



Research on factors influencing opinion polarization based on QCA algorithm

Jianhua Dai, Shaoqiang Shen*

School of Economics and management
Nanjing university of science and technology Nanjing. China

djh705@163.com, 12110722856@njjust.edu.cn*

Abstract. The phenomenon of opinion polarization in unexpected online public events constantly challenges the construction of clear cyberspace, which can also lead to social harms such as cyber violence, rumors, moral abduction and doxing, etc. Studying the influencing factors of opinion polarization can help clarify the causes of the opinion polarization and dissipate the negative effects as well. Taking 29 hit public events as cases, clear-set QCA algorithm is adopted to analyze the synergy effect among event, emotion and communication variables of opinion polarization, and the findings are as follow: (1) Single variable cannot lead to opinion polarization and opinion polarization is the result of synergy of variables.(2) Event attribute and Emotional attribute play the main role of influencing opinion polarization while the communication attribute plays a subordinate role. (3) Type of event, event trigger, degree of harm, emotional state and information ambiguity are the key influencing factors (4) The three main types of configuration paths for opinion polarization are: the type which harms the national sentiment, the government's breach of trust type and the celebrities' breach of morality type.

Keywords: Online public opinion, opinion polarization, synergy analysis, csQCA

1 Introduction

The anonymity of the Internet and the borderless and instantaneous interaction of the "network society" have facilitated the dissemination and sharing of social information, but the phenomenon of polarization of public opinion triggered by Internet public opinion has also become an important factor affecting social stability and has received widespread attention. In public opinion events, since most of them are socially sensitive phenomena and the subjects of the events are mostly government personnels or public figures such as celebrities, they are more likely to cause a greater degree of public opinion reactions on the Internet, which brings challenges to social governance and the construction of a clear cyber environment.

Research on factors influencing opinion polarization has become a hit topic through the years, existing studies mostly focus on the event and communication attribute of the

© The Author(s) 2023

X. Ding et al. (eds.), *Proceedings of the 2023 4th International Conference on Big Data and Social Sciences (ICBDSS 2023)*, Atlantis Highlights in Social Sciences, Education and Humanities 12, https://doi.org/10.2991/978-94-6463-276-7_10

events, but there are few studies that include users' sentiment attribute, and most of the existing studies ignore the effect of linkage and matching effects among many factors on opinion polarization. Based on this, this study collects hit public opinion events in the past four years and uses Crisp set Qualitative Comparative Analysis (csQCA) to study the influence of three factors, namely, event attributes, sentimental attributes and communication attributes, on the polarization of public opinion. The method seeks to analyze the synergy effect of the three attributes on the polarization of public opinion and identify the main driving paths leading to the polarization of public opinion.

2 Literature review and Configuration Model

2.1 Opinion Polarization

James Stoner studied the interesting phenomenon of group members' views becoming more extreme during group discussions. He introduced the concept of "group polarization," which emphasizes the cognitive shift toward one side of an opinion after a discussion among team members, resulting in an extreme opinion^[1]. Public opinion polarization" is derived from "group polarization", but it is more inclined to the differences in the perceptions of people about a certain fact or entity, and such deviations are unipolar, bipolar or multipolar, and their views are extreme and irrationalized^[2].

Existing studies have also looked at the influencing factors lead to opinion polarization. For example, some studies have suggested that in social media, like-minded people are more likely to engage in topic discussions and filter information more selectively^[3], thus more likely to lead to opinion polarization.

Existing studies shows that many elements are associated with the creation of opinion polarization^[4], but each of them is not a sufficient condition for the creation of opinion polarization. This is one of the shortcomings of the current research, which ignores the interactions between different elements. Therefore, this paper attempts to analyze the synergy effect of different elements on polarization of public opinion. By summarizing previous studies and shortcomings, this paper intends to carry out research in the following dimensions.

2.2 Public Opinion Event

According to the existing research, The characteristics of public opinion events are often taken as one of the important elements to measure the polarization of public opinion^[5]. Combining relevant theories and previous studies, this paper introduces the types of public opinion events, event triggers, the handling of the events by the triggers and the degree of harm of the events into the configuration model.

The type of event often determines the basis of Internet users' attitudes toward the event, and when the event involves elements such as government and officials, the public prone to have a critical tendency toward the event^[6]. This paper classifies public opinion events into four types: nationalism events, rights resistance events, moral privacy events and public power abuse events.

From the perspective of "relative deprivation", when comparing their own situation with that of event pointers, Internet users find that they tend to feel deprived when they are in a disadvantageous position, and this deprivation may lead to opinion polarization more easily, so this paper identifies event pointers as one of the factors influencing opinion polarization. Based on this, the responding of event triggers to events may lead to further fermentation of public sentiment, which may lead to further escalation of public opinion polarization in serious cases.

From the viewpoint of frustration-aggression theory, when people feel that their rights or emotions are violated or hurt, they are likely to have radicalized or extremized thoughts or emotions. In real life, when public opinion events are more harmful to society and netizens feel that their own rights or emotions are also hurt or infringed, they are likewise more likely to provoke emotional polarization reactions.

2.3 The Emotional State

The Psychology Of Emotion theory states that emotions always guide decision-making, consciously or unconsciously, and are the regulators of human behavior.^[7] Paul Ekman proposes six basic types of emotion including anger, disgust, fear, happiness, sadness and surprise^[8], where negative emotions are more likely to trigger radical statements such as hatred and anger, and more likely to lead to the polarization of public opinion. In previous studies on polarization, the observability and statistics of surface data are particularly emphasized in the research questions, and the emotional issues in the psychological black box are hardly addressed, and the consideration of the emotional states of Internet users in the public opinion field is also missing in the studies on opinion polarization.

2.4 Information Dissemination Process

This paper considers three variables in the process of information dissemination, including information ambiguity, media engagement and the guidance of opinion leaders.

In "Infotopia How Many Minds Produce Knowledge", Cass Sunstein argues that events with ambiguous information in real time at the beginning of exposure are more likely to trigger the phenomenon of group polarization^[9], so this paper considers the degree of information ambiguity as one of the key factors in the polarization of public opinion.

Media involvement: Resource mobilization theory suggests that the involvement of internal or external resources available to protest groups in collective action, including means of communication and media influence, can be a key resource for social movements and play an important role in the outbreak of events. This paper therefore considers the extent of media involvement in the study.

Opinion leaders guidance: The circle communication structure of social media gives more power to opinion leaders, whose views can increase the degree of opinion polarization^[10]. Both media and self-publishing can be seen as the core communication nodes in the dissemination of public opinion, i.e. opinion leaders. The extensive participation of opinion leaders can significantly increase the attention of public opinion events, and the thought of opinion leaders on public opinion events can also influence

the opinion sentiments of a large number of Internet users. This paper hypothesizes that the negative guidance of opinion leaders on public opinion events can incite a large number of netizens' emotions, which is more likely to lead to the polarization of public opinion.

To sum up, this study tries to analyze the influence of different elements on the polarization of public opinion from three dimensions: public opinion events, emotional states and information dissemination. We extract a total of 8 conditional variables from the three dimensions, and construct a configuration model as shown in Figure 1.

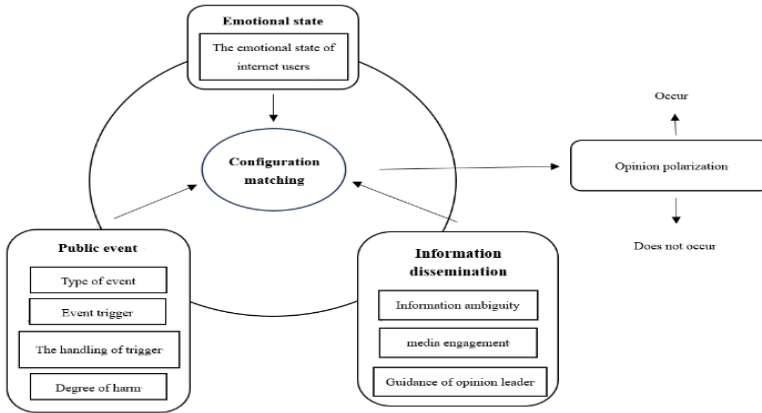


Fig. 1. Configuration model

3 Method and Data

3.1 csQCA

As a case study oriented research method, QCA can help the researcher to carry out theoretical and empirical dialogues and can systematically analyze data from small and medium samples [11], it is particularly suitable for explaining complex cause-induced phenomena and can obtain the best combination of causes that induced the formation of the phenomenon. While the study in this paper is intended to explore the form of the combination of factors and the role of the mutual coordination influence between them, and since most of the variables in this experimental sample are dichotomous, it was determined that csQCA (Clear Set Qualitative Comparative Analysis) was chosen as the final implementation method and the fsqca 3.0 software tool was chosen for the statistical analysis of the data.

3.2 Sample and Data

This paper uses the Weibo platform as the basic data source. As the largest social media platform in China today, Weibo has a huge user base, including a large number of

ordinary netizens, self-media, opinion leaders and major official media, and its platform nature and volume meet the requirements of case data acquisition. we obtain the text comment information related to cases through a crawler program, and collect a total of about 40414 related data in 29 cases.

3.3 Variable Setting and Data Process

Outcome Variable: The Occurrence of Opinion Polarization

In this study, We utilizes Size Parity^[12] to determine whether the opinion polarization occurs. When the value is greater than 0.7 or less than 0.3 it represents the emergence of polarization of public opinion

$$\text{Size Parity}_{\text{entropy}} = -\frac{1}{\ln G} \sum_{i=1}^G pi \ln pi$$

(1) Size parity

Notes: G stands for the number of viewpoint clusters and pi denotes the proportion of the total number of people in its viewpoint cluster.

The above selected cases are calculated based on the Size parity and the ratio of Aggressive comments, and the results are shown in Fig. 2

Time	Case	Size Parity	Aggressive comments	Opinion Polarization
2019.2	Zhai Tianlin Academic Fraud	0.25	>50%	Yes
2020.5	HongKong Riot	0.23	>50%	Yes
2020.5	TongZhuo Academic Fraud	0.83	>50%	Yes
2020.10	Three children policy case	0.51	<50%	No
2021.1	ZhengShuang Surrogate pregnancy	0.28	>50%	Yes
2021.2	H&M boycotts Xinjiang cotton case	0.26	>50%	Yes
2021.4	Testa Owners' Rights Defense case	0.46	<50%	No
2021.7	Chris Wu rape case	0.24	>50%	Yes
2021.7	100th anniversary of the founding of CCP	0.12	<50%	No
2021.7	Child abduction case	0.13	>50%	Yes
2021.7	Olympic Games Xiao Ruoteng silver medal case	0.26	>50%	Yes
2021.11	Leehom Wang cheating case	0.19	>50%	Yes
2022.1	'Chain Woman' case	0.19	<50%	No
2022.3	Shanghai Lockdown	0.8	>50%	Yes
2022.3	MU5735 crash case	0.78	<50%	No
2022.5	Porsche downsizing case	0.64	<50%	No
2022.5	Henan University female student brain hemorrhage case	0.59	<50%	No
2022.6	Tangshan violent injury case	0.09	>50%	Yes
2022.7	Henan village bank case	0.27	<50%	No
2022.7	Zhengzhou health code case	0.19	>50%	Yes
2022.8	Xuanzang Temple enshrines Japanese war criminals case	0.16	>50%	Yes
2022.8	Pelosi Taiwan case	0.2	>50%	Yes
2022.9	Sanya Lockdown	0.47	<50%	No
2022.9	Li Yifeng prostitute case	0.33	<50%	No
2022.10	Twentieth Congress of CCP	0.23	<50%	No
2022.11	Zhengzhou Foxconn Employees Returning Home case	0.69	<50%	No
2022.11	The death of Jiang Zemin case	0.18	<50%	No
2022.11	Urumqi fire case	0.16	>50%	Yes
2022.12	Anti-epidemic policy adjustment	0.59	<50%	No

Fig. 2. Polarized cases

Conditional Variable.

The types of public opinion event (TE), event triggers (ET), the handling by the event triggers (HET), degree of harm (DH), netizens’ emotional state (ES), information ambiguity (IA), Media Participation (MP) and opinion leader guidance (OLD) are considered condition variables in this study, which were assigned a value of “0” or “1” respectively, where “0” means false and “1” means true.

For TE and ET, since they do not have positive or negative attributes, 0 and 1 only represent their classification; for TE, rights protest events and entertainment banter events are marked as 1, while the other two types of events are marked as 0. For ET, 1 means the event is directed to the government or government related personnel, while 0 means the event is directed to a company or individual. For HET:1 means the event triggers a poor response to the event, 0 means a good response. For DH, 1 represents a high level of harm in public opinion events, while 0 vice versa. For ES, this paper classifies them into four categories according to the emotions that appear in public opinion events, anger, disgust, happiness, and sadness, and anger and disgust are marked as 1, while happiness and sadness are marked as 0. For IA, if there is missing or ambiguous information at the beginning of the event, it is judged to be 1, and 0 if the opposite is true. For MP 1 means a large number of official media are involved in public opinion events and 0 vice versa. OLD represents the influence status of the opinion leader on the event, with a mark of 1 if the opinion leader leads the event negatively and leads to further fermentation, and 0 if the opposite is true. The truth table is shown as Fig 3.

	TE	ET	HET	DH	ES	IA	MP	OLD	RESULT
EVENT1	0	1	1	1	1	1	1	1	1
EVENT2	1	0	1	1	1	1	1	1	0
EVENT3	0	1	1	1	1	0	0	0	1
EVENT4	1	1	1	1	0	0	0	0	0
EVENT5	1	0	0	0	0	1	0	1	0
EVENT6	1	0	0	0	0	1	1	1	1
EVENT7	1	0	1	0	0	1	1	1	0
EVENT8	0	0	1	1	1	0	1	1	1
EVENT9	1	0	1	0	1	1	1	1	1
EVENT10	1	1	0	0	0	0	1	0	0
EVENT11	0	0	1	1	1	0	1	1	1
EVENT12	0	1	1	1	1	1	1	1	1
EVENT13	1	1	1	1	1	0	1	1	1
EVENT14	0	1	1	1	1	1	1	1	0
EVENT15	1	0	1	0	1	0	0	1	0
EVENT16	1	1	1	0	1	0	0	1	1
EVENT17	0	1	0	1	0	1	1	0	1
EVENT18	1	1	0	0	0	0	1	0	0
EVENT19	1	0	0	1	0	0	1	0	0
EVENT20	0	1	0	0	0	0	1	0	0
EVENT21	0	0	0	0	0	0	1	0	0
EVENT22	0	1	1	1	1	1	1	1	1
EVENT23	1	0	1	0	0	1	1	1	1
EVENT24	1	0	1	1	0	1	1	1	1
EVENT25	1	0	1	0	0	1	1	1	0
EVENT26	1	0	1	1	1	1	1	1	1
EVENT27	0	1	0	0	0	0	1	0	0
EVENT28	1	0	1	0	0	1	1	1	0
EVENT29	0	0	1	1	1	1	1	1	1

Fig. 3. Truth table

Variable	Opinion polarization occurs	
	Consistency	Coverage
TE	0.466667	0.411765
~TE	0.533333	0.666667
ET	0.466667	0.538462
~ET	0.533333	0.500000
HET	0.866667	0.650000
~HET	0.133333	0.222222
DH	0.733333	0.733333
~DH	0.266667	0.285714
ES	0.733333	0.785714
~ES	0.266667	0.266667
IA	0.666667	0.625000
~IA	0.333333	0.384615
MP	0.866667	0.541667
~MP	0.133333	0.400000
OLD	0.866667	0.650000
~OLD	0.133333	0.222222

Fig. 4. Necessary condition analysis

4 Analysis

4.1 Necessary Condition Analysis

Before conducting the conditional grouping analysis, it is necessary to test the "necessity" of each condition individually. In this paper, we first test whether a single condition constitutes a necessary condition for opinion polarization to occur. With reference to domestic and international studies, this paper uses a consistency level greater than 0.9 as the criterion to determine whether a condition is necessary. Fig 4 shows the results of the necessity test for the emergence of opinion polarization using fsQCA3.0 software. As can be seen from the Figure, the level of consistency for all conditions is less than 0.9, so there is no single necessary condition that leads to the creation of opinion polarization.

4.2 Configuration Analysis

In contrast to necessary condition analysis, configuration analysis is generally used to reveal the adequacy of results arising from different groupings consisting of multiple conditions. Consistency is often used to measure the adequacy of a grouping, and Schneider and Wagemann^[13] pointed out that the consistency level for determining adequacy should not be lower than 0.75. However, the threshold value for consistency level was determined differently in different studies. Referring to the consistency threshold criteria of domestic and foreign scholars and the actual situation of this study, 0.8 was used as the consistency threshold criterion. Table 1 shows the configuration outcomes.

Table 1. Configuration analysis

Conditions	Configurations (opinion polarization occurs)			Configurations (opinion polarization does not occurs)
	C1	C2	C3	C4
TE	●	•	•	
ET	•	•	●	
HET	•	⊗	•	⊗
DH	●			⊗
ES	•		●	⊗
IA		●		⊗
MP		•		
OLD		•	•	⊗
Consistency	0.857143	1	1	1
Raw coverage	0.4	0.266667	0.266667	0.357143
Unique coverage	0.2	0.066667	0.066667	0.357143
Solution consistency	0.928571			0.923077
Solution coverage	0.866667			0.857143

Note: ● indicates the present of the core condition, ⊗ indicates the absent of the core condition, • indicates the present of the edge condition, ⊗ indicates the absent of the edge condition, and Blank indicates the condition is irrelevant.

4.3 Case Analysis

Configuration C1: TE*ET*HET*DH*ES.

In C1, TE and DH are the core conditions, and they both belong to the attributes of public opinion events. This path shows that when the event type is a breach of trust by public authority or a nationalistic event with a high degree of harm and the event initiator has a awful response and handling of the event, it will still lead to the polarization of public opinion under the effect of netizens' anger. The representative case is the "Urumqi fire incident". The emotions of netizens changed from sadness to anger. On the one hand, netizens were immersed in sadness and prayed for the victims of the incident; on the other hand, they expressed their dissatisfaction and criticism of the policy. After the incident, the Urumqi Fire Bureau's press conference, the relevant personnel said that the fire escape did not produce blockage, which conflicted with the information received by netizens, and part of the responsibility for the incident was "blamed" on the victims' own lack of escape ability, this excusing behavior also instantly ignited the negative emotions of netizens. and the public opinion field has also emerged a serious polarization phenomenon.

Configuration C2: TE*ET*~HET*IA*MP*OLD.

Configuration 2 highlights the role of information dissemination process, the core condition of this configuration path is information ambiguity, the event attribute in this path is a edge condition, and the event initiator has made a proper handling of the event, However, due to the nature of online information dissemination, when the information of an event is ambiguous or inflammatory, the online media will accelerate the dissemination and amplify the harm of the event, and coupled with the negative guidance of opinion leaders on the event, the phenomenon of polarization of public opinion will also occur. The majority of cases in this configuration are entertainment parodies, including "Wang Lihong's Cheating Case" and "Chris Wu's Rape Case".

Configuration C3: TE*ET*HET*ES*OLD.

The core conditions in configuration 3 are ET and ES, which show that the polarization of public opinion is more likely when the initiator of the event is the government and related personnel and the emotional state of Internet users is disgusted or angry. On one hand, government and government-related personnel are often seen as a class of power, with more "social privileges" than ordinary people, and therefore such people are more likely to be viewed critically. On the other hand, social opinion is an outward expression of netizens' emotions, and when netizens' emotions are mostly anger or disgust, it is more likely to lead to extremist remarks.

Configuration C4: ~HET*~DH*~ES*~IA*~OLD.

The elements in configuration C4 cover three attributes, the core conditions are ~ES and ~IA, C4 shows that when the event is not harmful to the society, and the information of the event is clear and not inflammatory and unknown, when the event initiator has explained or handled the event well, the netizens have not overreacted to the event, and the opinion leaders have not negatively guided the event, the phenomenon of polarization of public opinion basically does not occur under such circumstances.

5 Conclusion

Single Variable Cannot Lead To The Formation of The Phenomenon of Opinion Polarization.

On the whole, no single condition can be a necessary condition to influence the emergence of opinion polarization; behind the phenomenon of opinion polarization is the synergistic effect of multiple factors, and the effective combination of all factors together leads to the phenomenon of opinion polarization.

Event and Emotion Attributes Have More Influence on Opinion Polarization than Communication Attributes.

It is found that communication attributes have the least influence on the phenomenon of opinion polarization during the development of public opinion events. The reasons

for this are, firstly, as the results of previous related studies show the same, event attributes are more likely to stimulate the development of high-heat public opinion events. In the public opinion field, low-heat opinion events often fail to generate extensive discussions among online audiences, therefore opinion polarization is not possible. On the other hand, high-heat events can quickly grab the attention of online audiences and are more likely to trigger their emotional outreach and stand in line, so from this point of view, event attributes are more likely to drive opinion polarization. Second, group psychology can drive the spread of extreme emotions. In the Internet environment, Internet users are easily influenced by group contagion and group implication, and they easily accept the opinions of most people in the public opinion field. If there are already a lot of angry expressions or invectives in the public opinion field, the new public opinion audience is more likely to be infected by these emotions and show similar angry emotions. Since group emotions often show simple and exaggerated dual characteristics, polarized emotions escalate in both internal and external communication, shaping the emotional state and action direction of group members, and more easily leading to the polarization of public opinion.

Type of event, event trigger, degree of harm, emotional state and information ambiguity are the key influencing factors.

According to the analysis of the combination of each attribute, it can be concluded that the Type of event, event trigger, degree of harm, emotional state and information ambiguity play a major role in the generation of opinion polarization, and they are the "lower limit" that affects the generation of opinion polarization. The factors such as HET, MP and OLD must be combined with the core influencing factors to play a role and play a supplementary role in the generation of opinion polarization. The events most likely to lead to polarization of public opinion include but are not limited to: events damaging national sentiment, events of government's breach of trust and events of celebrities' breach of morality, which can explain more than 60% of the cases of polarization of public opinion in the case database.

Reference

1. J. Stoner, "A comparison of individual and group decisions involving risk," Massachusetts Institute of Technology, 1961.
2. W. Huang & J. Gao, "Analyzing the Generation Mechanism and Intervention Measures of the Audience's Emotional Polarization of Network Public Opinion," *Information studies: Theory&Application*, 42(5)122-127, 2019.
3. J. Lee, J. Choi, C. Kim & Y. Kim, "Social Media, Network Heterogeneity, and Opinion Polarization," *Journal of Communication*, pp. 702-722, 2014.
4. Y. Ge, Y. Qin & H. Zhao, "Social Media & Opinion Polarization in an Artificial Society: An AgentBased Modeling," *Chinese Journal of Journalism & Communication*, 42(02)67-99, 2020.
5. Y. Wu & F. Li, "A Review of Public Opinion Polarization," *JOURNAL OF INTELLIGENCE*, 41 (4) 98-103, 110, 2022.

6. L. Li & W. Zhen, "The Mechanism of Internet Mass Incident Outbreak: A Dual Modeling Study of "Communication Attributes" and "Event Attributes" - A Qualitative Comparative Analysis Based on 195 Cases," *Modern Communication (Journal of Communication University of China)*, 35 (2) 25-34, 2013.
7. Q. Xia & Y. Yuan, "From group polarization to public polarization: the way in and the turn of polarization research," *Journalism & Communication*, pp. 5-32+126, 2017.
8. P. Ekman, "Emotions Revealed," *Bmj*, p. 328, 2004.
9. C. R. Sunstein, *Infotopia: How Many Minds Produce Knowledge*, Oxford University Press, USA, 2006.
10. J. Shen & X. Wan, "The Characteristics and Formation Mechanism of Group Polarization in the Context of Network Stratification--Analysis of Microblog Comments Based on the "2-27 Incident"," *Modern Communication (Journal of Communication University of China)*, pp. 55-61, 2021.
11. R. Huang & Y. Gui, "The Internet and Homeowners' Collective Resistance: A qualitative comparative analysis," *Sociological Studies*, pp. 28(29-56), 2009.
12. A. Bramson, P. Grim, D. Singer & S. Fisher, "Disambiguation of social polarization concepts and measures," *The Journal of Mathematical Sociology*, pp. 80-111, 2016.
13. C. Q. Schneider and C. Wagemann, *Set-Theoretic Methods for the Social Sciences*, Cambridge University Press, 2012.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

