
122	Khabar India
123	khaskhabar.com à¤¹à¤¿à¤à¥ à¤¦à¥
124	Knn India
125	koimoi
126	L'Indro
127	Livemint
128	Livemint - Business & Finance News India
129	Loksatta
130	m.indiatvnews.com (blog)
131	maharashtra times
132	Mainstream

133	Market Realist
134	MensXP.com
135	Merinews
136	Mid-Day
137	Mint Market Info
138	Mobiletor.com
139	Moneycontrol.com
140	Moneylife
141	Morung Express
142	Mumbai Mirror
143	mydigitalfc.com
144	Nagaland Post
145	Nagpur Today
146	Nai Dunia
148	NDTV
149	NDTV (blog)
150	NDTVSports.com
151	New Kerala
152	New Nation
153	New Security Beat (blog)
154	News Nation
155	News Today
156	NewsClick
157	Newslaundry
158	Newstrack India
159	Newsx
160	NYOOOZ
161	Odisha News Insight
162	Odisha Samaya
163	Odisha Sun Times
164	OdishaChannel
165	Oherald
166	Oneindia
167	Oneindia Hindi
168	Oneindia Tamil
169	onmanorama
170	Orissadiary.com
171	Outlook
172	Outlook Hindi
173	Outlook India
174	PaGaLGuY
175	Patrika
176	Pitch (blog)
177	PR Newswire
178	PR Newswire (press release)

179	PR Newswire India
180	Pradesh18 Hindi
181	PrameyaNews7
182	PrepSure
183	Pune Mirror
184	Punjab News Express
185	RadioandMusic.com
186	Rajasthan News
187	Rajasthan Patrika
188	Reuters
189	Sahara Samay
190	SahilOnline
191	Sakshi Post
192	Scoop News
193	ScoopWhoop
194	Scroll.in
195	Siliconindia.com
196	Skymet Weather
197	Sportskeeda
198	SpotboyE
199	State Times
200	Storypick
201	Swarajya
202	TechShout!
203	Techstory
204	Tehelka
205	Ten Sport (blog)
206	The Arunachal Times
207	The Asian Age
208	The Better India (blog)
209	The Echo of India
210	The Hans India
211	The Hindu
212	The Hitavada
213	The Indian Awaaz
214	The Indian Express
215	The Indian Express (blog)
216	The Navhind Times
217	The New Indian Express
218	The Northlines
219	The Quint
220	The Sangai Express
221	The Sentinel
222	The Siasat Daily
223	The Statesman

224	The Statesman
225	The Wire
226	TheHealthSite
227	Times Now.tv
228	Times of India
229	Times of India (blog)
230	Times of Oman
231	Trak.in (blog)
232	Travel Trends Today
233	TwoCircles.net
234	UdaipurTimes
235	ummid.com
236	Web India
237	Web India 123
238	Webdunia Hindi
239	YourStory.com
240	Youth Ki Awaaz
241	Zee News
242	Zee News (blog)
243	Zee News Regional
244	Zigwheels.com

Table of Regional Publications

Number	Regional Publication
1	Ananda Viketan
2	Daily Thanthi
3	Dinakaran
4	Dinamalar
5	Dina Mani
6	Junior Vikataan
7	Kalki
8	Maalai Malar
9	Malai Murasu
10	Tamil Murasu
11	Thinaboomi
12	Vikatan

5. References

Blei, David, Carin, Lawrence, & Dunson, David. 2010, "Probabilistic Topic Models." IEEE Signal Processing Magazine, vol. 27, no.6, :55–65.

Bravo-Marquez, Felipe, Mendoza, Marcelo, & Poblete, Barbara. 2013, "Combining Strengths, Emotions and Polarities for Boosting Twitter Sentiment Analysis." Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining - WISDOM '13, no.August,1–9. Retrieved (http://dl.acm.org/citation.cfm?id=2502069.2502071).

Gebrekirstos Gebreselassie. 2012, "Introduction to Sentiment Analysis Overview." Retrieved (http://www.lct-master.org/getfile.php?id=114&n=1&dt=TH&ft=pdf&type=TH).

Godbole, N., Srinivasaiah, M., & Skiena, S. 2007, "Large-Scale Sentiment Analysis for News and Blogs." Icwsm, 219–22. Retrieved (http://www.uvm.edu/~pdodds/files/papers/others/2007/godbole2007a.pdf\nhttp://student.bus.olemiss.edu/files/conlon/Others/__BookChapter_SocialMEsia_EBusiness/Large-Scale Sentiment Analysis for News and Blogs.pdf).

O'Keefe, T. & Koprinska, Irena. 2009, "Feature Selection and Weighting Methods in Sentiment Analysis." Australasian Document Computing Symposium, 67.

Singh Tomar Pankaj Sharma VITM, Deepak & AsstProfessor, Gwalior. 2016, "A Text Polarity Analysis Using Sentiwordnet Based an Algorithm." vol. 7, no.1, :190–93.

Thiel, Killian, Kötter, T., Berthold, Michael, Silipo, R., & Winters, Phil. 2012, "Creating Usable Customer Intelligence from Social Media Data: Network Analytics Meets Text Mining." KNIME. com Report, 1–18. Retrieved (http://www.knime.org/files/knime_social_media_white_paper.pdf).

Tumasjan, Andranik, Sprenger, To, Sandner, Pg, & Welpe, Im. 2010, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment." Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 178–85. Retrieved (http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1441/1852).

Will, Lowe, Benoit, Kenneth, Slava, Mikhaylov, & Laver, Michael. 2011, "Scaling Policy Preferences from Coded Political Texts." Legislative Studies Quarterly, vol. 36, no.1, :123–55.