Frontiers in Economics and Management Volume 4 Issue 9, 2023
ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

Analysis of Public Opinion Impact based on the Star Collapsed House Incident

Junyang Pan

School of School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233000, China

*2875177934@qq.com

Abstract

The ongoing information revolution and industrial transformation are evolving rapidly, and the development of the internet holds significant importance in fostering innovative capabilities, promoting new aspects of digital cooperation, and establishing a new framework for network security. However, with the rapid development of social networking platforms, negative events related to celebrities in the entertainment industry, such as property collapses, have become increasingly common in recent years. This article comprehensively examines the various factors and characteristics of information propagation following the outbreak of public opinion. It presents an information dissemination model for sudden public opinion events. Furthermore, it focuses on the case of Wang Leehom's property collapse incident and collects corresponding Weibo evaluation data during that period using web scraping techniques. By optimizing the word selection structure based on the Hidden Markov Model (HMM) principle, and applying LDA and K-means for overall sentiment analysis, the study identifies three main topics of interest to the public: discussions about the central figures in the public opinion event, debates on the legal and moral values of celebrities, and discussions on family concepts related to female celebrities.The article highlights the repeated appearance of keywords related to family issues and women's responsibilities in the distribution of evaluation topics. This suggests that the ethical and moral controversies and conflicts in family concepts reflected in the context of celebrity property collapses are critical issues that should not be overlooked and warrant attention.

Keywords

Star Collapse; Evaluation Analysis; Data Crawler; LDA.

1. Introduction

In today's world, the new era of internet information technology has transformed various aspects of human life. It plays a significant role in implementing the new development concept, serving as a vanguard of this concept, leading innovation, promoting coordinated development, facilitating green development, enhancing openness, and contributing to shared development. However, as the saying goes, "Water can carry a boat, but it can also capsize it." The rapid development of internet technology has also brought about many potential crisis effects for human society, especially among the younger generation. Their ability to discern information may be relatively weak, and under the influence of "fan culture" and the carefully curated public personas of celebrities, it is possible for them to develop a narrow and extreme behavior orientation, as well as skewed values.Therefore, safeguarding a healthy online ecosystem and ensuring a clean online space is of paramount importance, particularly with regard to addressing the potential negative impacts on young people.

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

With the development and maturation of big data technology, networking technology, and 5G technology, new media platforms now have access to rich user data, allowing for personalized content recommendations based on user preferences. This also makes it easier for users to find specific communities that share their particular interests and common topics, facilitating the emergence of fan communities. As internet technology advances, it is transitioning from being a mere information technology tool to a significant player in the realm of new media.The rapid growth and widespread use of the internet provide the public with a vast space for discourse and free expression.

This paper conducts an analysis of how people react and what they think in the face of public opinion shocks and how these shocks lead to changes in societal perceptions. It takes a multidimensional approach, considering factors such as societal moral values, ethical perspectives, and areas of public concern, to comprehensively examine the phenomenon of public opinion shocks stemming from celebrity property collapses. Through this comprehensive analysis, the paper aims to effectively assess the extent and depth of the impact of these public opinion shocks, aiding policymakers in making informed decisions.

2. Literature Review and Analysis of Mechanisms

2.1. Literature Review

The development of China's internet, in just over three decades, and platforms like Weibo, with a history of just over a decade, have played a crucial role in information dissemination. This article conducts a comprehensive literature review, focusing on the societal context reflected by celebrity property collapses and their effects on public opinion. Notable studies include Xinxin Guan's (2019) implementation of an improved TF-IDF algorithm using Python for media article analysis, Li Jinye's (2021) design of an enhanced TF-IDF algorithm for English text keyword extraction, and Thiyagarajan's (2021) discussion on the automation of keyword extraction techniques for discussion board assessment, highlighting the TF-IDF algorithm's strong correlation with human evaluation. Wei Wei emphasize the significance of negative online public sentiment as a powerful expression of public emotion in the internet era, necessitating timely management and guidance. Various authors have shared their perspectives on social network public sentiment and analysis methods, while Zhang Yiwu (2021) advocates a multifaceted approach involving regulatory bodies, industry associations, and platforms to regulate the online environment and promote rational celebrity following through education. Regarding the effectiveness of clean-up initiatives, Shi Nan (2022) and Li Xueqin (2021) highlight the opportunities brought by industry transformation. Lastly, Yang Zhiyong, Vice Dean of the Institute of Finance and Strategic Studies at the Chinese Academy of Social Sciences, underscores the importance of regulation in the internet industry's transition from "wild growth" to orderly development. Collectively, these studies shed light on the multifaceted impact of public opinion, particularly in the context of celebrity property collapses, emphasizing the need for effective management, regulatory measures, and education to foster a healthier online environment.

2.2. Public Opinion Mechanism Analysis

The dissemination of public opinion, in contrast to the traditional era of print media, is distinguished by the rapid pace in the digital age. Any individual with access to online communication devices can now bridge the gap between receiving and sharing viewpoints within a matter of minutes. Notably, platforms like Weibo, which emphasize keywords, have further compressed the process-from reception to contemplation to expression-into a matter of seconds. The communal and networked nature of online communication means that every user not only receives a plethora of fragmented information but also continually generates

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

fresh perspectives.Simultaneously, public opinion dissemination exhibits characteristics of clustering. The emergence of online communities, groups, and trending topics intensifies the compartmentalization of discussions. Furthermore, the decentralization and fission of public opinion dissemination enable users with differing views to expand upon existing discussions, thereby broadening the impact nodes of information and viewpoints. This transformation can turn what initially began as localized idea exchanges into novel points of opinion conflict, resulting in a dissemination process that accommodates both dispersion and aggregation.As celebrities inherently embody both traffic and trending topics, their every move carries a ripple effect in public opinion. For the ten celebrity property collapse incidents examined in this paper, the factors contributing to the resulting public opinion effects can be analyzed from four perspectives:

Firstly, the intrinsic impact of celebrity public opinion: Celebrities, being one of the most prominent groups in the spotlight, naturally serve as symbols of attention and discussion. The disparity between their actions and their carefully curated public personas significantly influences the subsequent public opinion.Secondly, psychological resonance among the masses: In the current cultural context, marketing and promotion are pivotal for amplifying influence. Social media provides a platform for individuals to express their viewpoints. Celebrities, as members of society, are equally subject to societal ethics and morals. Faced with events like property collapses, public groups with shared values rapidly gather to voice their opinions, often utilizing moral high ground as a catalyst for the dissemination of public sentiment.Thirdly, shortcomings in platform supervision: Given the anonymity associated with personal information on online platforms, achieving consensus and mutual understanding among users with diverse viewpoints can be challenging. Heated debates and contentious discussions often dominate the landscape. Anonymity diminishes ethical oversight, making it easier for new waves of dissemination to emerge.Fourthly, deficiencies in government and platform analytical capabilities and emergency response mechanisms: The overall trend typically follows a camel's hump shape rather than a sudden, sharp spike. Both government and platform accounts lack professionals who can effectively utilize their authoritative status for opinion guidance.

3. Analysis of Public Opinion Evaluation based on Machine Learning

3.1. Optimisation of Word Selection and Statistical Analysis of Review Data based on Hmm Principles

The article focuses on analyzing the public opinion evaluations related to ten celebrities' property collapse incidents, with particular emphasis on Wang Leehom. Several reasons underlie this selection:

Firstly, the data collection for the celebrities studied in the article was influenced by macro-level policies, leading to the censorship of certain keywords and comments. However, the evaluation information regarding Wang Leehom is the most recent and less affected by such censorship.Secondly, the Wang Leehom incident did not trigger any definitive legal judgments based on national laws, unlike other celebrities whose cases resulted in legal proceedings. The nature of this incident primarily revolves around the moral aspects of public opinion. Consequently, the comments generated by this incident are influenced to a greater extent by individual perspectives, making them more suitable for the audience evaluation analysis in the article.Lastly, the duration of the Wang Leehom incident aligns with the typical characteristics of public opinion outbreaks. It has a concentrated timeframe, making it easier to collect comment data. The overall duration of the incident was determined based on trends observed in Baidu Index keyword searches, as shown in the graph see Figure 1:

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

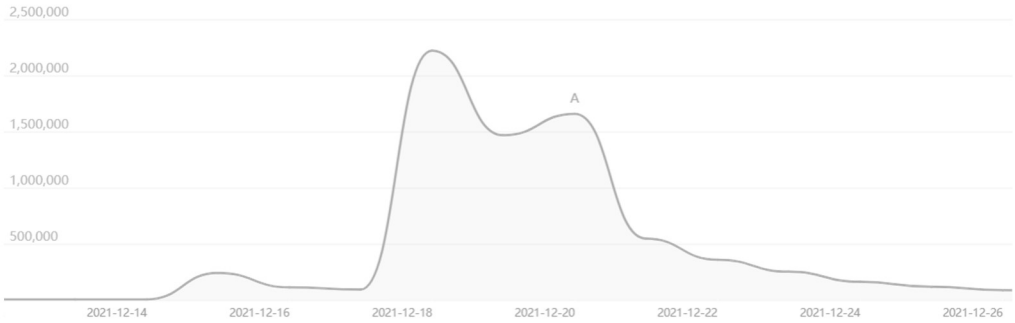

Figure 1. Baidu Index Trend of "Leehom Wang" Keyword Search

The trend chart reveals that the Wang Leehom incident had an overall duration from December 16, 2021, to December 26, 2021. To ensure data quality and focus on evaluations related to the incident itself, the article selected the period from December 17, 2021, to December 23, 2021, for the analysis of Weibo comments.During this timeframe, data was collected using web scraping tools, including information such as keywords, blogger IDs, blogger nicknames, blog post URLs, blog post content, publishing platforms, reposts, comments, and likes. In total, 9,236 pieces of relevant data were collected.Following this, the data underwent manual cleaning to remove irrelevant content, including videos, audio, advertisements, and some insufficiently informative comments. After the initial cleaning process, approximately 8,000 data points remained for further analysis.

The Hidden Markov Model (HMM) is essentially a statistical Markov model of a system that uses a hidden Markov chain to predict the next state. In this system, there is a Markov process combined with unobserved (hidden) states. The application of the HMM model in Chinese word segmentation and tagging primarily involves using observed words to infer the most probable latent states. By backtracking, the model can determine the segmentation components.Building upon the principles of this method, the article used a Python program to perform Chinese word segmentation on the collected 8,000 comments. The segmentation results and their corresponding word frequencies are presented in the table see Table 1:

Table 1. Preliminary disaggregated processing table for comment data

Numble	Tag words	Frequency	Category
11377	Wang Lee-Hom	10844	noun
11378	Li Lianglei	2966	noun
11379	microblog	2254	noun
11380	cucumbers	1416	noun
11381	full version	982	noun
...	...	...	...

All lexical quantities were also summarised and the results are shown in Table 2:

Table 2. word index

Lexical category	Pronoun	Other	Adverb	Prepositions	Nouns	Conjunctions
Number	288	11377	298	4843	11291	1041

While all the presented keywords clearly show the appearance of terms related to the event and individuals involved, it is inevitable that pronouns, verbs, meaningless nouns, and other

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

phrases may also appear. Therefore, considering the characteristics of the event, a filtering process was applied to the segmentation results to remove verbs, pronouns, adverbs, adjectives, as well as some meaningless and repetitive nouns.As a result, the original comments underwent cleansing and labeling to eliminate redundant vocabulary. The refined results are shown in Table 3:

Table 3. Filtered table of original comment splits

Serial number	Original comment	Tagged word
242374	Chow Tai Sang and Wang Leehom's collaboration expires without further cooperation	Chow Tai Sang,Wang Leehom
242376	Originally not rely on Wang Leehom, that is the strength of others Wahaha	Wang Leehom
242379	This time it's actually Leehom Wang. Seriously, I kinda like listening to him.	Wang Leehom
...	...	...

From the categorization of word types, it is evident that the keywords in the comments primarily revolve around Wang Leehom and Li Jiong's family-related topics, marital views, connections with other celebrities, moral perspectives, and other issues. These hot-topic keywords are mostly centered around related celebrities, friends, and family-related discussions. However, the words themselves may not provide substantial information.In the subsequent analysis, clustering will be performed on the selected keywords to obtain a distribution of the main evaluations. This process aims to group similar keywords and extract meaningful patterns or sentiments from the comments for further examination and understanding.

3.2. LDA-based Subject Classification and Individual Subject K-means Centres

In this paper, based on the LDA theme model training process, the use of python to visualise and analyse the theme acquisition, the use of data for the total number of screened sub-vocabularies, to reduce the interference of meaningless vocabulary, and at the same time set up on the data to remove the frequency of the words higher than 0.6 and lower than 0.2, so that the classification results are not affected by outliers. Considering that the sample number magnitude is small, and the distribution of the theme is more concentrated in the star, so the number of selected themes is 3, then the PyLDAvis visualisation results are divided into a total of 3 images, the specific classification results are shown in Figure 2-4:

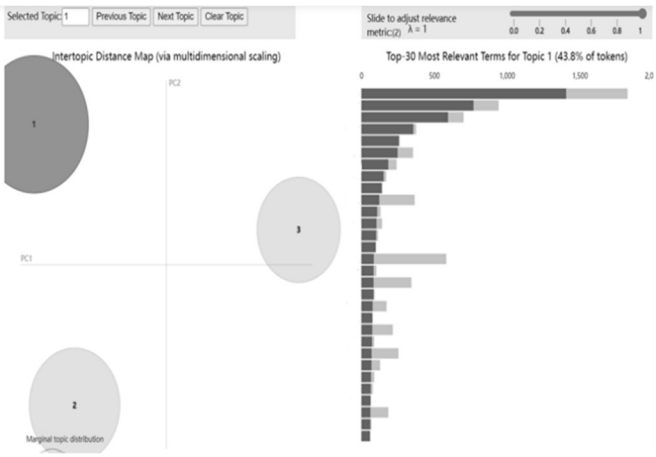

Figure 2. Theme 1 Visualisation Relationship Map

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

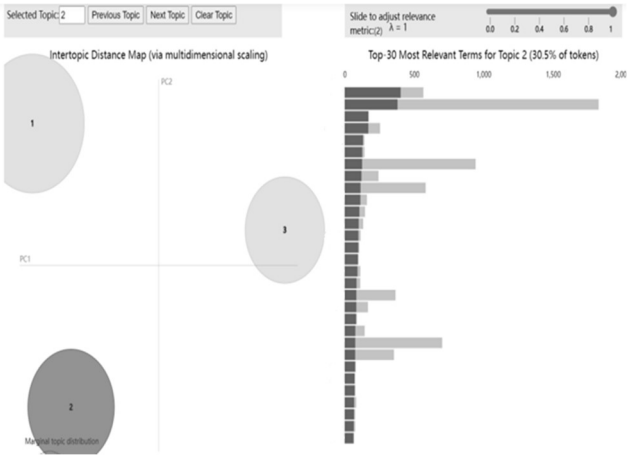

Figure 3. Theme 2 Visualisation Relationship Map

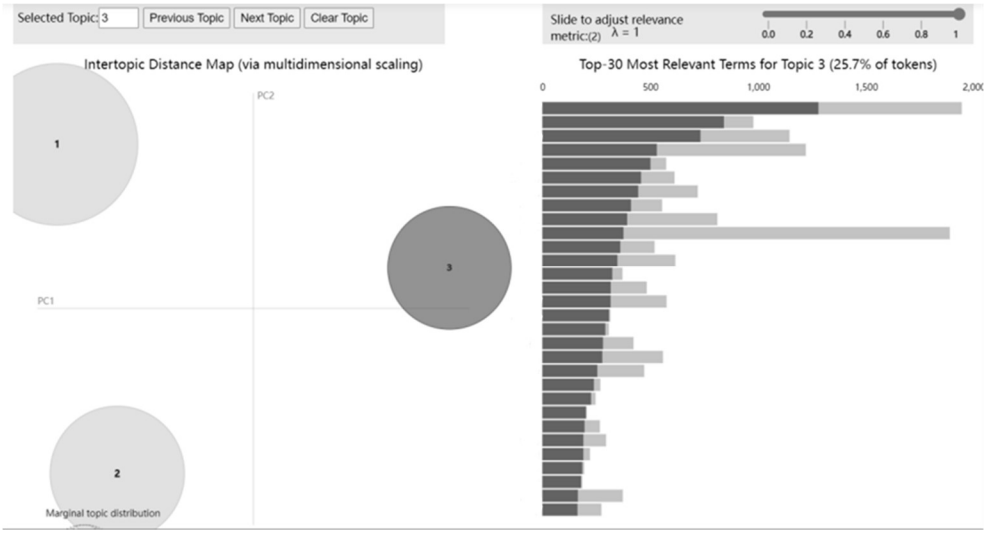

Figure 4. Theme 3 Visualisation Relationship Map

The results of the keyword distribution relationship when the number of clusters is 3 are as follows:

1st theme: 0.185*"Li Lianglei" + 0.034*"ex-wife" + 0.033*"star" + 0.028*"music" + 0.024*"children" + 0.022*"marriage" + 0.018*"China" + 0.017*"Via" + 0.017*"fiasco" + 0.017*"webpage"

2nd theme: 0.069*"netizens" + 0.064*"Li Lianglei" + 0.051*"Weiya" + 0.037*"Problems" + 0.024*"JJ Lin" + 0.024*"Marriage" + 0.024*"Yundi Li" + 0.023*"News" + 0.022*"Lightness of Being" + 0.020*"World"

3rd theme: 0.157*"Weibo" + 0.121*"Video" + 0.051*"Time" + 0.046*"Women" + 0.035*"Li Lianglei" + 0.027*"Fan Weiqi" + 0.026*"Family" + 0.019*"Children" + 0.018*"Marriage" + 0.016*"Xu Ruoxuan"

Taking into account the imagery and distribution relationship functions, the following observations can be made:

For Topic One, the primary aggregated keywords are "Li Jinglei", "marriage", and "celebrities". This indicates that Topic One is centered on the main characters of the public opinion event and represents the main focus of discussion among netizens. Therefore, Topic One can be identified as a discussion of the central figures in the public opinion event.

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

Regarding Topic Two, the main keywords include "Viya", "JJ Lin", "Li Yundi", and "issues". This suggests that Topic Two revolves around discussions of celebrity values and the intense debates over moral and legal issues related to celebrities. Hence, Topic Two can be categorized as a discussion of the legal and moral values associated with celebrities.

For Topic Three, the prominent keywords consist of "women", "Fan Weiqi", "family", and "children". This indicates that Topic Three predominantly focuses on traditional family values and ethical considerations from the perspective of female celebrities. Therefore, Topic Three can be characterized as a discussion of family values within the context of female celebrity perspectives.

Upon identifying the aforementioned thematic directions, to further explore the information distribution of the relevant keywords within each topic and gain a more detailed understanding of the themes of public discourse and the topics of interest, taking into consideration the interconnectedness of keywords, a k-means clustering analysis was employed using the aforementioned keywords as the primary dataset.

This study utilizes Python programming to conduct cluster analysis on the filtered tokenized data. It also extracts relevant keywords for the LDA topic model and calculates the proximity of these keywords to the central words, obtaining centrality values. Based on the distribution relationship between topics and subcategories, the results were dissected into three categories.

In the first theme of discussions regarding the central figures in the public opinion event, keywords related to "Li Jinglei" are associated with terms such as "infidelity," "accusations," "statements," and "former wife." Keywords related to "marriage" are linked to terms like "intrusion" and "intervention." These terms are commonly found in internet discussions, indicating a shared focus among the public on issues related to marriage and infidelity in the context of this central figure.

In the second theme concerning discussions about the family concepts of female celebrities, keywords related to "Xu Ruoxuan" and "Fan Weiqi" are associated with terms like "speaking out," "lawsuit," and "condemnation." In this theme, the central terms reflect that the public is more concerned with the attitudes of female celebrities, who are part of the celebrity group, towards family and emotional issues. As a relatively vulnerable group in the moral discourse, public attention is particularly drawn to their perspectives on marriage and emotions.

In the third theme, which focuses on discussions about the legal and moral values of celebrities, keywords related to "Wei Ya" and "career" are associated with terms like "tax evasion," "taxation," "family," and "men." This suggests that the public's interest in the Wang Leehom incident has also extended to other celebrity property collapse cases, with a particular emphasis on discussions about gender roles within family relationships.

4. Conclusion

Looking at the overall keyword processing process in the context of the Wang Leehom incident, it is evident that the public's primary mode of evaluation is based on a moral values perspective. Additionally, related events associated with this incident are often brought up in public discussions. This includes discussions involving friends from the same social circle as the involved parties and contradictions and phenomena related to real-life family weddings. From a temporal perspective, negative emotional evaluations consistently dominate the discourse, with the main focus of negative evaluations centering around emotional issues related to marriage. Notably, there is no observed decrease in negative emotions over time. Furthermore, as the incident continues to develop or when third parties, such as official media outlets, define the incident's consequences or when brands terminate contracts, new rounds of public opinion dissemination are triggered. The prevalence of negative evaluations throughout the timeline suggests that the emotional impact of celebrity property collapses does not spontaneously

ISSN: 2692-7608 DOI: 10.6981/FEM.202309_4(9).0015

dissipate. Monitoring and addressing this phenomenon is essential to prevent conflicts and misinformation. From the perspective of thematic scope, overall evaluations revolve around factors primarily related to the perspective of victims. This includes discourse from the female perspective and discussions about the role of women within the context of family values. These topics are the ones that the public is most concerned about on the Weibo platform. These discussions often give rise to contradictions in family values, which should be a point of concern and regulation for the government. The phenomenon of celebrities may serve as a focal point and magnifying glass for such issues, but the underlying problems are deeply rooted in the challenges of the real world. Examining the distribution characteristics of the audience of interest using decision tree models, it becomes evident that work environments and family factors are closely related to public behavioral perceptions. Behind the attention to the development of public opinion, it is important to explore these societal reflections further.

The correct understanding of positive values is crucial for the artistic performance of celebrity artists. As the cultural trends of our times continue to evolve, the benefits derived from being a celebrity are increasingly favored by capital and the market. However, the rapid development of the entertainment industry has led to insufficient industry regulations. In fact, there remains a lack of comprehensive training for artists in terms of their personal qualities. It is not difficult to observe that some celebrities exhibit lapses in their moral character, value consciousness, and personal behavior. The reason why celebrities may cross ethical boundaries is often due to deviations in their personal values. Emphasizing the cultivation of positive energy at the source of celebrity artistry is the primary task for establishing their social trustworthiness. It involves nurturing their foundational moral character, guiding them in distinguishing between right and wrong, and helping them develop a sense of responsibility as celebrity artists. This, in turn, should encourage them to discourage actions that violate public order and good morals. Only by actively promoting positive values to society can celebrities maintain the correct direction of their value system.

References

[1] Chong, Dennis, and James N. Druckman. Dynamic public opinion: Communication effects over time. American Political Science Review 104.4 (2010): 663-680.

[2] Hoffman, Lindsay H., et al. The role of communication in public opinion processes: Understanding the impacts of intrapersonal, media, and social filters. International Journal of Public Opinion Research 19.3 (2007): 287-312.

[3] Masngut, Nasaai, and Emma Mohamad. Association between public opinion and Malaysian government communication strategies about the covid-19 crisis: content analysis of image repair strategies in social media. Journal of medical Internet research 23.8 (2021): e28074.

[4] Boutyline, Andrei. Holding a position: Public opinion as cognition in a disorganized field. Poetics 95 (2022): 101710.

[5] Cacciatore, Michael A. Arthur M. Sackler Colloquium Advancing the Science and Practice of Science Communication: Misinformation about Science in the Public Sphere: Misinformation and public opinion of science and health: Approaches, findings, and future directions. Proceedings of the National Academy of Sciences of the United States of America 118.15 (2021).

[6] Dong, Xuefan, and Ying Lian. A review of social media-based public opinion analyses: Challenges and recommendations. Technology in Society 67 (2021): 101724.

[7] Jia, Lianrui. What public and whose opinion? A study of Chinese online public opinion analysis. Communication and the Public 4.1 (2019): 21-34.

[8] Younus, Arjumand, et al. What do the average twitterers say: A twitter model for public opinion analysis in the face of major political events. 2011 International Conference on Advances in Social Networks Analysis and Mining. IEEE, 2011.