

Impact Assessment of “Women, War and Peace”

Rezvaneh Rezapour, Jana Diesner

January 2015

Contents

1. [Executive Summary](#)
2. [Introduction](#)
3. [Background on Prior Work and Alternative Solutions](#)
4. [How our Approach works: Background and Methodology](#)
5. [Analysis of Media Coverage of Issue and Film](#)
 - a. [Data Collection and Curation](#)
 - b. [Semantic Networks](#)
 - c. [Digging Deeper: Summarization of Content of Text Data Sets](#)
 - [Afghanistan\(Peace Unveiled\)](#)
 - [Colombia\(The War We are Living\)](#)
 - [Liberia\(Pray The Devil Back To Hell\)](#)
 - [Serbia\(Came to Testify\)](#)
 - [War Redefined](#)
6. [Social Media Analysis](#)
 - a. [Twitter account \(account name: “Women, War and Peace”\)](#)
 - b. [Facebook Fan Page \(account name: “Women, War and Peace”\)](#)

1. Executive Summary

Impact assessment must begin with a research question. In the case of “Women, War and Peace” (A.Disney, P.Hogan, G.Reticker and N.Chaudry,). Our team was asked to investigate the following questions: How has the film moved the public awareness about Women and their role in the society as the key players in keeping the peace? And what does society think about “Women and War and the consequences of that in women’s lives?”

Using key terms provided by the filmmaker, we conducted an analysis of media and social media discourse on the topic of the film to determine its state before the film was released. This provided the baseline for any observable change. We then looked at the discourse on the topic after the release of the film. We define impact as any observable change in the media and social

media discourse between these two points: before and after the film's release. Within these parameters, we also looked for more specific indicators of change.

With regards to the first research question, if the film had an impact on media and social media discourse, we would expect to see several things. Because we must assume that the discourse of the topic is not static (meaning it is not completely dependent on the film), we can only determine that the film (and not something else) had an impact if the media discourse on the topic of the film ("**Women, War and Peace keeping**") intersects with the language or content of the transcript of the film in a way that it did not before the film's release. With regards to the second research question, we looked to see how the media discourse and social media discourse characterized the relationship between women and war and the consequences of that in different countries as mentioned in the keywords. To do this, we looked at the sentiments expressed, the key figures of the conversation, and other key indicators of observable difference in the structure and or content of the topic's networks before and after the release of the film.

It is important to say a brief word on data because all impact assessment is dependent on the type of data that is collected. A report is only as valid as the data is good. In other words, whether or not research questions can be asked or answered depends on the type of data the researcher (and filmmaker) collects and has access to. For the "WWP" film media analysis, we analyzed data that we collected from press coverage using the keywords for the film for a specific time period. WWP includes five different movies with different release dates, hence we needed to use different time periods for each of the movies, for instance for the movie "Pray the devil back to hell" which is the first movie of the WWP series, we had the following time period:

April 23, 2002 to April 23, 2008 [= Press Before] and April 24, 2008 to November 20, 2014 [= Press After]. The split point reflects the point in time when WWP was first screened at the 2008 Tribeca Film festival in NYC. These were found through LexisNexis Academic, one of the world's largest online electronic libraries for legal, business, news, and public information. For the "WWP" social media analysis, we analyzed data that we collected from Facebook and Twitter. In both cases, our team used **a)** network analysis to detect key agents and organizations and **b)** text mining techniques to find trends in current discussions (topics, sentiments, dynamics).

Lastly, it is important to note that the determination of a change in the network around a film and a film's topic may be different depending on the data source. For example, Twitter and Facebook are both social media databases, but the film has a greater engagement with social media users and a stronger network in terms of influential social media persons on Facebook as compared to Twitter. As we will discuss below, on Twitter, the film account is following and being followed by highly influential people. Out of the users that the account was able to attract, 275 users fall into categories of types of agents or organizations that are directly or indirectly relevant with respect to social change (Government/ legislation, NGO, Media, Celebrity and Others. **44** were affiliated with government/ legislation; there were **119** in NGO, **80** in Media, and **32** Celebrity with more than 100,000 followers,. On Facebook, the movie was able to engage users in a lively debate on the film's topic. Users not only picked up the topics of the posts, but also contributed

new, topic-related discussion points. Active users in terms of participation in commenting are mainly from the U.S. female.

With regards to interpretation of results, such as stated above and in much greater detail below, these differences or changes, which are observed in media or social media discourse, would be most meaningful to the filmmakers if contextualized within a larger plan for change, such as the action that the filmmakers would like to see Like role of Women in keeping Peace and Peace talks. In its current form, the impact assessment report tells the filmmaker that a change has or has not occurred in that media discourse, but it does not draw conclusions for the filmmaker as to their cause or value with regards to a particular action plan. To do that would require another set of analyses and data in addition to that of the media and social media. Such a project would represent an opportunity for future work.

2. Introduction

Documentaries are meant to tell a story, i.e. to create memories, imagination and sharing (Rose, 2012). Moreover, the goal with documentaries is to induce change in people's knowledge and/ or behavior (Barrett & Leddy, 2008). How can we know if a production has achieved these goals? We herein apply an empirical, scalable and systematic methodology that we have been developing for this purpose to data on the “**Women, War and Peace**” documentary (referred to as WWP herein).

In a nutshell, we approach this question by combining computational techniques from data mining and network analysis: we assume that documentaries are produced, screened and watched as part of larger and continuously changing ecosystems that involve multiple stakeholders and the flow of information between them. We track, map and analyze socio-semantic networks that represent these stakeholders and the information they disseminate (Diesner, Aleyasen, Kim, Mishra, & Soltani, 2013).

3. Background on Prior Work and Alternative Solutions

Prior work on assessing the impact of documentaries is limited in scope, depth and practical implementations (Barrett & Leddy, 2008; Figueroa, 2002). Major media institutes have proposed systematic frameworks, which are mainly of theoretical and/or normative nature (Barrett & Leddy, 2008; Clark & Abrash, 2011; Figueroa, 2002; KnightFoundation, 2011). Some frameworks include network related indicators, but fail to implement and measure them. Scholarly work on this topic is primarily confined to studies of psychological effects of films on individuals, and conceptualizes documentaries as a subcategory of mass media.

Overall, evaluation in this domain has typically been done by using (a) traditional, scalable and quantitative methods and metrics, such as the number of visitors of a screening or webpage, and/or (b) conventional, qualitative and small-scale methods for in-depth analysis of the perception of a topic or product by small numbers of people, such as interviews with focus groups. We integrate these two levels by jointly considering (a) the social network of stakeholders involved with the main topic of a movie - whether they have anything to do with a

particular production or not - and (b) the substance of the information produced and shared by these groups. The resulting socio-semantic networks allow for reasoning about two types of behavioral information - relationships and information (Diesner, 2013; Roth & Cointet, 2010).

4. How our Approach works: Background and Methodology

Our solution is based on a theoretical framework that we developed by synthesizing indicators of impact based on empirically tested theories from media effects, diffusion research, social and semantic network analysis, and collective action. The resulting CoMTI (content, medium, target, and impact) framework incorporates indicators specific to documentary evaluation that we identified in discussions with subject matter experts as well as additional impact metrics that we considered relevant (Diesner, Pak, Kim, Soltani, & Aleyasen, 2014). This framework considers a variety of stimuli that have been associated with cognitive, attitudinal, and behavioral change on the individual, communal and societal over time. In a nutshell, our methodology involves the following three steps (for details see (Diesner et al., 2013) :

Baseline model: First we map the current discourse about the main issue addressed in a movie prior to release. This is mainly to understand the existing ecosystem and where impact is possible. Main issues can be identified in a data driven way, e.g. by conducting text summarization techniques on the film transcript, or by film makers or funders. Herein we applied the latter strategy. Once the main issues are identified, we use a) network analysis to detect key agents and organizations and b) text mining techniques to find trends in current discussions (topics, sentiments, dynamics). For this step, we use ConText (<http://context.lis.illinois.edu/>) and NodeXL (<http://nodexl.codeplex.com/>) to collect, analyze and combine text data and network data based on news coverage data, social media data, and interviews with focus groups. Practitioners can use this procedure to understand the given opportunity space for connecting campaign work to relevant stakeholders and themes; helping them to strategically allocate scarce resource and mobilizing social capital.

Ground truth model: This represents the information contained in the actual documentary or media product, i.e. the message that the film *can* communicate. We understand that there is much more to a film than the actual content, e.g. the cast, images, sound and other aesthetic elements, which are not yet considered with our methodology. At the same time, we argue that the film's content is the smallest common denominator that anybody watching/ listening to a media product could take away. For this purpose, we apply the same text mining techniques as in step one, but this time to the film transcript.

Change model: Finally, we measure whether the needle has moved, i.e. trying to capture the measurable impact of a film. For this purpose, we reassess the public discourse and key players related to/ co-mentioned with from the film's release onward in time. The same types of data analysis techniques as in steps 1 and 2 are used. We analyze the overlap of:

- The ground truth model with the change model. This is information from the actual film that made it into public discourse, i.e. had an impact on public discourse.
- Any change in the topic (change model) and (social) media coverage of the film. This is

the impact of social media and press on the perception of the topic addressed in the film.

- Subtracting the ground-truth model and (social) media coverage from the change model let's use identify the amount of change in the topic of interest that can be attributed to any other things going and in the world. Because they do.

5. Analysis of Media Coverage of Issue and Film

a) Data Collection and Curation

To assess the media coverage of the core issue addressed in WWP and the movie itself, we collected and analyzed news wire data from LexisNexis Academic. To map the debate around the main issue addressed in the movie, we decided to divide the search into 4 parts. As the keywords that we received from the producer show WWP includes 5 movies which focus on four countries; Afghanistan, Serbia, Liberia and Colombia. We used these countries and the main problems addressed in each of the movies to search the press for the General theme of each movie. The keywords which were used are as below(Diesner & Rezapour, 2015):

Country:	Keywords: (general Theme of the movies)	Before	After
Serbia	woman-women-war-wartime-peace*-Serbia-rape-sexual violence	(10/10/2008)to (10/10/2011) 54	(10/11/2011)to (11/20/2014) 66
Afghanistan	woman-women-war-wartime-peace*- Afghanistan-peace talks-peace talks-peacetalk*	(10/24/2008)to (10/24/2011) 450	(10/24/2011) to (11/20/2014) 1069
Colombia	woman-women-war-wartime-peace*-Colombia- gold-mine-displace*	(10/31/2008)to (10/31/2011) 80	(11/1/2011)to (11/20/2014) 109
Liberia	woman-women-war-wartime-peace*-Liberia- protest*-Charles Taylor	(04/23/2002)to (04/23/2008) 493	(04/23/2008)to (11/20/2014) 605
Movie(Press on Film)	Keywords Used in LN	No. Docs from LN	
Pray the Devil Back to Hell	pray-back-hell-documentary-film-Abigail Disney -Pamela Hogan-Gini Reticker-Nina Chaudry	85	
I came to testify	I-Came-Testify-documentary-film-Abigail Disney-Pamela Hogan-Gini Reticker-Nina Chaudry-Bosnia	22	
Peace unveiled	peace-unveiled-documentary-film-Abigail Disney-Pamela Hogan-Gini Reticker-Nina Chaudry-Afghanistan	4	
The war we are living in	war-living -documentary-film-Abigail Disney-Pamela Hogan-Gini Reticker -Nina Chaudry-Women-war -Peace -Colombia	3	
War redefined	War Redefined-documentary –film-Abigail Disney-Pamela Hogan-Gini Reticker -Nina Chaudry	3	

The data cleaning, management and analysis of all LexisNexis data was then performed in

ConText¹. To the best of our knowledge, ConText is the only publicly available tool that supports the conversion of raw data from LexisNexis into a controlled and curated dataset and database per query. The database that gets automatically created and populated with data includes information about an article's release date, source and publication venue, language, author, images and graphics, length, and key words, among other categories.

In summary, we have four text data sets:

- Transcript of the movie. From this we construct the ground truth model.
- Press coverage on the main theme addressed in the movie before (baseline model) and after the screening of WWP at the Sundance Film Festival.
- Press coverage on the movie from the Sundance Film Festival onwards.

b) Semantic Networks

b. 1. Semantic Networks from Meta Data

The database we construct from the LexisNexis data entails keywords that were automatically assigned and indexed by LexisNexis. The key words represent the main, high-level individuals, organizations, locations and issues addressed in every article if applicable. A percentage value per keyword (also determined by LexisNexis) indicates the strength of association of a keyword with the article. Using ConText, we construct semantic networks from these keywords based on their co-occurrence per article.

Semantic networks are structured representations of information and knowledge that are assumed to represent the knowledge that some person or group has about a topic, and are typically used for reasoning and inference purposes (Diesner & Carley, 2011; Woods, 1975). At a minimum, semantic networks entail nodes, which are also referred to as concepts, and links or edges between the concepts. In this case, concepts represent key words that summarize or synthesize the information provided in news articles, and links are formed for any concepts that co-occur per article. The cumulative link weight represents the number of articles for which a link was observed based on the disambiguated articles. In a semantic network, the meaning of a concept is the ego-network of that concept, i.e. the nodes (alters) and links that get activated when the focal node (ego) is mentioned. The meaning of the entire network emerges from the collective meaning of the nodes as well as the structural properties or patterns of the entire graph. In

¹ When downloading data from LexisNexis, only batches of up to 500 articles including meta data can be stored at a time (for downloading larger search results, multiple batches need to be saved one by one). We built a module for ConText that pre-processes these downloaded batches by:

- splitting them up into individual text files,
- Deduplicating the articles, which is necessary because the same or highly similar articles might have been published by multiple outlets,
- storing each article as an individual text file, and
- Storing all data pertaining to each article in a database.

ConText, the user decides between which categories links shall be formed. For example, connections among and between agents and/or organizations represent a social network. The user also sets the strength value from which on key words are considered; the higher the value the smaller and more focused the resulting network. The networks images below were generated in Gephi (<https://gephi.org/>) based on the output files from ConText. In these pictures, the node colors indicate clusters (different flavors of meaning in this case) that naturally emerge from the data. Node size correlates with the degree centrality (number of direct neighbors) per node; the bigger the node the more it connects and relates to other themes.

b. 2. Semantic Networks from body of the articles

After exploring meta-data and keyword networks, we are digging deeper into the substance of the transcript and media data by analyzing the content of those documents. We use the most significant words in the documents and create a codebook which is the source of our next analysis. We create another network based on that codebook which shows the words in the body of the document.

c) **Digging Deeper: Summarization of Content of Text Data Sets:**

After exploring meta-data and keyword networks, we are digging deeper into the substance of the transcript and media data by analyzing the content of those documents. We use the following three summarization techniques to gain an efficient understanding of the gist of information contained in these different datasets: corpus statistics, topic modeling and sentiment analysis. Each of these techniques is explained in more detail below.

Corpus Statistics

One highly effective and efficient way to measure prominence or salience of terms explicitly mentioned in some text data set is to look at cumulative and weighted term frequencies. We list the most frequent terms and the terms that scored highest on the tf-idf metric per dataset after removing stop words before and after the release of the film to create a code book for digging deeper as mentioned before. (We used the stop word removal technique in ConText to eliminate those highly frequent yet not content bearing terms). In short, tf-idf scores terms high that occur often in a few documents. Therefore, this technique does not apply to the transcript, which contains only one document

Topic Modeling and Sentiment Analysis

Topic modeling is an unsupervised machine learning technique that summarizes the content of a corpus of unstructured, natural language text data in terms of the most salient topics that are explicitly or implicitly contained in the data (Blei, Ng, & Jordan, 2003; Griffiths, Steyvers, & Tenenbaum, 2007). Each topic is represented by a fit value that indicates how strongly a topic describes a text set, as well as the most salient terms per topic in the underlying data. The terms per topic are sorted by their fit with a topic.

Topic modeling is a parameterized method, i.e. the user has to set the number of topics to be identified (we used 4-7 depending on the size of the dataset), the number of terms per topic

(again, 5-10), the iteration rate for the routine (300-500), and a list with non-content bearing terms to be excluded from the analysis. We identified the best settings per data set by running topic modeling with different parameter configurations multiple times and comparing the results.

In the resulting images (Figures 7-10), each of the most salient topics is displayed as a cluster (typically 5 clusters per Image). Think of these clusters as different topics addressed in a dataset. The size of a cluster correlates with its prevalence within a single dataset (scaled according to the dirichlet parameter for fit value per topic). The same goes for individual terms within clusters: the larger a word, the more strongly it represents the topic it belongs to.

We have also added sentiment analysis into these visualizations: sentiment analysis basically helps to answer the following question: How does an individual or a group of people perceive a topic, product, etc. The underlying assumption with topic modeling is that human language provides windows into people's emotions and cognition. In the images below, neutral words are shown in blue, words with a negative connotation in red, and words with a positive connotation in green. Words that have no sentiment or opinion value are shown in black.

Since we have four different movies, we are going to analyze each of them separately below based on the countries they are related to:

- **Afghanistan:**

The debate around the main theme of the movie "Peace Unveiled" is centered on Women, Peace talks and Taliban. As "Figure 1" shows the main focus and general debate is around "Taliban", "War and conflicts" and "Terrorism" before the release of the movie. Women are in the margin. The fact that these nodes occupy central positions in the networks means that these themes are driving the public discourse on the given issue.

Change in baseline (from film release onwards): The thematic coverage from after the release onwards still has a strong focus on "Taliban" and "Peace talks" and "women"(Figure 2), women are not yet in the center but we can see a significant change in the network. This means that the public discourse connects these topics areas more strongly.

From digging deeper in the body of the document we can see a significant change between before and after the movie release. The connection between "Peace, talk, government, Taliban and women" are much stronger than before.

Press coverage of the film: About "Peace Unveiled" the press coverage on the movie itself is about the movie with "Taliban and War and Women" being the central nodes in this cluster. The green cluster is about the general theme of the movie and the purple one is focusing on the movies related to the WWP series. This means that the film is successful in tapping into the public discourse on the theme addressed in the film and connecting the issues discussed in the film to the central societal debates. We consider this successful media coverage. Note that the networks on the coverage of the theme are larger and denser than those for the press coverage of the film because the underlying datasets contain more articles.

Here are our findings: In agreement with the findings from meta-data analysis and digging deeper in the body of the articles, the coverage of the issue addressed in “Peace Unveiled” are as below:

Baseline and change in baseline: A focus on “Taliban, war and conflicts and government”. Before the release of the film and “peace talks and women, government and Taliban” after the release of the film.

Ground truth: The film’s content aligns with these themes, i.e. it picks up the public (discourse) on the main issues addressed in the film, which are related to women and their role in keeping the peace and being involved in the process of peace talks.

The press coverage of the film covers the issue using the same concepts that are already present in the public discourse on the film. In other words, the film’s content and press coverage on the film tap into the existing public discourse on the given issue; making it much easier for the film to connect to relevant stakeholders and adding a new dimension into the discourse.

In conclusion about the movie “Peace unveiled” we can see the shift from “Taliban and Government and war conflicts” to the more highlighted role of the women in the society in “peace talks and government and Taliban”

- **Colombia:**

The debate around the main theme of the movie “The War We Are Living” is centered on Afro-Colombian Women, War and their struggle over their gold-rich land. As “Figure 1” shows the main focus and general debate is around “Armed-forces”, “War and conflicts” and “Human rights violations” before the release of the movie. Women are not the center of the debates. The fact that these nodes occupy central positions in the networks means that these themes are driving the public discourse on the given issue.

Change in baseline (from film release onwards): The thematic coverage from after the release onwards has a strong focus on “War and Conflicts” and “Human Rights” and “women”(Figure 2), women are not yet in the center but we can see a significant change in the network. This means that the public discourse connects these topics areas more strongly.

From digging deeper in the body of the document we can see a connection between “Men”, “Women” and “War”. However we do not see any terms related to African Colombian women and their gold-rich lands yet.

Press coverage of the film “The War We Are Living”: the press coverage on the movie itself is about the movie with “Women, War and conflict, war crimes and women’s rights” being the central nodes in this cluster. The green cluster is about the general theme of the movie and the red one is focusing on the movies related to the WWP series. This means that the film is successful in tapping into the public discourse on the theme addressed in the film and connecting the issues discussed in the film to the central societal debates. We consider this successful media coverage. Note that the networks on the coverage of the theme are larger and denser than those

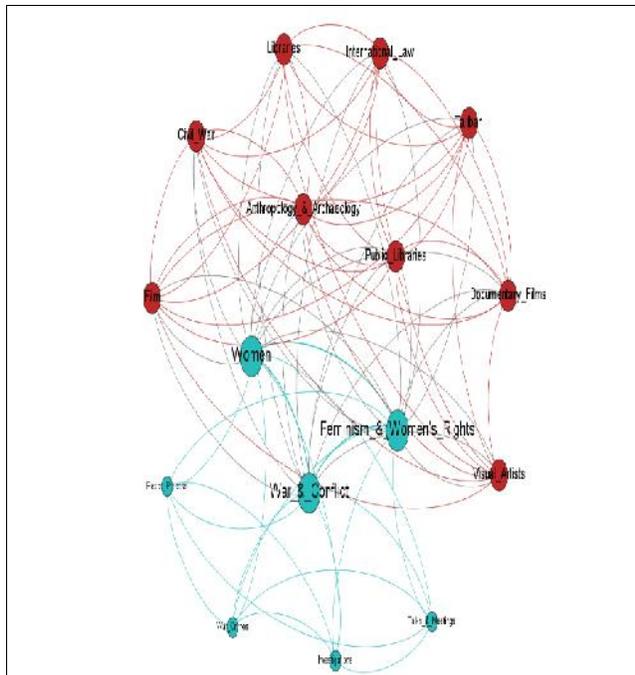


Figure 5: News Coverage of Film “The War We Are Living “: Semantic Network of Meta Data

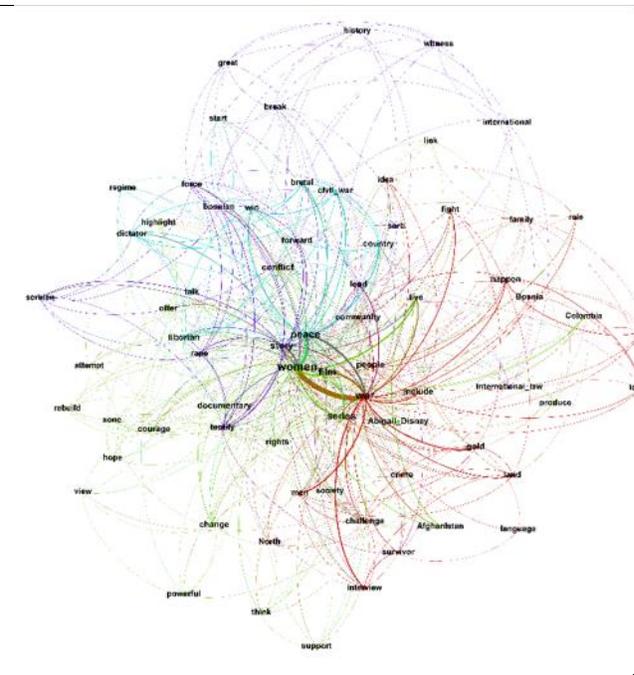


Figure 6: News Coverage of Film “The War We Are Living “: Semantic Network of the Body of the article

Topic Modeling: The press coverage of the issue before (Figure 7) shows “government, country and people” as the most important terms in the middle of the cluster. As we can see in = (Figure 10) after the release of the film “Work and war” are in the middle of the cluster “women, Peace and Colombia” are in one cluster. We can also see some terms like “Human Right, Peace, care” in the topic model after the release of the movie which is a significant change.

After running the topic modeling on the transcript paired with sentiment analysis (Figure 8) we can see that the movie addresses negatively loaded terms, e.g. “war” and “death”, which are also prominent in the press coverage of the movie or the debate on the topic.

Consistent with our prior results, the press coverage on the film “The War we are living “brings together the general theme of the “Women, War and peace” as well as “Abigail Disney” the producer of the series and a description on WWP as war series. (Figure 9).

As shown below “War”, “Death”, and “Attack” are negative terms and “Peace”, “Rights”, “Care” and “Support” are positive terms in all documents.

Conclusion: In agreement with the findings from meta-data analysis and digging deeper in the body of the articles, the coverage of the issue addressed in “The War we are living” are as below:

Baseline and change in baseline: A focus on “Armies, war and conflicts and government”. Before the release of the film and “Human rights and women, Women’s rights” after the release of the film.

Ground truth: it doesn’t seem that the film picks up the public (discourse) on the main issues addressed in the film, which are related to women and their struggle in keeping their gold-rich lands and in general the Afro-Colombian women and war in general.

In digging deeper on the press coverage of the film we found the issue using that film aimed to cover which were related to women, war and gold.

In conclusion about the movie “The War we are living” we can see the shift from “armed forces and war conflicts” to “women, peace and human rights”

- **Liberia:**

The debate around the main theme of the movie “Pray the Devil back to Hell” is centered on Women, civil war, Peace prize and Dictator Charles Taylor. As “Figure 1” shows the main focus and general debate is around “War and conflicts”, “Peace keeping” and “Civil War” before the release of the movie. Women are not yet in the center. The central nodes are driving the public discourse on the given issue.

Change in baseline (from film release onwards): The thematic coverage from after the release onwards still has a strong focus on “Women” and “War and conflicts” and “Nobel Prize”(Figure 2), women are in close and strong connection with the main themes in the network which is a significant change in the network. This means that the public discourse connects these topics areas more strongly.

From digging deeper in the body of the document we can see “Charles Taylor”, “war” and “crime” are connected before the release of the film. After the release of the film “women” are much highlighted in the middle of the network with close and strong connections with “Peace”, “Rape”, “democracy”, “violence” and “campaign”. We can also see a connection between “Charles Taylor”, “war”, “trail” and “crime” Which shows the public awareness about the main issue addressed in the film.

Press coverage of the film “Pray the Devil back to Hell”: the press coverage on the movie itself is about the movie as an art showed in the center of cluster in red. The green cluster is about the general theme of the movie and the Nobel Prize. The purple one is focusing on the movies related to the WWP series and different religions. This means that the film is successful in tapping into the public discourse on the theme addressed in the film and connecting the issues discussed in the film to the central societal debates. We consider this successful media coverage. Note that the networks on the coverage of the theme are larger and denser than those for the press coverage of the film because the underlying datasets contain more articles.

After analyzing the body of the articles we found that the general debate on the film “Pray the Devil back to Hell” is similar as above. “Women” are in the center of the cluster and in relation with “peace”, “Leymah Gbowee” the noble prize winner and “Abigail Disney” the producer of the film.

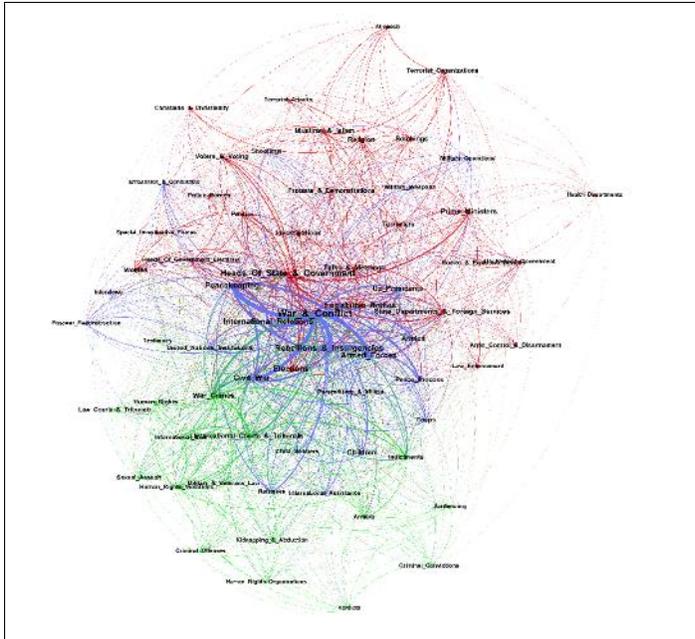


Figure 1: News Coverage on Topic before Film “Pray the Devil Back to Hell” Release: Semantic Network of Meta Data

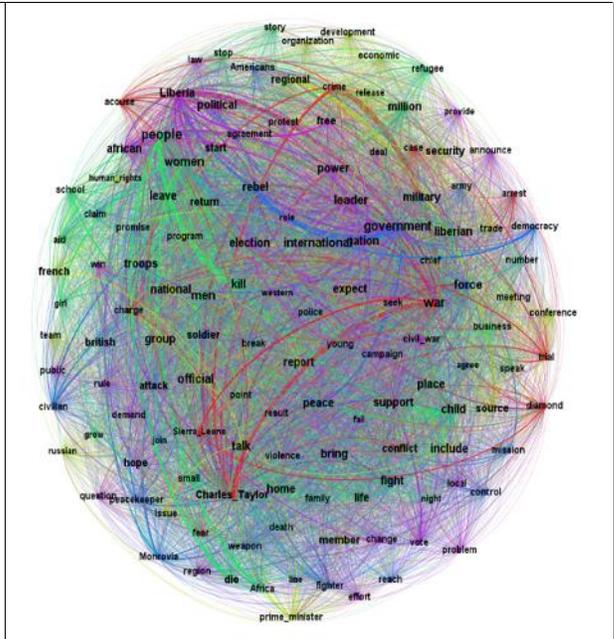


Figure 2: News Coverage on Topic before Film “Pray the Devil Back to Hell” Release: Semantic Network of the Body of the article

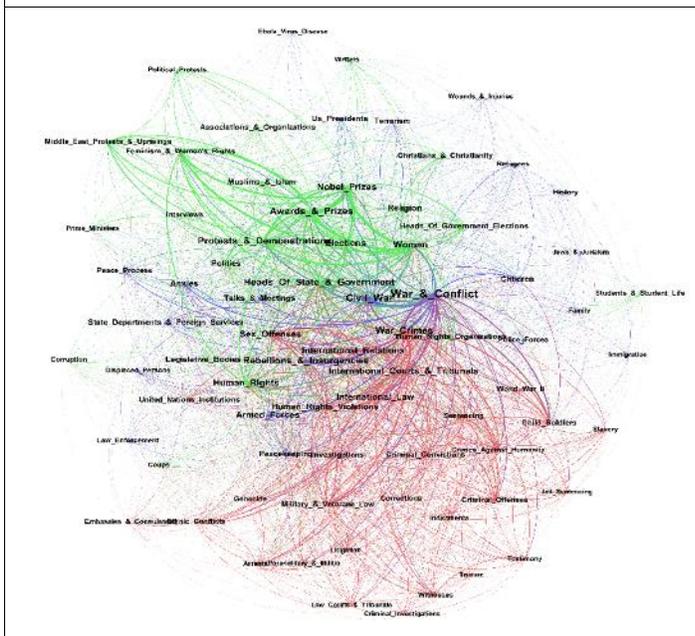


Figure 3: News Coverage on Topic After Film “Pray the Devil Back to Hell” Release: Semantic Network of Meta Data

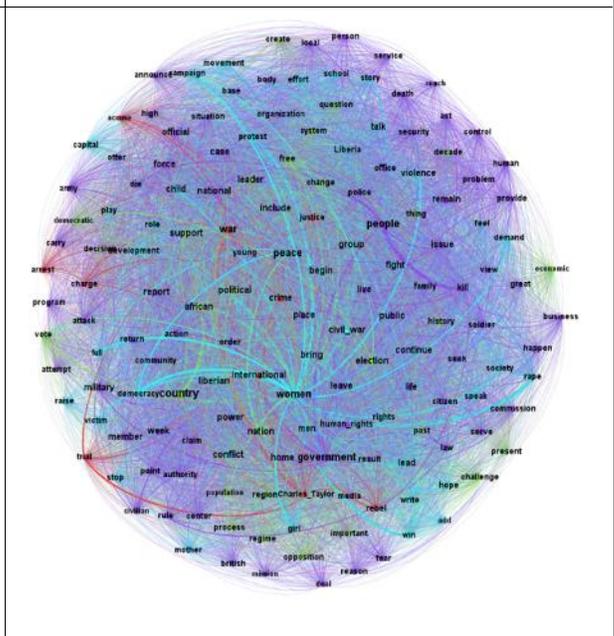


Figure 4: News Coverage on Topic After Film “Pray the Devil Back to Hell” Release: Semantic Network of the Body of the article



Figure 9: Summary of Press Coverage of Press on Film: Topic Modeling and Sentiment Analysis

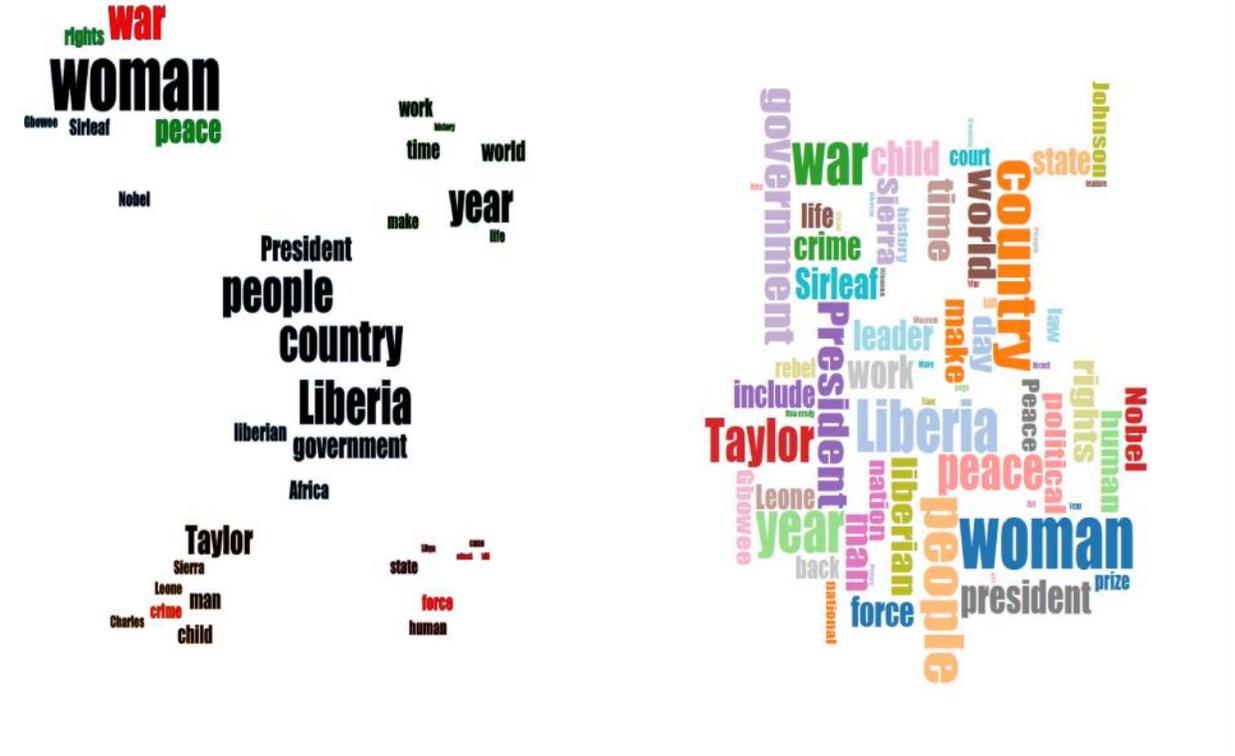


Figure 10: Summary of Press Coverage of Theme after Film Release: Topic Modeling and Sentiment Analysis

Conclusion: In agreement with the findings from meta-data analysis and digging deeper in the body of the articles, the coverage of the issues addressed in “Pray the Devil back to Hell “are as below:

Baseline and change in baseline: A focus on “War conflicts, civil war and Charles Taylor” before the release of the film and “Women , peace talks , Nobel Prize, Human Rights and Talks and meetings ” after the release of the film.

Ground truth: The film’s content aligns with these themes. It picks up the public (discourse) on the main issues addressed in the film, which are related to women and the Noble Prize and war crimes and the role of Women in Liberia and being involved in the process of peace talks.

The press coverage of the film covers the issue using the same concepts that are already present in the public discourse on the film.

In conclusion about the movie “Pray the Devil back to Hell” we assume this movie is the most successful one among the others since the public debates and the main issues address in the movie are very well aligned after the release of the film.

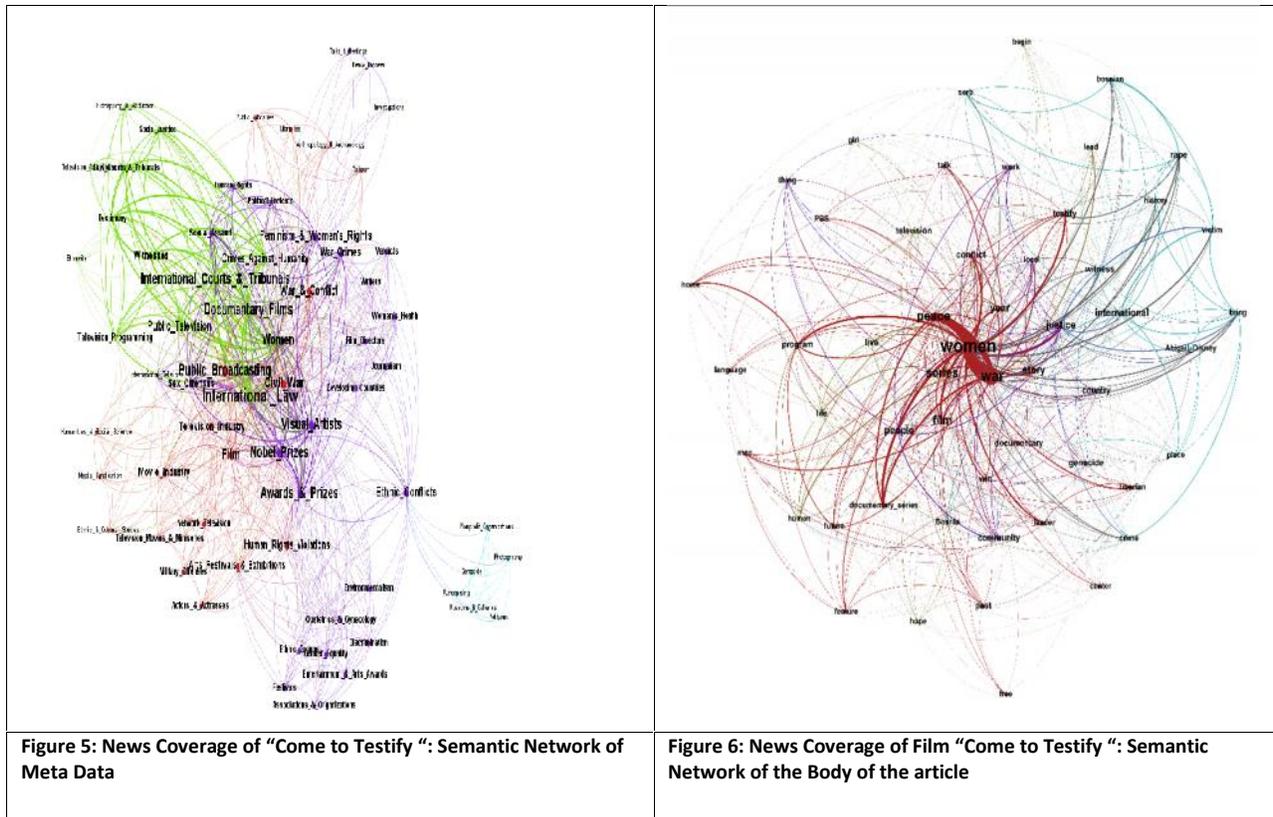
- **Serbia (Bosnia):**

The debate around the main theme of the movie “Came to Testify” is centered on Women, sexual violence during the war and their courage to go to international courts as witnesses. As “Figure 1” shows the main focus and general debate before the release of the movie are “Ethnic Conflicts”, “war and Conflicts”, “Genocide”, “Sex Offenses” and “War crimes” . Women are not in the center of the network yet but it is in the same cluster, Blue Cluster, with “sex offenses and human rights”. There are 4 clusters in the network with different focuses. The blue cluster as was mentioned before is more focusing on women and family and human rights, the green cluster is related to international law and judiciary and the purple cluster is talking about army and military forces and different religions.

Change in baseline (from film release onwards): The thematic coverage from after the release onwards focuses on “War and Conflicts”, “Genocide”, “war crimes” “International law” and “International courts and tribunals”(Figure 2),women are not yet in the center but we can see a connection between women and war crimes and . This means that the public discourse connects these topics areas more strongly since the focus of the film is also on Women as witnesses and war crime and the international courts.

From digging deeper in the body of the document we can see a significant change between before and after the movie release. Women are in the center of the graph and in related to “War, Rape, Justice and peace”

Press coverage of the film “Came to Testify”: the press coverage on the shows 3main clusters. The central focus is on Women international law sex offenses and documentary films. The green cluster is talking about the general theme of the film “Witnesses, testimony and social justice” are the highlighted subjects. The purple cluster is focusing on the other movies in WWP series, Noble prize and Art. This means that the film is successful in tapping into the public discourse



Topic Modeling: Women, War, men, Force and Rape are center topic in the press coverage of the movie before the release(Figure 7) .after the release of the film, we see human rights and international in the same cluster which may be related to international courts and public awareness of the war crimes in Bosnia(Figure 10).

To understand the main message that a viewer might get from watching a movie in a purely text data driven way, we also run topic modeling paired with sentiment analysis on the transcript. (Figure 8) indicates that the movie addresses negatively loaded terms, e.g. “war” and “rape” “entangled with Women and witnesses , which are also prominent in the press coverage of the movie or the debate on the topic.

Consistent with our prior results, the press coverage brings together the general theme of the “Women, war testify and justice” as well as description on WWP as documentary series. (Figure 9).

As shown below “War”, “rape” and “crime” are negative terms and “Peace”, “Rights” and “justice” are positive terms in all documents.



Figure 9: Summary of Press Coverage of Press on Film: Topic Modeling and Sentiment Analysis

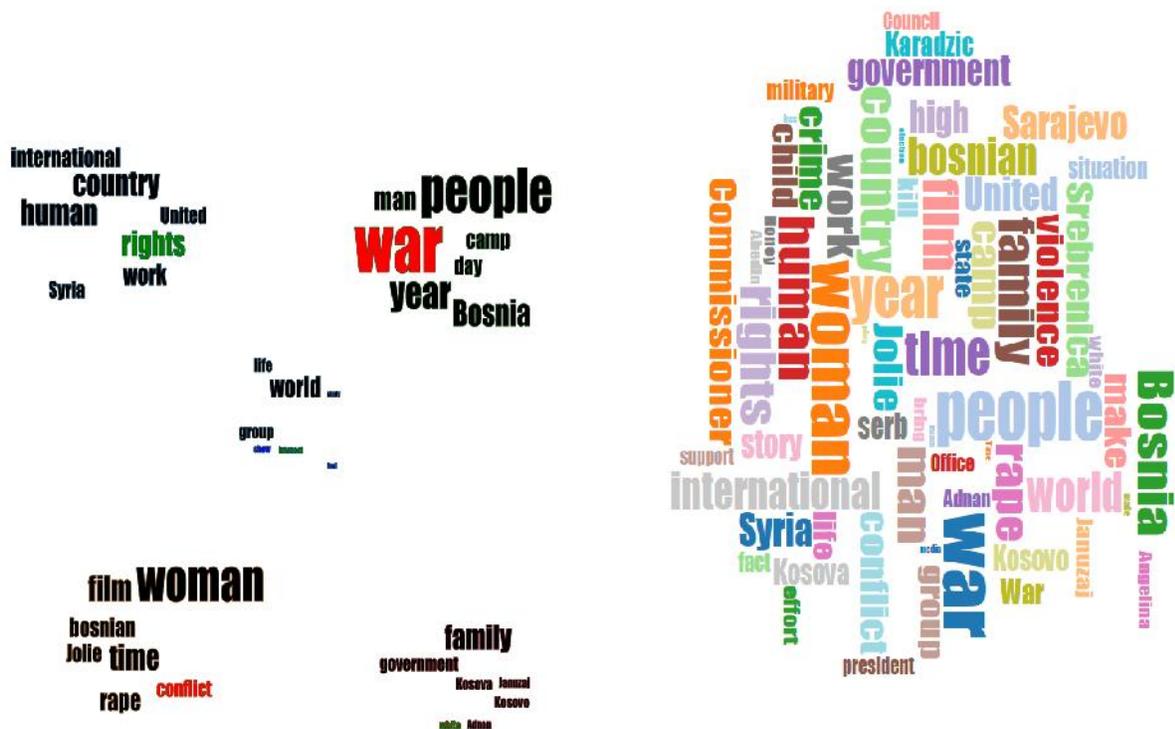


Figure 10: Summary of Press Coverage of Theme after Film Release: Topic Modeling and Sentiment Analysis

Conclusion: In agreement with the findings from meta-data analysis and digging deeper in the body of the articles, the coverage of the issue addressed in “Came to Testify” are as below:

Baseline and change in baseline: A focus on “genocide, war and conflicts and sexual offense”. Before the release of the film and “justice, women, testify, international court, and also human rights” after the release of the film.

Ground truth: The film’s content picks up the public (discourse) on the main issues addressed in the film, which are related to women and their courage in being present in international courts as witnesses and testifying on the war crimes.

The press coverage of the film covers the issue using the same concepts that are already present in the public discourse on the film. In other words, the film’s content and press coverage on the film tap into the existing public discourse on the given issue; making it much easier for the film to connect to relevant stakeholders and adding a new dimension into the discourse.

In conclusion about the movie “Came to Testify” we can see the shift from “War and conflict, genocide and Sex offenses” to “international law and court, justice and Human rights”

- **War Redefined:**

This movie is not related to any specific country. It includes the main issues and general themes covered above in all other four movies. Hence we just searched the press on the movie and analyzed the transcript.

In general the movie “War redefined” reveals how leading thinkers and politicians are looking at war to include women’s strategic role in global conflict, peacemaking and postwar reconciliation.

As mentioned before the role of Women have been more highlighted in past few years especially after the release of the movies analyzed above. The procedure of peace keeping and peace talk have also been more highlighted. The analysis above showed that women and peace talks have been entangled in last few years, hence we assume that the WWP series could perfectly highlight the strategic role of women in different societies.

As we can see in Figure 1 the press on movie “War redefined” focuses on the general theme of the series like “Sex offenses”, ”Awards” and ”Nobel prize”. In comparison to the result of the semantic network from the body of the article “Women, war and Peace” are centered in the network and subjects related to all the other films in the series can be seen as well as the “Abigail Disney” and “Leymah Gbowee” as the producer and Nobel Prize winner.

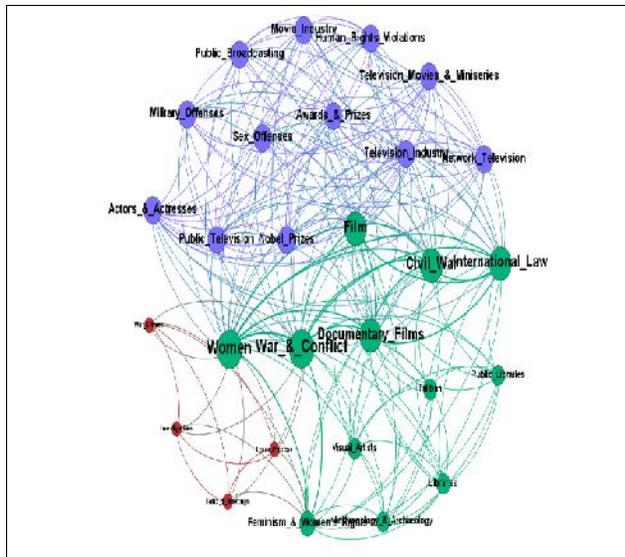


Figure 1: News Coverage of “War Redefined “: Semantic Network of Meta Data

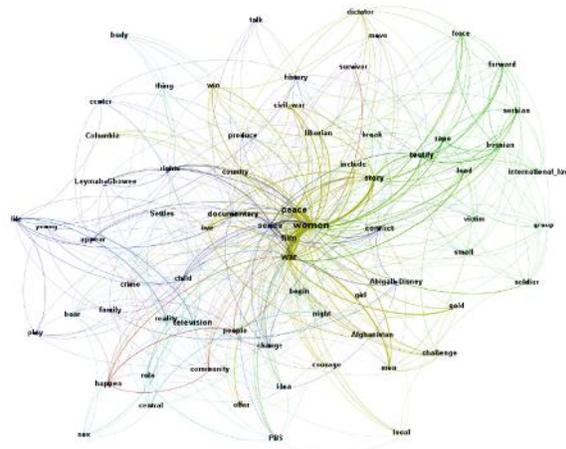


Figure 2: News Coverage of Film “War Redefined “: Semantic Network of the Body of the article

Topic modeling: after analyzing the transcript of the movie “War redefined” (Figure 3) we found that the movie addresses Women and the other general theme of the movie like war, conflict, rape and violence.

Consistent with our prior results, the press coverage brings together the general theme of the “Women, war testify and justice” as well as subjects related to other movies in the WWP series. (Figure 4).

As shown below “War”, “conflict” and “crime” are negative terms and “Peace” and “Security” are positive terms in all documents.

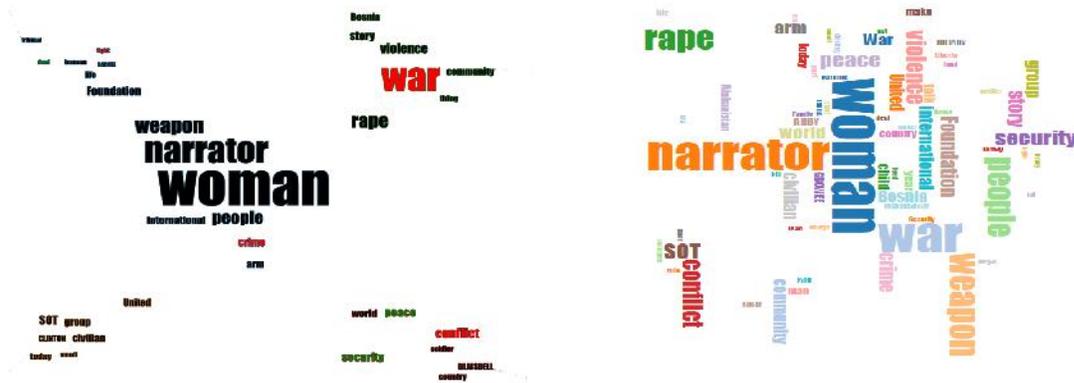


Figure 3: Summary of Transcript: Topic Modeling and Sentiment Analysis

6. Social Media Analysis

a. Twitter:

As of November 18, 2014:

Followers (Minus Intersection)	13,907
Followees (Minus Intersection)	1,679
Reciprocated followers (intersection)	707

Followers:

Of the **13,907** users following "Women, War and Peace" movie (but not followed by the movie) **55** of them are power users. (Power users are defined as users with more than 100,000 followers). On Twitter, power users gain their power from their followers, the number of followers represents the potential power of a user to broadcast information and affect others.

Intersection:

The overlap between followees and followers (reciprocated followers) is **707** users, with **42** of them being power users. This type of reciprocated relationships implies stronger social capital than both types of unidirectional relations (*Hansen, Shneiderman, & Smith, 2010*).

Followees:

The film account is following – but not followed by – 1,679 user accounts. Followees represent the group of users the film account is interested in and the possible pool of followers of the film account in the future. Among them, **101** accounts are identified as power users.

Follower's types:

Next, we took a closer look at the aforementioned followers. We divided all followers into five types of account:

1. Government/ legislation
2. NGO
3. Media
4. Celebrity

5. Others

We assume that the first four types: government/ legislation, NGO, Celebrity and Media to be directly or indirectly relevant with respect to social change. We choose the users with more than **1000** followers to reduce the dataset and have the most informative information. **275** users belong to three of these four types. **44** were affiliated with government/ legislation; there were **119** in NGO, **80** in Media, and **32** Celebrity.

The choosing criteria for these different types in shown in the table below:

Account Type	Criteria
Government/ legislation	Legal, legislation, Government
NGO	Non-Profit, charity
Media	Social media influencer, TV host, author, producer, movie, Editor, blogger, writer
Celebrity	Actor, Actress, director, singer, model

Out of the **1,975** users who has more than **1000** followers there was **1,700** in the “others” category where they did not fit in any of the above categories.

b. Facebook Fan Page:

We do not assume all social media platforms to lead to the same impact Thus, we complement the Twitter analysis with an analysis of user activity on Facebook. On Facebook Fan Pages, users can provide comments to posts. This provides a valuable source for analyzing stimulus (posts) and responses (comments to posts) Which together form a public discourse.

As of November 19, 2014, there are **250** posts on the Facebook of the film Most of which contain some image and text data. On this page there is **22,122** likes overall. A total of **3,932** users has contributed on the page. Out of that, **3,401** users have contributed in liking the posts. On these posts, there were **903** comments.

User Profiling:

We can profile users based on more detailed descriptions they have provided in their profiles:

1. Gender:

- 3315 female- 84%
- 430 male- 11%
- 187 not specified- 5%

2. Locale: top three

- 3466 are English US- 90%

- 213 are English British- 5.5%
- 30 are Spanish Laos- 0.7%

Posts and Comments Analysis:

Figure 1 shows a clustered semantic network generated from the posts on this page, where nodes are words in the posts, and links are formed if any two words co-occur at least twice within and/or across posts. The width of the ties is proportional to the frequency of co-occurrence. Each emerging cluster is indicated by a separate box and color.

These clusters represent the different themes that emerge from the discussion on this page. The most prevalent clusters are focused on the information presented in the movie. The topic of the movie is found but in peripheral clusters. We can compare this network to the one generated from the user's reactions to those posts. Figure 2 shows the semantic network constructed from the comments on posts, Using the similar rules (comment threshold: three or more times) and logic as above. Users' core discussions center on:

- 1) Standing against women and children violence
- 2) Sexual violence campaign

Overall, the discourse on the Facebook Fan Page suggest a lively debate in which users pick up on the provided stimuli (posts) by addressing additional themes in their comments. We conclude that, this page has focused on facilitating debate about the movie's topic, it has been successful in stimulating a public debate on the core topic of the movie and other related topics.

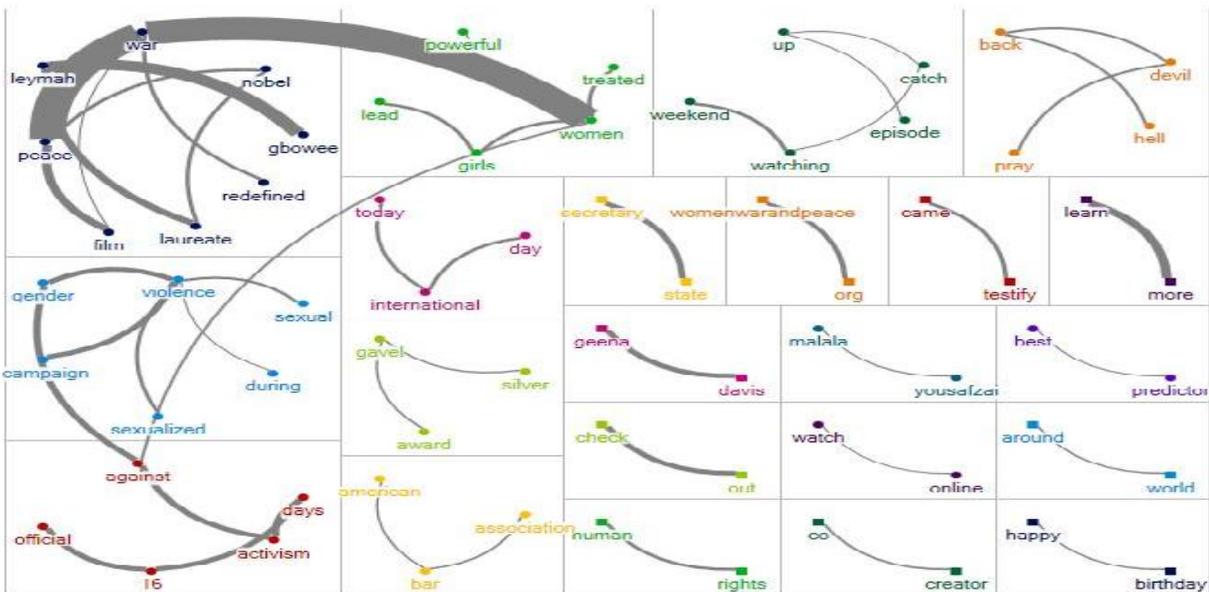


Figure 1: Semantic Map of Posts on the Facebook Fan page

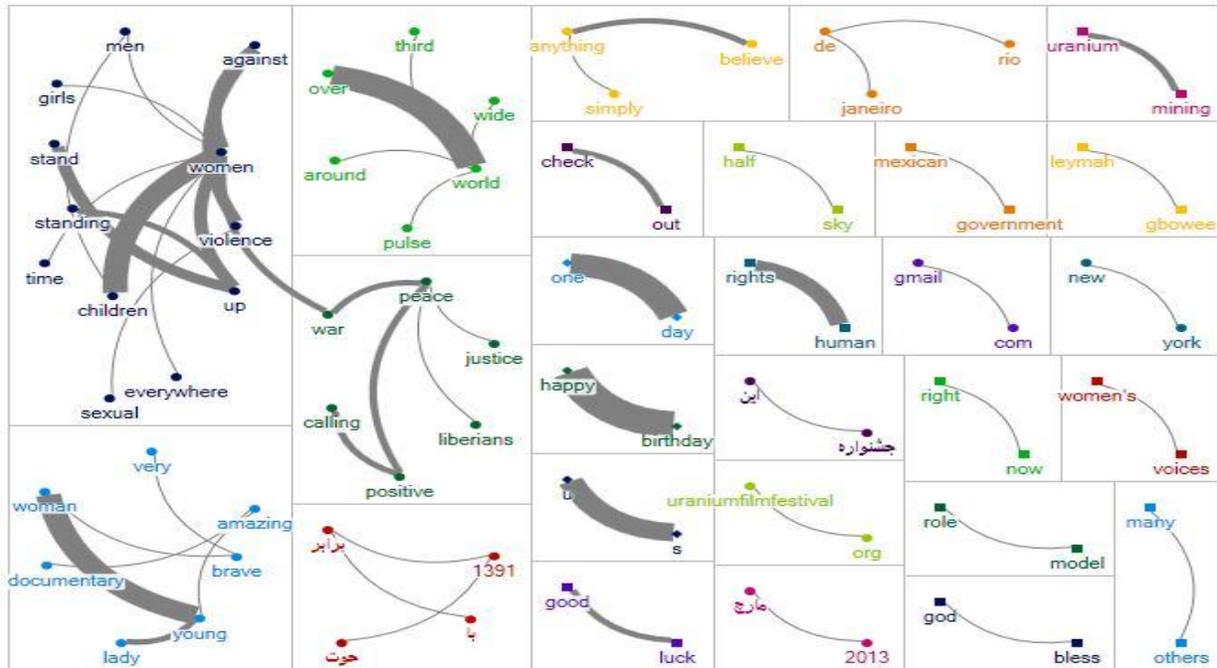


Figure 2: Semantic Map of Comments to Posts on the Facebook Fan page

c. Summary:

In summary, the social media analysis for the “Women, War & Peace” movie suggests the following:

- **Twitter:**

The account is successfully attracting influential people. **275** users were categorized as legal, NGO, Celebrity or Media. NGO category had **119** users where these are organizations that is related to the movie subject. One important follower of the account is the president of United States, which implies that the topic discussed in the movie is an interest of his. Beside him there are many influential users who follow the account these users have the power to spread the word about the movie are following the account.

- **Facebook:**

The fan page on Facebook also was a successful in engaging the users into a debate to discuss different topics presented in the movie. Active users in terms of participation in commenting came mostly from United States, and were female. After analyzing users, comments and post we so that users have engaged in a dissection about the movie and started a conversation about the topics presented. For the Facebook fan page account, strategic posting to facilitate debate on the movie’s topic seems to be needed. For example, controversial or unknown statistics relate to the topic can spark interests of users to participate in the movie’s message and disseminate it to their social media friends

Acknowledgment

This work is supported by the FORD Foundation's JustFilms Fund and Women, War and Peace group.

References

- Barrett, D., & Leddy, S. (2008). Assessing Creative Media's Social Impact. *The Fledgling Fund*.
- Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993-1022.
- Clark, J., & Abrash, B. (2011). Social justice documentary: Designing for impact: Center for Social Media.
- Diesner, J. (2013). From Texts to Networks: Detecting and Managing the Impact of Methodological Choices for Extracting Network Data from Text Data. *Künstliche Intelligenz/ Artificial Intelligence*, 27(1), 75-78. doi: 10.1007/s13218-012-0225-0
- Diesner, J., Aleyasen, A., Kim, J., Mishra, S., & Soltani, K. (2013). *Using Socio-Semantic Network Analysis for Assessing the Impact of Documentaries*. Paper presented at the WIN (Workshop on Information in Networks), New York, NY.
- Diesner, J., & Carley, K. M. (2011). Semantic Networks. In G. Barnett & J. G. Golson (Eds.), *Encyclopedia of Social Networking* (pp. 766-769): Sage.
- Diesner, J., Pak, S., Kim, J., Soltani, K., & Aleyasen, A. (2014). *Computational Assessment of the Impact of Social Justice Documentaries* Paper presented at the iConference, Berlin, Gemany.
- Figueroa, M. E. (2002). *Communication for social change: An integrated model for measuring the process and its outcomes*: Rockefeller Foundation.
- Griffiths, T., Steyvers, M., & Tenenbaum, J. (2007). Topics in semantic representation. *Psychological Review*, 114(2), 211-244.
- KnightFoundation. (2011). *Impact: A Guide to Evaluating Community Information Projects*.
- Rose, F. (2012). *The Art of Immersion: How the digital generation is remaking Hollywood, Madison Avenue, and the way we tell stories*: WW Norton.
- Roth, C., & Cointet, J. (2010). Social and semantic coevolution in knowledge networks. *Social Networks*, 32(1), 16-29.
- Woods, W. (1975). What's in a link: Foundations for semantic networks. In D. Bobrow & A. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science* (pp. 35-82). New York, NY: Academic Press.
- Diesner, J., & Rezapour, R. (2015). *Social Computing for Impact Assessment of Social Change Projects*. Paper presented at the Social Computing, Behavioral-Cultural Modeling, and Prediction (SBP), Washington, DC, USA.