Platform values in times of crises and war

A study on the duality between platform and public values
during the Ukraine war

Lin Chang Meijer (4446372)

Master's thesis

Supervisor: Dr. Anne Helmond

Second reader: Dr. David Gauthier

Word count: 10996

Citation: Chicago Manual of Style

Utrecht University

MA New Media and Digital Culture 2022/2023

Utrecht, April 2023

Abstract

Today, the circulation of misinformation on platforms is threatening the public life of citizens, especially during times of crises and war. The framework of regulation *by* platforms provides insight into how platforms regulate content through their policies. This thesis examines the content policies of the platforms TikTok, Facebook, and Twitter during the Ukraine war through a textual analysis. It is argued that these policies reveal the dual nature of platform values, in which platforms operate as private actors in a public sphere. To illustrate, this study shows the difference in platform values across these platforms, such as *safety, authenticity,* and *freedom of speech.* However, these policies reveal that *freedom of speech* is used to serve the interests of Facebook and Twitter but clashes with the public value of *responsibility* and *accountability*. Besides that, the Ukraine war also constructed other public values such as *trust* and *transparency*, which interfere with the opaque and ambiguous policy enforcements of these platforms. As a result, it has been shown that during the war in Ukraine, TikTok, Facebook, and Twitter shifted their self-regulatory systems into co-regulatory practices.

Keywords: *Platforms, content regulation, misleading content, Ukraine war, platform governance*

Table of contents

Introduction

Social media platforms have become increasingly important subjects of study, especially in times of crises and war. To illustrate, the EU has raised its concerns about the spread of disinformation. Disinformation is defined by the EU as inaccurate or misleading content with the purpose of deceiving a financial advantage that could be harmful to the public.[1] With that being said, the EU emphasizes the role of online platforms in this debate.[2] In 2018, the EU established the Code of Practice for online platforms, trade organizations, and significant actors in the advertising industry, in which all groups enhance their online policies to reduce misinformation.[3] The Code of Practice aimed for better responsibility and openness on online platforms, as well as enhancing the platforms' disinformation policies. Particularly during the COVID-19 crisis, inaccurate information about vaccines, incorrect medical information, and false statements put the public's health in peril.[4] Since this occurrence is especially happening on platforms, media scholars Michael Dieter et al. illustrate the need to study regulations by platforms as different fields regarding public communication, education, and healthcare get intertwined.[5] However, the spread of misleading content was not only a problem during COVID-19, but it is also an ongoing problem that the EU now recognizes as a threat to citizens, as the EU has acknowledged the use of aggressive pro-Kremlin disinformation and war propaganda as a strategy of Russia against Ukraine.[6] Similarly, with COVID-19, platforms have dealt with enormous amounts of misleading content, including videos, memes, and photos related to the war. News outlets even called it the 'TikTok war' in which social media platforms are used to document wars and conflicts.[7] Here, a lot of this content contains misinformation, and huge platforms like Facebook, Twitter, and TikTok are flooded with false content.[8] Again, the need to study platform policies is especially evident when looking at times of disruption when misleading content is threatening the lives of citizens.

---

[1] European Commission, 'Tackling Online Disinformation | Shaping Europe's Digital Future', 9 February 2023, https://digital-strategy.ec.europa.eu/en/policies/online-disinformation.
[2] European Commission.
[3] European Commission, 'A Strengthened EU Code of Practice on Disinformation', accessed 22 March 2023, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan/strengthened-eu-code-practice-disinformation_en.
[4] European Commission.
[5] Michael Dieter et al., 'Pandemic Platform Governance: Mapping the Global Ecosystem of COVID-19 Response Apps', *Internet Policy Review* 10, no. 3 (2021): 1–28. 2.
[6] European Commission, 'A Strengthened EU Code of Practice on Disinformation'.
[7] Kyle Chayka, 'Ukraine Becomes the World's "First TikTok War"', *The New Yorker*, 3 March 2022, https://www.newyorker.com/culture/infinite-scroll/watching-the-worlds-first-tiktok-war.
[8] Sardarizadeh, 'Ukraine War: False TikTok Videos Draw Millions of Views', *BBC News*, 25 April 2022, sec. Reality Check, https://www.bbc.com/news/60867414.

As a result, the discipline of platform governance studies how platforms regulate content outside of legal and political studies. For instance, media scholars have looked at the politics of platforms, in which communication scholar Tarleton Gillespie has looked at the term 'platform' itself as a discursive tool to investigate how platforms frame their services and technology to present themselves as neutral facilitators towards their end-users, advertisers, and content creators.[9] While platforms want to appear as neutral as possible, many media scholars such as Gillespie have contested this portrayal, as they are facilitators of public speech.[10] In addition, media scholars José van Dijck et al. have studied the implications that arise when private actors such as platforms get intertwined with societal issues."[11] Here, to understand regulation *by* platforms, Gillespie argues that content guidelines deliver important discursive work, as they show not only the dual nature of platform policies but also the challenges of content regulation over time.[12] With that being said, media scholars Rebecca Scharlach et al. illustrate how guidelines also reveal the duality of platforms through values, as they research how platforms influence their decisions about what is permitted and prohibited through platform values.[13] Here, they use the loose definition of describing platform values as "the underlying principles governing and expressed through social media."[14] As a result, the platform guidelines show the values of platforms to protect the internet and the commercial values of platforms, such as the strategies and purposes of operations.[15] Because of this, it is crucial to look at the guidelines of platforms to reveal these dualities.

Here, I argue that studying platform values illustrates the dual positioning of platforms, where they have a social responsibility to the public as arbiters of public speech on the one hand but also want to pursue their commercial values as private actors. Therefore, to further study regulation *by* platforms, I question in what ways the content policies of platforms surrounding misleading content illustrate the duality between commercial and public values during the Ukraine war. This study will be an addition to the study of Scharlach et al., in which it also

---

[9] Tarleton Gillespie, 'The Politics of "Platforms"', *New Media & Society* (2010): 4, https://doi.org/10.1177/1461444809342738. 7.

[10] Tarleton Gillespie, 'Regulation of and by Platforms', in *The SAGE Handbook of Social Media*, ed. Jean (Jean Elizabeth) Burgess, Alice Marwick, and Thomas Poell, 1st ed., 1 online resource vols (Los Angeles: SAGE reference, 2018), 254–78, https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5151795. 255.

[11] José van Dijck, Thomas Poell, and Martijn de Waal, *The Platform Society: Public Values in a Connective World* (New York, NY: Oxford University Press, 2018). 3.

[12] Gillespie, 'Regulation of and by Platforms'. 264.

[13] Rebecca Scharlach, Blake Hallinan, and Limor Shifman, 'Governing Principles: Articulating Values in Social Media Platform Policies', *New Media & Society*, 7 March 2023, 1–20, https://doi.org/10.1177/14614448231156580. 7.

[14] Scharlach, Hallinan, and Shifman. 3.

[15] Scharlach, Hallinan, and Shifman. 2.

studies platform policies but also addresses their need to examine additional platform value formation by examining how values are included in moderation practices.[16] The research question will be broken down into sub-questions based on examining three different layers: 1) The first sub-question covers which values are evoked by TikTok, Facebook, and Twitter when regulating misleading content. 2) The second sub-question covers what position TikTok, Facebook, and Twitter have regarding the Ukraine war while being moderators of public speech. 3) The last sub-question covers which platform enforcements have resulted in changes within TikTok, Facebook, and Twitter when combating misleading content during the Ukraine war. Here, the policies on misleading content by TikTok, Facebook, and Twitter will be analyzed in comparison to each other through a textual analysis, a method that looks similar to the method by Scharlach et al. Here, a textual analysis will allow us to critically engage with media texts to identify how texts construct values and beliefs.[17] As a result, the community guidelines, transparency center, and newsroom/blogs on TikTok, Facebook, and Twitter will be examined. First, the theoretical framework will introduce three main subjects, in which the neutrality of platforms, regulation *of* and *by* platforms, and self-regulatory practices will be discussed and why these concepts will be needed. Thereafter, the methodology will explain the case, corpus, and method. In the analysis, the sub-questions will be answered, and the main findings of the analysis will be presented. The thesis will end with a discussion and conclusion.

[16] Scharlach, Hallinan, and Shifman. 15.

[17] Marie Gillespie and Jason Toynbee, *Analysing Media Texts*, Understanding Media, [4] (Maidenhead, Berkshire, England ; Open University Press in association with the Open University, 2006), http://catdir.loc.gov/catdir/toc/fy0703/2006298258.html. 2.

1. Theoretical framework

The theoretical framework is used to explore important concepts and frameworks that illustrate the existing literature on platform governance. Here, the theoretical framework first introduces platforms as neutral facilitators, regulations *of* and *by* platforms, and finally the self-regulatory practices of platforms during times of crises. Thereafter, each section ends with a discussion of why these concepts and frameworks are needed to further investigate platform regulations during times of crises and war.

Defining the word 'platform'

To understand platform governance, it is important to first discuss the term 'platform' itself. While doing so, various media scholars use different definitions, which means that the word 'platform' itself is an ambiguous term. For instance, Plantin et al. illustrate how the word 'platform' is used for game design, as well as for websites and social media applications.[18] Media scholars José van Dijck et al. define platforms as "a programmable architecture designed to organize interactions between users."[19] Media scholar Jean Burgess refers to platforms as "*computational architectures* on which features and services can be built and as *discursive spaces* for cultural expression and audience engagement."[20] Communication scholar Tarleton Gillespie describes social media platforms as "new opportunities to speak and interact with a wider range of people, organizing them into networked publics."[21] Others call social platforms media or social intermediaries, facilitators, or hosts, as they all present and facilitate content on their sites. While it is acknowledged that platforms have multiple names and meanings within different sectors, this thesis uses the word 'platform' to refer to programmable social media intermediaries that organize interaction between different audiences, while facilitating the circulation of information and content. Following this, the section below reveals how platforms use this ambiguous positioning as a strategy.

---

[18] Jean-Christophe Plantin et al., 'Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook', *New Media & Society* 20, no. 1 (1 January 2018): 293–310, https://doi.org/10.1177/1461444816661553. 2.

[19] José van Dijck, Thomas Poell, and Martijn de Waal, *The Platform Society: Public Values in a Connective World* (New York, NY: Oxford University Press, 2018). 9.

[20] Jean Burgess, 'Platform Studies', in *Creator Culture* (New York University Press, 2021), 26, https://doi.org/10.18574/nyu/9781479890118.003.0005.

[21] Tarleton Gillespie, 'Regulation of and by Platforms', in *The SAGE Handbook of Social Media*, ed. Jean (Jean Elizabeth) Burgess, Alice Marwick, and Thomas Poell, 1st ed., 1 online resource vols (Los Angeles: SAGE reference, 2018), 254, https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5151795.

1.1 Platforms as neutral facilitators

Before the study of platform governance is introduced, it is important to understand the crucial role that platforms have in the daily lives of citizens. Therefore, this section covers how platforms are viewed as a discursive practice where platforms position themselves as non-neutral. However, this positioning is contested by many media scholars, as platforms reshape public practices in society. Both communication scholars Tarleton Gillespie, and Rasmus Nielsen and Sarah Ganter use the economic framework of platforms as a *multi-sided market* in which the word platform is used strategically by platforms to appear to the many audiences that they serve.[22] Gillespie argues that platforms position themselves strategically in a way that their definition of what they are is concrete enough to make clear what they offer but is vague enough so that their positioning appeals to multiple contexts and audiences.[23] Here, the word 'platform' is a descriptive term for digital intermediaries, which is on the one hand based on its meaning within the computational and architectural industries, but on the other hand, also based on figurative and political concepts.[24] To illustrate, platforms such as YouTube need to present themselves strategically to their users, advertisers, and professional content producers, define a role and a set of standards that are acceptable to all audiences while also serving their financial interests and resolve or at the very least smooth over any contradictions that may exist.[25] However, as Gillespie argues, it is not only favorable for platforms to position their services like this but also interesting in a legal sense, which gives platforms the freedom to conduct their businesses with fewer restrictions and less liability.[26] As a result, this suggests that platforms want to remain as neutral as possible in this regard, so that they are not only attractive to as many audiences as possible but also ensure that they are not legally responsible for the content on their platforms.[27]

This discursive positioning of 'neutrality' by platforms is contested by media scholars. One of the reasons is that these platforms have immense power over society. To illustrate, Nielsen and Ganter use the framework of 'platform power' in which they illustrate the dependency of society on these platforms, as many citizens use social media to access news and retrieve

---

[22] Rasmus Kleis Nielsen and Sarah Anne Ganter, *The Power of Platforms: Shaping Media and Society*, 1 online resource vols (New York, NY: Oxford University Press, 2022), https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=3224265. 167.
[23] Tarleton Gillespie, 'The Politics of "Platforms"', *New Media & Society*, 2010, 3, https://doi.org/10.1177/1461444809342738.
[24] Gillespie. 3.
[25] Gillespie. 7.
[26] Gillespie. 9.
[27] Gillespie. 12.

information.[28] Here, Nielsen and Ganter argue that platform power is a generative form of power, in which businesses use platforms to carry out their interests, while also becoming entangled in a highly asymmetric relationship with the platform.[29] As a result, this power is reshaping the media landscape, where platforms influence how news is received, produced, and distributed, how politics works, and how citizens connect.[30] To elaborate on this framework of 'platform power', media scholars José van Dijck et al. have studied the implications that arise when private actors such as platforms get intertwined with societal issues, in which they also argue that "platforms are neither neutral nor value-free constructs".[31] This is because platforms' algorithms, business models, and user activity amplify ideological tenets.[32] Here, challenges arise when platforms' ideological systems contest public values such as accountability and transparency.[33] Therefore, they question who is or should be guarding public values in societies that are structured through online intermediaries.[34] Here, public values refer to the value that actors add to society, which is developed through collective involvement and a common set of standards and ideals.[35] One example of that is the internet. Communication scholars Jeffrey Treem et al. state that platforms position themselves as democratic sites in which the public can exchange and discuss ideas deliberately.[36] Here, platforms position themselves as democratic, neutral, and value-free sites in the name of the 'common good'. However, Treem et al. state the opposite and argue that social platforms are a place where public speech is shaped, developed, and debated and therefore they should not be seen as a realistic representation of society's opinion.[37] As a result, platforms not only connect social and economic actors but also develop new sorts of value systems.[38] Media scholar Tijana Milosevic adds that these regulations by platforms raise questions of responsibility and transparency since the digital public sphere is monitored by private actors.[39]

---

[28] Rasmus Kleis Nielsen and Sarah Anne Ganter, *The Power of Platforms: Shaping Media and Society*, 1 online resource vols (New York, NY: Oxford University Press, 2022), 1.
https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=3224265.
[29] Nielsen and Ganter. 2.
[30] Nielsen and Ganter. 167.
[31] Dijck, Poell, and Waal, *The Platform Society*. 3.
[32] Dijck, Poell, and Waal. 22.
[33] José van, 'Governing a Responsible Platform Society', in *The Platform Society*, ed. José van Dijck, Thomas Poell, and Martijn de Waal (Oxford University Press, 2018), 137,
https://doi.org/10.1093/oso/9780190889760.003.0008.
[34] Dijck, Poell, and Waal, *The Platform Society*. 8.
[35] Dijck, Poell, and Waal. 22.
[36] Jeffrey W. Treem et al., 'What We Are Talking About When We Talk About Social Media: A Framework for Study', *Sociology Compass* 10, no. 9 (2016): 773, https://doi.org/10.1111/soc4.12404.
[37] Treem et al. 773.
[38] Dijck, Poell, and Waal, *The Platform Society*. 25.
[39] Tijana Milosevic, 'Social Media Companies' Cyberbullying Policies', *International Journal of Communication* 10, no. 0 (13 October 2016): 5167.

Therefore, this discursive positioning by platforms is used in this thesis to study how platforms position themselves during the Ukraine war and if this resonates with their strategic positioning as neutral facilitators. Especially with the Ukraine war, the discursive positioning of platforms is important to stud since platforms have an immense power on the lives of citizens, what has been revealed by the framework of 'platform power'. Similar to any great event, many citizens use platforms to retrieve information about the war, which makes the role of platforms even greater in this case. However, while platforms want to appear as democratic, neutral, and value-free sites, it is argued that platforms want to appear as neutral as possible. This has been shown by Gillespie, in which platforms not only position themselves strategically toward their audience but also profit from this in a legal sense. Therefore, it is important to look at how platforms position themselves through platform regulations.

1.2 Regulations *of* and *by* platforms

The study of platform governance offers insightful frameworks to help examine platform regulations. For instance, both frameworks -regulation *of* platforms and regulation *by* platforms- are used to illustrate the legal positioning that platforms profit from. Especially the regulation *by* platforms shows which policies platforms have and enforce, which can be read in the 'Terms of Service' and the 'Community Guidelines'. This does not only illustrate what is and what is not allowed on the platforms but also the challenges that platforms face. Here, studying platform regulations has gained a special interest from political science and legal scholars outside media disciplines, which has resulted in the field of platform governance. This study looks especially at the governance and power relationships facilitated through platforms. To illustrate, legal scholar Robert Gorwa argues that these platform services are partly shaped by the governance processes that involve policy and regulatory constraints.[40] Gorwa states that platform governance is a term borrowed from the original definition of governance, in which institutions make and enforce rules and provide services. However, platform governance focuses more on a broader definition in which not only governments but also platforms organize, structure, and regulate online infrastructures.[41] For instance, policies, terms of service, algorithms, interfaces, and regimes all fall under the study of platform governance.[42] Here, these scholars elaborate on the economic framework of platforms as a *multi-sided market*, discussed

---

[40] Robert Gorwa, 'What Is Platform Governance?', *Information, Communication & Society* 22, no. 6 (12 May 2019): 4, https://doi.org/10.1080/1369118X.2019.1573914.
[41] Gorwa. 3.
[42] Gorwa. 4.

in de previous section. They do this by using a political science framework that sees platforms as corporate actors competing in the field of contested global governance.[43] For instance, media scholars such as Tarleton Gillespie, José van Dijck, and Taina Bucher were among the first to be interested in 'platform politics' in which they analyzed the governance and power relations mediated by platforms. Internet governance scholars Laura DeNardis and Andrea Hackl focus on the importance of internet governance through social media platforms. When looking at social media through the lens of internet governance, it allows insight to how social media actors can be regulated through the enforcement of intellectual property rights, antitrust, privacy, or other public interest issues.[44] This is important since the technological design and policies made by these social media platforms function as a form of *privatized governance* that enacts rights and controls the flow of information online.[45] As a result, platform governance encompasses how platform practices, policies, and affordances interact with political forces to make platforms more democratically accountable.[46]

The study of platform governance looks not only at the regulations *of* platforms but also at the regulations *by* platforms. Here, the regulation *of* platforms studies how platforms are governed by institutions through legislation. Regulation *by* platforms, on the other hand, looks at the policies that platforms make themselves. First, he illustrates that the regulations *of* platforms remain a grey area due to the protection that platforms legally have. Here, Gillespie stated earlier that platforms position themselves ambiguously, which is made possible due to the regulations *of* platforms. To illustrate, the online policies that apply to platforms were established before what is now known as 'platforms'.[47] Platforms, often referred to as 'internet intermediaries', have one main character as they connect or host interactions between third parties on the internet.[48] Because of this, platforms can be seen as moderators between different audiences. In the United States, social media platforms still fall under the communication law, called Section 230. However, this law is twofold: first, intermediaries will not be regarded as 'publishers' of content but are considered a service that offers access to the internet.[49] Because of this, they cannot be responsible for the speech of their users. Second, when intermediaries

---

[43] Gorwa. 7.
[44] L. DeNardis and A. M. Hackl, 'Internet Governance by Social Media Platforms', *Telecommunications Policy* 39, no. 9 (1 October 2015): 2, https://doi.org/10.1016/j.telpol.2015.04.003.
[45] DeNardis and Hackl. 2.
[46] Gorwa, 'What Is Platform Governance?' 1.
[47] Tarleton Gillespie, 'Regulation of and by Platforms', in *The SAGE Handbook of Social Media*, ed. Jean (Jean Elizabeth) Burgess, Alice Marwick, and Thomas Poell, 1st ed., 1 online resource vols (Los Angeles: SAGE reference, 2018), 255, https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5151795.
[48] Gillespie. 256.
[49] Gillespie. 257.

do intervene, they are not required to meet the standards of effective policing, since they are, again, not publishers of content.[50] However, as read before, platforms not only moderate but also sort, distribute, and reshape public speech through their infrastructures, affordances, and algorithms.[51] Here, platforms don't just moderate speech but also create it, which is revealed when looking at what information is presented, allowed, and discouraged. Because of this legal positioning, platforms can make their own rules and policies on how they regulate content and therefore might undermine public values such as responsibility and transparency.

With that being said, the regulation *by* platforms provides insight into the role that platforms give to themselves and how platforms govern online speech. Although it is important to know how platforms are governed through legislation, looking at the regulation *by* platform reveals which policies platforms make on their own. Here, these policies are mostly written in the 'terms of service' and the 'community guidelines'. The terms of service offer a legal document with the terms that platforms enact for their services and the conditions, obligations, and rights that users have while using the platform. The community guidelines, on the other hand, portray norms and rules, directing how users should behave on a platform.[52] Therefore, these guidelines reveal how platforms govern through their policies. Most of these social media platforms have in their community guidelines similar rules that prohibit content that includes hate speech, sexual content, violence, self-harm, or illegal activities. As Gillespie argues, these guidelines offer important discursive work for researchers, as platforms regard themselves as facilitators of public speech.[53] In other words, the guidelines also reveal which values platforms consider important. To elaborate on this, media scholars Scharlach et al. use platform guidelines to reveal the duality of these platform values. Although scholars have been studying public values that are related to platforms, conceptualizing the word 'values' in relation to platforms has been a challenge because cognition and action by platforms can be contradictorily.[54] With this in mind, Scharlach et al. use the loose definition of describing platform values as "the underlying principles governing and expressed through social media".[55] Here, they build on the concept of *values as principles,* which looks at how platforms guide their judgment in regard

---

[50] Gillespie. 258.
[51] Gillespie. 257.
[52] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 46, https://doi.org/10.12987/9780300235029-003.
[53] Gillespie, 'Regulation of and by Platforms'. 264.
[54] Rebecca Scharlach, Blake Hallinan, and Limor Shifman, 'Governing Principles: Articulating Values in Social Media Platform Policies', *New Media & Society*, 7 March 2023, 3, https://doi.org/10.1177/14614448231156580.
[55] Scharlach, Hallinan, and Shifman. 3.

to what is allowed and what is not allowed in their policies.[56] As a result, they illustrate the importance of analyzing the public values platforms hold, as it shows the duality between commercial and public values of platforms.[57] Especially within the community guidelines, it is made clear how especially the U.S. platforms want to protect online speech.[58] This is illustrated by media scholars Jessica Maddox and Jennifer Malson, who introduce 'the marketplace of ideas metaphor' for speech. This has been widely used by platforms based in Silicon Valley, where the metaphor promotes a competitive, free market in which all ideas are heard and welcomed.[59] As Maddox and Malson state, platforms use this marketplace of ideas to justify their legislation and thereby ensure that they can operate freely.[60] Therefore, platforms urge users to embrace freedom of expression within the bounds of the community guidelines in which they present themselves as a place to do just that. However, it also illustrates commercial values, such as the strategies and purposes of operations, causing an amplification of radical material, privacy issues, and political prejudice.[61] Therefore, by understanding policy preferences and governance philosophies, the ideologies and imaginaries of platforms are illuminated.[62]

The study of platform governance illustrates the extraordinary need to study regulations *by* platforms to further understand platform power. Here, it is argued that conversation on platforms can only happen within the boundaries of platforms, which are fabricated through platform policies. However, as we have seen with COVID-19, the public only intervenes with these policies when things go wrong.[63] This reveals the role of platforms as arbiters of public speech, which create and enforce their platform values. As discussed earlier, the legal positioning of platforms allows them to make their own rules and policies, which can undermine public values such as responsibility and transparency. Especially during crises and war, these public values are important. As the EU Code of Practice advocates for these values, regulation *by* platforms illustrates that platforms make this more difficult with their policies. As stated

---

[56] Scharlach, Hallinan, and Shifman. 7.
[57] Scharlach, Hallinan, and Shifman. 1.
[58] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 47, https://doi.org/10.12987/9780300235029-003.
[59] Jessica Maddox and Jennifer Malson, 'Guidelines Without Lines, Communities Without Borders: The Marketplace of Ideas and Digital Manifest Destiny in Social Media Platform Policies', *Social Media + Society* 6, no. 2 (1 April 2020): 3, https://doi.org/10.1177/2056305120926622.
[60] Maddox and Malson. 4.
[61] Scharlach, Hallinan, and Shifman, 'Governing Principles'. 2.
[62] Pawel Popiel, 'Digital Platforms as Policy Actors', in *Digital Platform Regulation: Global Perspectives on Internet Governance*, ed. Terry Flew and Fiona R. Martin (Cham: Palgrave Macmillan, 2022), 136.
[63] European Commission, 'A Strengthened EU Code of Practice on Disinformation'.

earlier, guidelines expose the calculated and self-serving behavior of platforms, which leaves the question of responsibility open and illustrates the standards and values that platforms amplify through their policies. Hence, studying platform values reveals the dual positioning of platforms, where they have a social responsibility to the public as arbiters of speech, on the one hand, but also want to pursue their commercial values as private actors, especially during uncertain times where misleading content is causing social unrest. Therefore, it is important to not only study platform guidelines, as these policies illuminate not only what is and is not allowed on the platform but also the inner conflicts of platforms themselves.

1.3 Self-regulation in times of crises

This section sheds light on how platforms regulate content through self-regulation. By doing this, platforms intervene by themselves, which raises concerns surrounding their lack of transparency. Here, the framework of platform transience by Barret and Kreiss illustrates that platforms change continuously, which often means unnoticeable changes in their policies, procedures, and affordances.[64] Although platforms are provided with legal immunity, they still interfere with their content policies. This current governance strategy is called 'self-governance'.[65] Self-governance means that platforms not only manage the public area but also implement and respond when there are complaints by third parties about content. A disadvantage to this is that platforms' decision-making remains internal, making it unclear to third parties how platforms regulate content.[66] This is legally acceptable because, within the boundaries of the law, platforms are not responsible for the material that is being shared on their platforms. However, a downside of this governance strategy is the lack of transparency, since providing insight into the regulatory systems is often voluntarily done by platforms.[67] To illustrate, media scholar Rotem Medzini illustrates the implications of self-regulatory systems that platforms use, which result in systems that allow for private censorship and a lack of democratic responsibility.[68] Due to public concerns, platforms slowly start to implement changes within their technical infrastructure or become more transparent regarding their policies. As a result, Medzini developed the concept of 'enhanced self-regulation', which is used to describe the enhancement of self-regulation and the execution of policies by rule-

---

[64] Bridget Barrett and Daniel Kreiss, 'Platform Transience: Changes in Facebook's Policies, Procedures, and Affordances in Global Electoral Politics', *Internet Policy Review* 8, no. 4 (31 December 2019): 2.
[65] Gorwa, 'What Is Platform Governance?'. 9.
[66] Gorwa.. 10
[67] Gorwa. 9.
[68] Rotem Medzini, 'Enhanced Self-Regulation: The Case of Facebook's Content Governance', *New Media & Society* 24, no. 10 (1 October 2022): 2228, https://doi.org/10.1177/1461444821989352.

makers of self-regulatory regimes.[69] For instance, Zuckerberg has developed an independent oversight board to provide responsibility and control over Facebook's policies. This led to the creation of a council made up of Facebook users, partners from the public and private sectors, rule-makers, third-party implementers, and auditors. Here, third-party implementors, including fact-checkers, maintain their autonomy on how to respond to content that they flag.[70] As a result, these decisions made by fact-checkers reshape how Facebook enforces its rules, which makes its regulations more co-regulatory.

However, this framework of enhanced self-regulation focuses on infrastructural changes by platforms due to the pressure of public concerns. This indicates that platforms only change due to outside pressure. While this may appear to the public eye, political communication scholars Bridgett Barrett and Daniel Kreiss state that platforms change continuously, which is seen in their use of machine learning, artificial intelligence, and algorithms.[71] This study also shows a literature gap regarding the causes and repercussions of why platforms change. As a result, they offer a framework called 'platform transience' that looks at the continual and fast change of platforms by specifically looking at their changing policies, procedures, and affordances.[72] Here, media scholars Natali Helberger et al. emphasize the importance of studying the architecture of platforms since this shapes how users communicate and connect, engage, criticize, evaluate, follow, and interact with one another.[73] Analyzing design decisions such as flagging systems, algorithms, and interaction levels with content reveals to what extent platforms take responsibility for the content that is being displayed.[74] However, the paradigm of platform transience also has some implications, as this fast, continuous pace of change lacks transparency. Therefore, it can be difficult for public representatives to determine the social and political repercussions of platforms, hold them accountable, or create policy interventions.[75] For instance, during the 2016 U.S. presidential elections, it was revealed that changes within the policies, procedures, and affordances of Facebook have led to an unequal information environment and reshaped the fairness of electoral competition.[76] Although Barret and Kreiss

---

[69] Medzini. 2229.
[70] Medzini. 2240.
[71] Bridget Barrett and Daniel Kreiss, 'Platform Transience: Changes in Facebook's Policies, Procedures, and Affordances in Global Electoral Politics', *Internet Policy Review* 8, no. 4 (31 December 2019): 2.
[72] Barrett and Kreiss. 4.
[73] Natali Helberger, Jo Pierson, and Thomas Poell, 'Governing Online Platforms: From Contested to Cooperative Responsibility', *The Information Society* 34, no. 1 (1 January 2018): 5, https://doi.org/10.1080/01972243.2017.1391913.
[74] Helberger, Pierson, and Poell. 5.
[75] Barrett and Kreiss, 'Platform Transience'. 3.
[76] Barrett and Kreiss. 17.

used the U.S. elections as a case study, it illustrates how platforms constantly alter their strategies without giving third parties insight or notice.

While looking at the framework of platform transience, the need to study why and how platforms change their policies, procedures, and affordances is especially relevant for generated content during times of crisis. During COVID-19, the circulation of misinformation has been apparent, which has often been referred to as the 'infodemic' and has been raising concerns about platforms amplifying misinformation.[77] As a result, the improvement of content moderation has been of extraordinary need during the COVID-19 pandemic. This concern has been shared among many scholars, as media scholars Michael Dieter et al. state how platforms play a crucial societal and infrastructural role in daily life, especially when public communication, education, and health care get intertwined on platforms during times of crises.[78] Besides that, legal scholar Evelyn Douek demonstrates the extensive control that platforms have over online speech as a result of the platform's visible implementation of proportionate speech restrictions during the pandemic.[79] For instance, platforms have removed false content during the epidemic in an exceptionally aggressive manner, despite Facebook's claim that doing so simply expanded the categories of content that were prohibited and were in violation of their existing policies.[80] These policies, which are often found under the platforms' community guidelines, all refer to the spreading of fake news, malevolent hearsay, threats and taunts, hate speech, and propaganda. However, Helberger et al. state that the prohibition of misleading content is a subject that platforms struggle with because the connection between public values like public safety and the dissemination of reliable information is not always as clear-cut as it is depicted in public discourse.[81] Thus, especially with misleading content, it is difficult to make a distinction between what the public consensus is and what is considered false or factual information. As a result, Twitter broadened its definition of harmful content and intervened when content went against the global and health guidelines of COVID-19.[82] Here, an affordance that has been widely used to identify harmful content is that of 'flagging'.[83] 'Flagging' has been introduced by social media platforms as a tool to report harmful content,

---

[77] Leticia Bode and Emily K. Vraga, 'Correction Experiences on Social Media During COVID-19', *Social Media + Society* 7, no. 2 (1 April 2021): 1, https://doi.org/10.1177/20563051211008829.

[78] Michael Dieter et al., 'Pandemic Platform Governance: Mapping the Global Ecosystem of COVID-19 Response Apps', *Internet Policy Review* 10, no. 3 (2021): .2.

[79] Evelyn Douek, 'Governing Online Speech: From "Posts-As-Trumps" to Proportionality and Probability', SSRN Scholarly Paper (Rochester, NY, 23 August 2020), 800. https://doi.org/10.2139/ssrn.3679607.

[80] Douek. 800.

[81] Helberger, Pierson, and Poell, 'Governing Online Platforms'. 6.

[82] Douek, 'Governing Online Speech'. 7.

[83] Helberger, Pierson, and Poell, 'Governing Online Platforms'.

which has been a possible solution to curate large collections of user-generated content.[84] Media scholar Elmie Nekmat illustrates that, in combination with a platform's flagging features, fact-checkers examine content for any misinformation.[85] Hence, platforms self-regulatory systems are mostly revealed during social unrest, as platforms not only moderate more expensively they introduce new forms of content moderation.

As stated earlier in the previous section, looking at the guidelines of platforms is crucial in revealing platform values. However, the framework of enhanced self-regulation by Medzini dives deeper into the changes within the policies of platforms' self-regulatory practices. By doing so, this framework helps to understand the changes that platforms enforce by increasing their accountability in issuing some of their content regulations to third parties. Therefore, it is argued that the framework of platform transience offers an insight into this, as misleading content is not only a public concern but also a challenge where platforms try to improve their content moderation, often in opaque manners. Especially with the issue of the increasing flow of misinformation during crises, platforms change their policies, procedures, and affordances. However, misleading information is not only a problem during COVID-19 but also has been around during the Ukraine war, making it important to examine what changes platforms are enforcing, which eventually reshapes the public conversation during this event.

---

[84] Kate Crawford and Tarleton Gillespie, 'What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint', *New Media & Society* (5 August 2014): 410.
[85] Elmie Nekmat, 'Nudge Effect of Fact-Check Alerts: Source Influence and Media Skepticism on Sharing of News Misinformation in Social Media', *Social Media + Society* 6, no. 1 (1 January 2020): 1, https://doi.org/10.1177/2056305119897322.

2. Methodology

In the theoretical framework, it is argued that content guidelines can be used to understand the regulation *by* TikTok, Facebook, and Twitter during the Ukraine war. Therefore, the analysis focuses on the content policies of TikTok, Facebook, and Twitter to reveal the dualities between the public and commercial values of these platforms. In this section, the methodology is introduced, in which the choice of case, corpus, and method is explained to construct the analysis.

2.1 Case and corpus

The framework of platform power shows how dependent people are on social media platforms since platforms constitute how information is received and how people interact with each other.[86] Because of this power, the EU has urged platforms to take responsibility and prevent the spread of misleading content, especially during times of crises and war.[87] Therefore, the largest social media platforms, TikTok, Facebook, and Twitter were analyzed. According to media scholar Joanne Gray, TikTok is the first social media platform that is exponentially growing and is therefore seen as a serious rival to Facebook and Twitter, which are based in Silicon Valley.[88] Originally, the most influential platforms were formed in Silicon Valley, where they have great cultural, economic, and political power regarding managing and controlling platform infrastructure and data. However, TikTok is also enjoying great success internationally, which causes a shift in this power. Because of this, it is argued that TikTok is an interesting platform for this research as it opens a debate whenever its values and policies align with the values of Silicon Valley. Here, Facebook and Twitter were used since these platforms are one of the largest platforms globally, and therefore, as Gillespie argues, these platforms constitute public speech.[89] As a result, this research not only looked at which policies various platforms hold but also looked at how platforms such as TikTok might differentiate in values from U.S. platforms. Since the analysis looked at the values that platforms constitute, the community guidelines were used, as these mostly state the norms and rules and how users

---

[86] Nielsen and Ganter, *The Power of Platforms*. 167.

[87] European Commission, 'Tackling Online Disinformation | Shaping Europe's Digital Future'.

[88] Joanne E. Gray, 'The Geopolitics of "Platforms": The TikTok Challenge', *Internet Policy Review* 10, no. 2 (2021): 2.

[89] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 73, https://doi.org/10.12987/9780300235029-003.

should behave on the platform.[90] However, only looking at the content guidelines for the Ukraine war would not be enough since all three platforms use their transparency centers, newsrooms, or blogs to elaborate on their policies during the Ukraine war. Especially the transparency center is a form of self-regulatory documentation that is made by platforms to increase their transparency on how they curate their content and what repercussions they take if content doesn't meet their guidelines. Therefore, the transparency center and the newsrooms of TikTok, Facebook, and Twitter were used to provide additional information on how these platforms moderate misleading content or content related to the war in Ukraine. Finally, the texts from these sites were placed into one document and categorized into three themes consisting of *the Ukraine war, misinformation, and values*. This resulted in a document in which 13 texts by TikTok, 17 texts by Facebook, and 16 texts by Twitter were used for the analysis, which can be found in *Table 1* below. *Appendix 7.1* provides the exact sources with weblinks.

Used texts	TikTok	Facebook	Twitter
Content guidelines			
Ukraine war			1
Misleading content	1		4
Public values			2
Transparency center			
Ukraine war	1		
Misleading content	2	3	
Public values	7	4	3
Newsroom/Blogs			
Ukraine war	1	5	2
Misleading content	1	5	4
Public values			
Total	13	17	16

Table 1. Used texts for analysis

---

[90] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 46, https://doi.org/10.12987/9780300235029-003.

2.2 Method

As the COVID-19 crisis has already shown, irruptive events reshape platforms' policies. Here, these events reshape the digital realm to such an extent that platforms have adapted their content guidelines to the situation of the pandemic. Therefore, it is crucial to look at how platforms changed their content policies during the Ukraine war. For this research, the method of Scharlach et al. was used as guidance for this analysis since they also studied the duality that platform values amplify through the guideline policies of platforms. Here, Scharlach et al. offer a method that closely looks at content guidelines through texts, which can be identified as a textual analysis. To elaborate, media scholars Marie Gillespie and Jason Toynbee state that a textual analysis analyzes the structure of a media text, as this method illustrates how text produces meaning.[91] This is a suitable method since a textual analysis not only looks at the words that are used but also at 1) the power of text, 2) how media texts construct knowledge, values, and beliefs, and 3) the change and continuity of media systems.[92] As they continue, meanings are not created by media writings alone, which means that words are produced in a certain context in which the audience creates meaning by engaging with the texts.[93] This is also in alignment with what Gillespie says about content guidelines, as platforms often strategically make their guidelines ambiguous and generic.[94] Hence, a textual analysis allows for a better understanding of which platform values the content guidelines of TikTok, Facebook, and Twitter reveals.

Following the method of Scharlach et al., a mixed-method approach was used for this study, which consisted of two parts. First, this part identified and clustered value terms, based on the *values as principles* by sociologist Nathalie Heinich discussed in the theoretical framework. Here, the *values as principles* were identified by their 1) guidance of behavior, 2) positive associations, and 3) applicability to multiple texts.[95] This means that in this part, the *values as principles* were used to reflect the platform values that TikTok, Facebook, and Twitter hold by identifying how platforms guide their judgment regarding what is and what is not allowed in their policies. The next step focused on encoding each document's platform values by identifying its stylistic patterns, recurrent words, themes, and ideas. In doing so, these were marked in the same color to identify different categories. After this, a close reading of all the

---

[91] Marie Gillespie and Jason Toynbee, *Analysing Media Texts*, Understanding Media, [4] (Maidenhead, Berkshire, England ; Open University Press in association with the Open University, 2006), 1. http://catdir.loc.gov/catdir/toc/fy0703/2006298258.html.
[92] Gillespie and Toynbee. 2.
[93] Gillespie and Toynbee. 3.
[94] Gillespie, 'Regulation of and by Platforms'. 264.
[95] Scharlach, Hallinan, and Shifman, 'Governing Principles'. 7.

categories was executed, using the different frameworks and concepts of the theoretical framework to illustrate how these categories relate to the sub-questions. The following shows the three different layers that the analysis is constructed around, and which categories were identified in these layers. Here, these three layers were *misleading content,* the *Ukraine war,* and *self-regulatory practices*. The first layer focused on which generic values are evoked by TikTok, Facebook, and Twitter when regulating misleading content. The role of misleading content is important in the Ukraine war, so it would made sense to first study the policies on misleading content. This revealed which behaviors are currently allowed or prohibited around misleading content by all three platforms. Here, the identified categories in this layer were *safety, community,* and *freedom of speech*. The second layer looked at the positions TikTok, Facebook, and Twitter have regarding the Ukraine war. As not every subject fit into these guidelines, the transparency center, newsrooms, or blogs of TikTok, Facebook, and Twitter were used to give additional information on the Ukraine war since the community guidelines are mostly used for generic rules. The identified categories in this layer were *platforms as a place* and *neutral platform policies*. The last layer focused on the changes these policies have had on the self-regulatory systems of TikTok, Facebook, and Twitter. This helped to answer which platform enforcements have resulted in changes within TikTok, Facebook, and Twitter when combating misleading content during the Ukraine war. Here, a mix of community guidelines, the transparency center, and the newsroom or blogs were used. The identified categories in this third layer were *fact-checking*, *labelling,* and *reducing the visibility of content*. This resulted in an analysis that identified the commercial and public values of TikTok, Facebook, and Twitter and compared them to each other.

3.4 Limitations

While studying platform governance, a lot of information on how platforms operate depends on the information that platforms give themselves. While Gillespie states that most guidelines look similar to each other, he also contends that arbitration is not the main goal of content guidelines.[96] Instead, they serve as a discursive tool to illustrate the protection of online speech while excluding harm and abuse.[97] Because of this, the main goal is not to explore the technical aspects of TikTok but to add them to the analysis as an additional layer, which resulted in the

---

[96] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 47, https://doi.org/10.12987/9780300235029-003.
[97] Gillespie. 47.

creation of a third layer in the analysis. Hence, the focus of this analysis is revealing the constraints and conflicts that TikTok, Facebook, and Twitter encounter while being curators and facilitators of public speech by analyzing platform policies as texts.

3. Analysis

The analysis looks at the duality between commercial and public values of TikTok, Facebook, and Twitter during the Ukraine war. This has been done with three different layers to illustrate the regular policies on misleading content (*3.1 Misleading content*), the relation to the Ukraine war (*3.2 Ukraine war*), and how this has resulted in enforcements and changes within the self-regulatory practices of these platforms (*3.3. Self-regulatory practices*).

3.1 Misleading content

To answer the first sub-question, 'Which public values are evoked by TikTok, Facebook, and Twitter when regulating misleading content?', this section reveals the platform values of TikTok, Facebook, and Twitter in the categories of *harmful content, community,* and *freedom of speech*. Here, TikTok, Facebook, and Twitter all prioritize the value of *safety*, as misleading content harms the platform. However, TikTok differentiates on why misleading content is harmful in comparison to Facebook and Twitter. To illustrate, TikTok's platform values lie mostly in protecting the *integrity* and *authenticity* of the community. However, TikTok explicitly focuses on the implementation of co-regulatory practices to enhance content moderation and enlarge the public values of *responsibility* and *accountability*. Facebook's and Twitter's platform values differ from those of TikTok, as both platforms want to protect *freedom of speech*. As a result, this illustrates the duality of Facebook's and Twitter's the content moderation, as both foster productive dialogue on the platform, which results in moderating content as little as possible. Here, Facebook and Twitter see misleading content as a complex situation in which they focus on informing and contextualizing misleading content rather than removing it.

Harmful content

All platforms define their understanding of misleading content differently, showing room for interpretation in which platforms can decide for themselves what they consider as 'misleading'. Here, TikTok, Facebook, and Twitter all define misleading content ambiguously, which Malson and Maddox introduce as the 'ambiguous moderation line'. This refers to the notion that platforms never formally define what constitutes a violation of their content policies.[98] This

---

[98] Jessica Maddox and Jennifer Malson, 'Guidelines Without Lines, Communities Without Borders: The Marketplace of Ideas and Digital Manifest Destiny in Social Media Platform Policies', *Social Media + Society* 6, no. 2 (1 April 2020): 6, https://doi.org/10.1177/2056305120926622.

ambiguous moderation line has also been identified within the policies of TikTok, Facebook, and Twitter on misleading content. To illustrate, TikTok uses the word 'harmful misinformation' to refer to misleading content, which they define as "content that is inaccurate or false".[99] Twitter defines misleading content as 'misinformation', that "has been confirmed to be false by external, subject-matter experts or include information that is shared deceptively or confusingly."[100] Facebook defines misleading content as "content with a claim that is determined to be false by an authoritative third party."[101] However, these definitions by all platforms lack detail and precision, as there is no clear distinction being made between true and false. However, Facebook acknowledges the duality of misleading content, and therefore, misinformation can't be comprehensively addressed in their community guidelines. In Facebook's words, information changes continuously and rapidly, which means that they can't define what is true or untrue. This means that users can't follow the same guidelines, in contrast with how Facebook defines their policies regarding hate speech or graphic violence, as these violations are more universally outlined.[102] Besides that, Facebook also makes clear that people have varying levels of knowledge about the world, which they mistakenly consider to be true. As a result, they believe that a policy that merely forbids misleading content would be inefficient. However, as TikTok, Facebook, and Twitter don't precisely define what is allowed and what is not, they do encounter misleading content, which threatens their platform values of *safety* due to the harm it causes. TikTok, for instance, removes misleading content "that causes significant harm to individuals, our community, or the larger public regardless of intent".[103] TikTok brings harm in context with civic processes, public health, or safety.[104] Facebook states that it removes misleading content "as it directly contributes to the risk of imminent physical harm."[105] Hence, these policies on misleading content suggest that TikTok, Facebook, and Twitter all articulate the platform value of *safety* when they explain why misleading content is not allowed.

Community

---

[99] TikTok, 'Community Guidelines', 2023, https://www.tiktok.com/community-guidelines?lang=en#37.
[100] Twitter, 'How We Address Misinformation on Twitter', 2023, https://help.twitter.com/en/resources/addressing-misleading-info.
[101] Facebook, 'Misinformation | Transparency Centre', 2023, https://transparency.fb.com/nl-nl/policies/community-standards/misinformation/.
[102] Facebook.
[103] TikTok, 'Community Guidelines'.
[104] TikTok, 'Combating Misinformation', TikTok, 2023, https://www.tiktok.com/transparency/en/combating-misinformation/.
[105] Facebook, 'Misinformation | Transparency Centre'.

Another reason why misleading content is not allowed is that it negatively affects the community. As stated earlier, the policies of platforms are stated in the 'terms of service' and the 'community guidelines'. Here, the community guidelines illustrate the platform's expectations of what is allowed and what is not, explained in often simple language for users to understand its policies.[106] This is done deliberately since these guidelines illustrate what is expected from their users, which are referred to as 'the community'. TikTok, for example, centers its platform value around protecting the *authenticity* of its community. TikTok: "We believe that trust forms the foundation of our community, and we strive to keep TikTok a *safe* and *authentic* space where genuine interactions and content can thrive. We do this by countering misinformation and disinformation, and tackling deceptive behavior that may cause harm to our community or society at large."[107] In another newsroom blog, TikTok states that they empower its community to share their creativity, knowledge, and passion and that misinformation irrupts these activities.[108] Besides that, TikTok also enforces its policies through co-regulatory systems, in which they share their part in the Code of Practice of Disinformation, which is an EU initiative that encounters disinformation by bringing together platforms, industry participants, fact-checkers, and researchers.[109] Through co-regulatory practices, TikTok ensures that it offers *authentic* online experiences for its community. TikTok hereby increases its democratic *accountability* by not only actively combatting misleading information through these commitments themselves but also by insisting on its partners to do the same. These commitments focus on the transparency, integrity, and empowerment of their services, users, and researchers.[110] This implementation looks similar to what Medzini introduced as the framework of 'enhanced self-regulation' to describe the enhancement of self-regulation and the execution of policies by rule-makers of self-regulatory regimes.[111] However, the analysis discovered that TikTok takes this a step further by not only taking *responsibility* to improve their its regulatory practices but also demanding that their audiences responsibly use TikTok to combat misinformation. In contrast, Facebook engages with their 'community' differently as they are trying to create a community that is better educated and to stop the dissemination of

---

[106] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 46, https://doi.org/10.12987/9780300235029-003.
[107] TikTok, 'Combating Misinformation'.
[108] Cormac Keenan, 'An Update on Our Work to Counter Misinformation', Newsroom | TikTok, 28 September 2022, https://newsroom.tiktok.com/en-us/an-update-on-our-work-to-counter-misinformation.
[109] TikTok, 'European Economic Area (EEA) - Code of Practice on Disinformation', TikTok, 2023, https://www.tiktok.com/transparency/en/copd-eu/.
[110] TikTok.
[111] Medzini, 'Enhanced Self-Regulation'. 2229.

false news.[112] For instance, Twitter combats misinformation to educate "the wider community".[113] Here, Facebook and Twitter use the word 'community' not as a value or principle but to name their users. As a result, protecting the community is not seen as their main value when encountering misleading content but rather as part of their end goals in comparison to TikTok.

Freedom of speech

While looking at these guidelines across all three platforms, Facebook and Twitter differ in their platform values in comparison to TikTok. This has to do with the values that are grounded in Silicon Valley. As Gillespie stated earlier, U.S. platform guidelines often want to amplify the freedom of online speech.[114] Maddox and Malson also state this, as most platforms promote themselves as places where people can exercise their right to free speech and urge users to do so.[115] Facebook evokes this value of *free speech* through its choice of words. Here, Facebook seems to be aware that its platform produces discourse. For instance, they don't use the word content but rather use the word 'speech' which resembles the user's core activity.[116] Twitter takes a similar stance on misleading content but rather focuses on providing the right information. Here, Twitter doesn't use the word 'speech' but uses the word 'conversation'. Therefore, they want to support their users so they can engage in a healthy public conversation.[117] Facebook is suggesting the same through its content policies on misinformation, with phrases such as "we focus on reducing its prevalence or creating an environment that fosters a productive dialogue".[118] Facebook also claims that it wants to assist users in retrieving information without suppressing productive public discourse in its policies on false news.[119] As a result, Facebook is taking into consideration to what extent misleading content will not only cause serious harm but also how much it will disrupt the production of speech. Similarly, Twitter's use of regulation also focuses on the users' right to free expression

---

[112] Facebook, 'False News | Transparency Centre', accessed 23 March 2023, https://transparency.fb.com/en-gb/policies/community-standards/false-news/.
[113] Twitter, 'Our Synthetic and Manipulated Media Policy | Twitter Help', accessed 23 March 2023, https://help.twitter.com/en/rules-and-policies/manipulated-media.
[114] Tarleton Gillespie, 'Community Guidelines, or the Sound of No', in *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018), 47, https://doi.org/10.12987/9780300235029-003.
[115] Maddox and Malson, 'Guidelines Without Lines, Communities Without Borders'. 6.
[116] Facebook, 'Misinformation | Transparency Centre'.
[117] Twitter, 'How We Address Misinformation on Twitter'.
[118] Facebook, 'Misinformation | Transparency Centre'.
[119] Facebook, 'False News | Transparency Centre'.

and conversation.[120] Therefore, Twitter is not only aware that it is a facilitator of public conversation, but it also insists on protecting the *freedom of speech* on its platform.[121]

As both Facebook and Twitter amplify the value of *freedom of speech* through their guidelines, they also both embrace the *production of public speech*. Here, social media platforms depend on the 'free marketplace of ideas' to defend themselves and permit the greatest possible quantity of free expression.[122] As a result, both Facebook and Twitter want to limit the removal of content as much as possible. Facebook's aim in regulating misinformation is not only because it causes harm but also because it disrupts the production of speech.[123] The same reason is given by Facebook when looking at the section that covers false news, as false news interrupts public discourse. However, false news will not be removed by the platform. This has to do with the subjectivity of fake news, as it can also be considered satire or opinion. Because of this, Facebook decides to remove harmful misinformation, while also reducing the appearance of false news.[124] Twitter, on the other hand, focuses on informing and contextualizing content rather than limiting or removing it immediately. Back in 2018, Twitter acknowledged that it doesn't have any guidelines for regulating the *authenticity* or *accuracy* of the content.[125] There, Twitter values *openness* which they describe as the way journalists, experts, and other engaged citizens naturally correct and question public debate: "This is important because we cannot distinguish whether every single Tweet from every person is truthful or not. We, as a company, should not be the arbiter of truth."[126] By stating this, both Twitter and Facebook use the *freedom of speech* as a metaphorical sphere in which they position themselves as neutral facilitators.

3.2 The Ukraine war

When looking at the sub-question 'What position do TikTok, Facebook, and Twitter have regarding the Ukraine war while being moderators of public speech?', this layer reveals how all three platforms address the Ukraine war, which is discussed below in the categories *platforms as place* and *neutral platform policies*. By using 'the marketplace of ideas metaphor'

---

[120] Twitter, 'Our Synthetic and Manipulated Media Policy | Twitter Help'.
[121] Twitter, 'How We Address Misinformation on Twitter'.
[122] Maddox and Malson, 'Guidelines Without Lines, Communities Without Borders'. 4.
[123] Facebook, 'False News | Transparency Centre'.
[124] Facebook.
[125] Del Harvey, 'Serving the Public Conversation During Breaking Events', 5 April 2018, https://blog.twitter.com/en_us/topics/company/2018/Serving-the-Public-Conversation-During-Breaking-Events.
[126] Colin Crowell, 'Our Approach to Bots and Misinformation', 14 June 2017, https://blog.twitter.com/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.

for speech, all three platforms legitimize why people use these platforms in times of war with values of *free expression* and the *open internet*. For instance, TikTok presented its platform as a *creative outlet* for its community during the war in Ukraine. Facebook states that its platform is used as a *site of resistance* in which free speech is used as a form of self-defense. Twitter uses a similar metaphor in which it illustrates itself as a *facilitator of information* during the war. However, as Facebook and Twitter claim not to interfere with public speech as much as possible, the Ukraine war challenges this dual frame. Here, during the war, the platform values of Facebook and Twitter regarding *neutrality* and *freedom of speech* interfere with the public value of *trust*. Therefore, as a response to the war, Twitter introduced a new set of policies called 'crisis information policy'.

Platforms as a place

First, all three platforms refer to the war in Ukraine with different words and references. TikTok uses the words 'the war in Ukraine', 'the Ukraine war', or 'during the war'.[127] Facebook uses the words 'Ukraine war' too, however, they also use other references such as 'Putin's invasion of Ukraine', the 'humanitarian crisis', 'the crisis', and the 'war and humanitarian crisis'.[128] Twitter uses similar words to not only refer to the war in Ukraine but also draws this example into a broader context, such as 'during periods of crisis', 'periods of conflict', and 'armed interstate conflict'.[129] Hence, all three platforms illustrate that the Ukraine war is a new form of crisis that is occurring and that it affects the global world. Besides that, TikTok, Facebook, and Twitter create a metaphorical sphere in which they emphasize the need for platforms during the Ukraine war. Here, they limit the showcase of their political opinions and instead focus on how this war reshapes the activities of users on the platform. For instance, TikTok presents its platform as *a creative outlet* for the community during the war in Ukraine, which is in alignment with its platform values. TikTok states that its main challenge is that it can't function as a bridge for people across the globe, so people can't express themselves and share their experiences during the Ukraine war.[130] In addition, TikTok wants to provide a site for creativity and entertainment, which affords relief and human connection during times of war.[131] Facebook, on

[127] TikTok, 'Bringing More Context to Content on TikTok', Newsroom | TikTok, 4 March 2022, https://newsroom.tiktok.com/en-us/bringing-more-context-to-content-on-tiktok.

[128] Facebook, 'Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine', *Meta* (blog), 26 February 2022, https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/.

[129] Sinéad McSweeney, 'Our Ongoing Approach to the War in Ukraine', 16 March 2022, https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine.

[130] Twitter, 'Community Guidelines Enforcement Report', TikTok, 30 June 2022, https://www.tiktok.com/transparency/en-au/community-guidelines-enforcement-2022-1/.

[131] TikTok, 'Bringing More Context to Content on TikTok'.

the other hand, creates a metaphorical sphere where the platform is used as *a form of resistance*. They emphasize how their policies uphold the right to free speech as a form of self-defense in response to an armed attack.[132] However, platforms also reveal a duality in the situation of the Ukraine war, as the war interferes with public values of *free expression* and *open internet*. Here, it is argued that this legitimizes the intervention of platforms with repercussions outside their normal policies. For instance, Facebook recognizes that the Ukraine war should be seen as an exception outside of their regular policies, as Ukrainian citizens use Facebook as a site for resistance and to speak up about the invasion.[133] Facebook: "Our starting point is always to defend people's ability to make their voices heard and to resist attempts to clamp down on the use of our services — especially during times of war and social unrest."[134] Twitter takes a similar stance, in which they consider themselves a facilitator of information during the Ukraine war: "As the conflict – and online conversation – evolves, we want to equip people on Twitter with context and enable informed experiences on the service."[135] Here, Twitter's positioning regarding the Ukraine war is mostly guided by its platform value regarding *freedom of speech*. Here, it emphasizes that everyone should have equal access to information, which can only be enforced through *a free and open internet*.[136] Therefore, Twitter recognizes the importance of context and factual and reliable information to create a healthy and open discourse, which it takes on to facilitate. Therefore, this section reveals that TikTok, Facebook, and Twitter use 'the marketplace of ideas metaphor' by Malson and Maddox for speech again, in which they explain the cruciality of platforms during the Ukraine war. In doing so, it reveals the commercial values with which they convince users to keep on using the platform.

Neutral platform policies

The case of the Ukraine war also reveals the implications of the neutral positioning of these platforms. Although Facebook and Twitter state that they want to intervene as little as possible in public debate, the case of the Ukraine war reveals the implications that come with this when the public value of *trust* interferes with Facebook's and Twitter's platform values of *neutrality* and *freedom of speech*. Facebook accentuates its *neutrality* as a platform when it defends itself against allegations of bias. To illustrate, an official statement from Facebook suggests that

---

[132] Facebook, 'Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine'.
[133] Facebook.
[134] Facebook, 'Protecting Public Debate During the War in Ukraine and Protests in Iran', *Meta* (blog), 23 February 2023, https://about.fb.com/news/2023/02/protecting-public-debate-in-ukraine-and-iran/.
[135] McSweeney, 'Our Ongoing Approach to the War in Ukraine'.
[136] McSweeney.

public *trust* decreases during times of conflict. As a result, it addressed how false assertions were made against the platform as it reviewed content during the war. Here, Facebook explained that these accusations mostly consisted of censoring Ukrainian content and allowing Russian propaganda.[137] In response, the platform accentuated the importance of its community standards guidelines, which are intended to be universally comprehensible for everyone, as Facebook's content reviewers also use these to examine content.[138] In their words: "False assertions designed to undermine *trust* in both public and private institutions are not new and are to be expected during a time of conflict."[139] While Facebook accentuates existing content policies, Twitter, on the other hand, developed a new set of policies in response to the war in Ukraine, as during a crisis, *public trust* is very fragile, as misinformation especially affects already vulnerable communities.[140] In May 2022, Twitter introduced its crisis information policy, in which they state: "In times of crisis — such as situations of armed conflict, public health emergencies, and large-scale natural disasters — false and misleading information has a special capacity to bring harm to vulnerable populations. Here, Twitter defines 'crisis' as "situations in which there is a widespread threat to life, physical safety, health, or basic subsistence."[141] As a result, Twitter states that extending the range of actions that are included in content moderation is a useful way to lessen the damage while maintaining *freedom of speech* and records of critical global events.[142] This new policy is further discussed in the third layer. This suggests that Twitter recognizes the Ukraine war as critical calling for new forms of policies. As a result, Twitter also enhanced its self-regulatory systems, like TikTok, by introducing new policies to encounter misinformation during crises. While TikTok and Facebook also acknowledge the Ukraine war as a the 'humanitarian crisis' or a 'crisis' in general, it is argued that Twitter incorporates the public value of *responsibility* to not only facilitate public conversation but also to actively verify public conversations accuracy, in contrast to what Twitter earlier stated that they "should not be the arbiter of truth". Hence, these new policies show the duality between the public value of *trust* and the platform values of *neutrality* and *freedom of speech*.

---

[137] Facebook, 'Why Claims of Bias in Our Content Review Process Are Wrong', *Meta* (blog), 1 February 2022, https://about.fb.com/news/2023/02/why-claims-of-bias-in-our-content-review-process-are-wrong/.
[138] Facebook.
[139] Facebook.
[140] Yoel Roth, 'Introducing Our Crisis Misinformation Policy', 19 May 2022, https://blog.twitter.com/en_us/topics/company/2022/introducing-our-crisis-misinformation-policy.
[141] Roth.
[142] Roth.

3.3 Self-regulatory practices

The last sub-question answers, 'Which platform enforcements have resulted in changes within TikTok, Facebook, and Twitter when combating misleading content during the Ukraine war?'. This layer reveals to what extent platforms take *responsibility* for the content that is being displayed, in terms of the public values of *transparency* and *accountability*. The main enforcement actions of the platforms have been categorized as *fact-checking, labelling,* and *reducing the visibility of content*. Here, especially the changes in self-regulatory practices during the Ukraine war make it explicit that platforms constitute public speech, although Facebook and Twitter claim otherwise. Especially Facebook and Twitter operate in opaque manners by reducing the *visibility of content,* which causes friction with the public value of *transparency*. However, this layer also reveals the absence of the user's role, which leaves the agency of the user open to question in times of crises and war.

Fact-checking

All three platforms recognized the role that misleading content plays during the Ukraine war. TikTok states that they are aware of the increased danger and consequences of providing misleading information during times of crisis.[143] Similarly, Twitter presents that several media outlets spread false, misleading, or incorrect content, which includes old videos that imitate battles as if they are currently happening in Ukraine.[144] Therefore, Twitter developed its crisis misinformation policy, which was introduced earlier in the analysis. Here, Twitter is taking action to reduce or remove misleading information that has the potential to alter crisis dynamics and target vulnerable people in targeted regions.[145] This includes but is not limited to, false events, allegations, and information.[146] Besides that, TikTok proactively removed 41,191 videos that contained harmful misinformation from February 24 until March 31, as a response to the start of the war.[147] Similarly, during COVID-19, TikTok, Facebook, and Twitter tended to shift away from entirely self-regulatory practices to a more co-regulatory system in which they worked together with independent fact-checkers to evaluate content. For instance, TikTok used mainly the assistance of fact-checking partners that are part of independent organizations to detect misleading war-related content.[148] Here, TikTok states that collaborating with experts

---

[143] TikTok, 'Bringing More Context to Content on TikTok'.
[144] McSweeney, 'Our Ongoing Approach to the War in Ukraine'.
[145] Twitter, 'Crisis Misinformation Policy', May 2022, https://help.twitter.com/en/rules-and-policies/crisis-misinformation.
[146] Twitter.
[147] TikTok, 'Community Guidelines'.
[148] TikTok, 'Bringing More Context to Content on TikTok'.

is key to battling misleading content.[149] Facebook also took significant action to combat false information, removing material from the area that had been fact-checked by outside organizations.[150] Both platforms use partners that speak both the Russian and Ukrainian languages and other dialects to add the needed cultural context to their content revisions. Twitter is less outspoken about the role of their fact-checking partners when regulating misleading content, as they once mentioned this when introducing its crisis misinformation policy: "We require verification from multiple credible, publicly available sources, including evidence from conflict monitoring groups, humanitarian organizations, open-source investigators, journalists, and more."[151] Therefore, this makes it unclear exactly how much of a role Twitter's fact-checking partners play in its crisis information policy. As a result, the role of fact-checking partners seems to be a strategy that has been widely used to encounter misinformation during the Ukraine war. In doing so, all three platforms react proactively to the war in Ukraine by regulating content extensively against misleading content. This suggests that platforms took *responsibility* and recognized their role in delivering accurate information during the Ukraine war.

Labelling

Labelling is also a tool that platforms extensively make use of. This is often done in conjunction with a fact-checker who reviews the content first. If information is judged to be misleading or subjective, the content is then labelled with the appropriate context. All three platforms make use of these labels, especially for state-controlled content. This labelling is different from 'flagging', in which users themselves report harmful content.[152] Here, TikTok adopted a media state policy as a response to the war in Ukraine, labelling 49 Russian state-controlled media accounts to provide context to its community.[153] Facebook also did this by labelling Facebook links and Pages coming from Russian state-controlled media websites so people are more informed before sharing and clicking on them.[154] Twitter expanded its approach by adding labels to tweets that shared Russian state-affiliated media websites. In doing so, they labelled more than 910,175 unique Tweets.[155] Twitter has also labelled one Ukrainian state-affiliated

---

[149] Keenan, 'An Update on Our Work to Counter Misinformation'.
[150] Facebook, 'Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine'.
[151] Roth, 'Introducing Our Crisis Misinformation Policy'.
[152] Kate Crawford and Tarleton Gillespie, 'What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint', *New Media & Society*, 5 August 2014, 411.
[153] TikTok, 'Bringing More Context to Content on TikTok'.
[154] Facebook, 'Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine'.
[155] McSweeney, 'Our Ongoing Approach to the War in Ukraine'.

account, in which they stated: "As the conflict – and online conversation – evolves, we want to equip people on Twitter with context and enable informed experiences on the service."[156] Therefore, this tool is widely used by Facebook and Twitter to give context to misleading content, since they do not have to remove content that contradicts their platform value of *free speech*.

Reducing the visibility of content

Here, the power of platforms is especially applicable when platforms decide which content is made visible and which isn't. While platforms proactively implement new content policies, they also have less transparent enforcement policies that are not visible to the public eye. Facebook and Twitter impose speech restrictions when encountering harmful content related to the Ukraine war. For instance, Facebook not only labels any Russian state-controlled media but also demotes posts containing this type of content.[157] Twitter's approach is based on a two-fold policy for providing reliable information. Here, Twitter still uses its content guidelines, which means that if the content violates Twitter rules, Twitter acts similarly to any other situation. However, Twitter also focuses on decreasing a misleading Tweet's reach, rather than amplifying the content on people's timelines. This is based on situations when the risks of imminent damage are low, but people could still be deceived if left uncontextualized.[158] This includes not only government accounts but also users who limit access to reliable information. While Twitter is less detailed about the role of fact-checkers, they do advocate for the effectiveness of reducing war-related content. Twitter: "We've long believed that content moderation should extend beyond the removal of content or accounts. Not amplifying – meaning the content isn't recommended on Home Timelines and other surfaces – and adding contextual labels quickly reduces content's reach and visibility on Twitter, at scale."[159] In contrast to the use of fact-checking partners and labelling, the analysis shows that reducing the visibility of content requires less accountability for Facebook and Twitter since this policy enforcement is invisible to the regular user, which results in little to no user involvement. Therefore, when looking at these enforcements, it reveals not only to what extent TikTok, Facebook, and Twitter adhere to public values of *responsibility* and *transparency* but also reveals the duality of involving users themselves during the Ukraine war. Here, it is argued that

---

[156] McSweeney.
[157] Facebook, 'Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine'.
[158] McSweeney, 'Our Ongoing Approach to the War in Ukraine'.
[159] McSweeney.

the interaction between different involved parties, such as the platforms, users, and moderators, in relation to platform policies is complicated and operated in a strategic manner that makes it difficult for the public to understand. For instance, the absence of a feature such as flagging is remarkable, as it is an affordance that relies on the attendance of users, which has been used beneficially to regulate numerous pieces of content in fast and changing environments and to legitimize the removal of content on behalf of the community.[160] That, and the other enforcements, suggest little to no user participation in these policies during the war in Ukraine. Therefore, the agency of users during this war is open for debate and leaves questions open about what role users should have to enhance the self-regulatory systems of platforms during times of crises and war.

---

[160] Crawford and Gillespie, 'What Is a Flag For?' 412.

4. Discussion

The thesis examined the dualities that occur between commercial values and public values during times of crisis. Public values such as *transparency, accountability, and responsibility* that have been identified by scholars such as van Dijck et al. and Milosevic are important for this study to examine. As several scholars such as Gillespie, van Dijck et al., and Gorwa contributed to the contention that platforms are not neutral facilitators, it is proven in the analysis that this notion is to still relevant when looking at the policies of Facebook and Twitter. However, the policies illuminate this narrative of *neutrality*, although this has not been explicitly articulated, by identifying the platform value of *freedom of speech* throughout this study. Here, the value of *freedom of speech* has been extensively studied by scholars, as Scharlach et al. also used 'the marketplace for ideas' for speech by Maddox and Malson similarly. Here, these policies reveal how this metaphor is used widely by Facebook and Twitter, which makes the difference in platform values different in comparison to TikTok. Besides that, the thesis identified the same platform values of *community* and *safety* that were also presented by Scharlach et al. Therefore, in alignment with Scharlach et al., it has been argued that platforms use these values to legitimize their policy enforcements, by using these platform values for their private interests.[161] Here, I argue that public values such as *accountability* and *transparency* are not supported especially by Facebook and Twitter, as these were poorly revealed in the policies of these platforms. Although this study confirms many concepts and existing values already, it has been shown that specific public values are considered more important during times of crises and war. As a result, the public value of *trust* has been a value that has not been identified as an often-mentioned public value and has not been identified by Scharlach et al. With that being said, this study shows the need to reveal the duality between commercial and public values by analyzing content policies, which remains a relevant topic for media scholars.

---

[161] Rebecca Scharlach, Blake Hallinan, and Limor Shifman, 'Governing Principles: Articulating Values in Social Media Platform Policies', *New Media & Society*, 7 March 2023, 14 https://doi.org/10.1177/14614448231156580.

5. Conclusion

The study of platform governance emphasizes the need to study regulations *by* platforms further by media scholars. Therefore, this study aimed to fill the gap between existing studies that focus on the regulation *by* platforms, and those that focus on the self-regulatory practices of platforms especially during crises and war. Here, Gillespie shows how regulations *by* platforms can be used as a discursive tool to study platforms as neutral facilitators.[162] While studying this, the positioning of platforms illustrates some implications, in which José van Dijck et al. argue that when private actors such as platforms get intertwined with societal issues, platforms always amplify their values which contest their positioning as neutral-free constructs.[163] As a result, many scholars, such as van Dijck et al. advocate for more *responsibility* and *transparency* on platforms, especially as private actors operate on societal issues. Besides that, Scharlach et al., elaborate on the importance of analyzing public values and platform values through content policies.[164] In this thesis, I have studied this duality, asking in what ways the policies of platforms surrounding misleading content illustrate the duality between commercial and public values during the Ukraine war. In particular, the role of misleading content during the Ukraine war is analyzed in this study, in which not only researchers but also the EU recognize it as a public threat.[165] Hereby, I argue that studying the policies of platforms reveals the dualities in public and commercial values during times of crises and war.

As a result, this thesis is an addition to the existing study by Scharlach et al. In doing so, the Ukraine war has been used as a case study to illustrate the dualities of content moderation of TikTok, Facebook, and Twitter when regulating misleading content. The analysis revealed several findings, as the first layer identified the difference in platform values between TikTok in comparison to Facebook and Twitter. To illustrate, TikTok emphasized protecting the *authenticity* of its community, whereas Facebook's and Twitter's platform values were based on *freedom of speech.* As a result, both Facebook and Twitter foster productive dialogue. This positioning makes the platforms want to regulate content as little as possible which results in a duality in their content policies. The second layer illuminated this duality further, as Facebook's official statements and Twitter's new crisis information policy revealed a conflict between the public value of *trust* and the platform values of *neutrality* and *freedom of speech*. The third layer revealed how TikTok, Facebook, and Twitter took *responsibility* for encountering

---

[162] Tarleton Gillespie, 'The Politics of "Platforms"', *New Media & Society*, 2010, 4, https://doi.org/10.1177/1461444809342738. 7.

[163] Dijck, Poell, and Waal, *The Platform Society*. 3.

[164] Scharlach, Hallinan, and Shifman, 'Governing Principles'. 1.

[165] European Commission, 'Tackling Online Disinformation | Shaping Europe's Digital Future'.

misleading content during the Ukraine war with a new set of policies that consisted of fact-checking, labelling, and reducing the visibility of content. By doing this, all platforms shifted from self-regulatory practices to co-regulatory systems, involving third parties to examine content. However, Facebook's and Twitter's opaque policies of reducing the visibility of content left questions open about *transparency* but also the role of users during times of crises and war. Therefore, this thesis contributes to the need for media scholars to enrich their understanding of regulations *by* platforms. In addition, this study has done this by studying the content policies of TikTok, Facebook, and Twitter, which revealed the duality of commercial and public values during the Ukraine war. With that being said, the limitation on whenever platforms construct their architectural design according to its platform values is acknowledged in this research. Therefore, a recommendation for future research on how public speech is reconstructed through the architectural design of platforms and the role of algorithms would expand this research's interest.

6. References

Barrett, Bridget, and Daniel Kreiss. 'Platform Transience: Changes in Facebook's Policies, Procedures, and Affordances in Global Electoral Politics'. *Internet Policy Review* 8, no. 4 (31 December 2019): 1–22.

Bode, Leticia, and Emily K. Vraga. 'Correction Experiences on Social Media During COVID-19'. *Social Media + Society* 7, no. 2 (1 April 2021): 1–11. https://doi.org/10.1177/20563051211008829.

Burgess, Jean. 'Platform Studies'. In *Creator Culture*, 21–38. New York University Press, 2021. https://doi.org/10.18574/nyu/9781479890118.003.0005.

Chayka, Kyle. 'Ukraine Becomes the World's "First TikTok War"'. *The New Yorker*, 3 March 2022. https://www.newyorker.com/culture/infinite-scroll/watching-the-worlds-first-tiktok-war.

Crawford, Kate, and Tarleton Gillespie. 'What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint'. *New Media & Society*, 5 August 2014, 410–28.

Crowell, Colin. 'Our Approach to Bots and Misinformation', 14 June 2017. https://blog.twitter.com/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.

DeNardis, L., and A. M. Hackl. 'Internet Governance by Social Media Platforms'. *Telecommunications Policy* 39, no. 9 (1 October 2015): 761–70. https://doi.org/10.1016/j.telpol.2015.04.003.

Dieter, Michael, Anne Helmond, Nathaniel Tkacz, Fernando van der Vlist, and Esther Weltevrede. 'Pandemic Platform Governance: Mapping the Global Ecosystem of COVID-19 Response Apps'. *Internet Policy Review* 10, no. 3 (2021): 1–28.

Dijck, José van, Thomas Poell, and Martijn de Waal. *The Platform Society: Public Values in a Connective World*. New York, NY: Oxford University Press, 2018.

Douek, Evelyn. 'Governing Online Speech: From "Posts-As-Trumps" to Proportionality and Probability'. SSRN Scholarly Paper. Rochester, NY, 23 August 2020. https://doi.org/10.2139/ssrn.3679607.

European Commission. 'A Strengthened EU Code of Practice on Disinformation'. Accessed 22 March 2023. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan/strengthened-eu-code-practice-disinformation_en.

———. 'Tackling Online Disinformation | Shaping Europe's Digital Future', 9 February

2023. https://digital-strategy.ec.europa.eu/en/policies/online-disinformation.

Facebook. 'False News | Transparency Centre'. Accessed 23 March 2023.
https://transparency.fb.com/en-gb/policies/community-standards/false-news/.

———. 'Meta's Ongoing Efforts Regarding Russia's Invasion of Ukraine'. *Meta* (blog), 26
February 2022. https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/.

———. 'Misinformation | Transparency Centre', 2023. https://transparency.fb.com/nl-nl/policies/community-standards/misinformation/.

———. 'Protecting Public Debate During the War in Ukraine and Protests in Iran'. *Meta* (blog), 23 February 2023. https://about.fb.com/news/2023/02/protecting-public-debate-in-ukraine-and-iran/.

———. 'Why Claims of Bias in Our Content Review Process Are Wrong'. *Meta* (blog), 1
February 2022. https://about.fb.com/news/2023/02/why-claims-of-bias-in-our-content-review-process-are-wrong/.

Gillespie, Marie, and Jason Toynbee. *Analysing Media Texts*. Understanding Media, [4].
Maidenhead, Berkshire, England ; Open University Press in association with the Open
University, 2006. http://catdir.loc.gov/catdir/toc/fy0703/2006298258.html.

Gillespie, Tarleton. 'Community Guidelines, or the Sound of No'. In *Custodians of the
Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social
Media*, 45–73. New Haven: Yale University Press, 2018.
https://doi.org/10.12987/9780300235029-003.

———. 'Regulation of and by Platforms'. In *The SAGE Handbook of Social Media*, edited by
Jean (Jean Elizabeth) Burgess, Alice Marwick, and Thomas Poell, 1st ed., 254–78. Los
Angeles: SAGE reference, 2018.
https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5151795.

———. 'The Politics of "Platforms"'. *New Media & Society*, 2010, 1–19.
https://doi.org/10.1177/1461444809342738.

Gorwa, Robert. 'What Is Platform Governance?' *Information, Communication & Society* 22,
no. 6 (12 May 2019): 854–71. https://doi.org/10.1080/1369118X.2019.1573914.

Gray, Joanne E. 'The Geopolitics of "Platforms": The TikTok Challenge'. *Internet Policy
Review* 10, no. 2 (2021): 1–26.

Harvey, Del. 'Serving the Public Conversation During Breaking Events', 5 April 2018.
https://blog.twitter.com/en_us/topics/company/2018/Serving-the-Public-Conversation-During-Breaking-Events.

Helberger, Natali, Jo Pierson, and Thomas Poell. 'Governing Online Platforms: From Contested to Cooperative Responsibility'. *The Information Society* 34, no. 1 (1 January 2018): 1–14. https://doi.org/10.1080/01972243.2017.1391913.

Keenan, Cormac. 'An Update on Our Work to Counter Misinformation'. Newsroom | TikTok, 28 September 2022. https://newsroom.tiktok.com/en-us/an-update-on-our-work-to-counter-misinformation.

Maddox, Jessica, and Jennifer Malson. 'Guidelines Without Lines, Communities Without Borders: The Marketplace of Ideas and Digital Manifest Destiny in Social Media Platform Policies'. *Social Media + Society* 6, no. 2 (1 April 2020): 1–10. https://doi.org/10.1177/2056305120926622.

McSweeney, Sinéad. 'Our Ongoing Approach to the War in Ukraine', 16 March 2022. https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine.

Medzini, Rotem. 'Enhanced Self-Regulation: The Case of Facebook's Content Governance'. *New Media & Society* 24, no. 10 (1 October 2022): 2227–51. https://doi.org/10.1177/1461444821989352.

Milosevic, Tijana. 'Social Media Companies' Cyberbullying Policies'. *International Journal of Communication* 10, no. 0 (13 October 2016): 5164–85.

Nekmat, Elmie. 'Nudge Effect of Fact-Check Alerts: Source Influence and Media Skepticism on Sharing of News Misinformation in Social Media'. *Social Media + Society* 6, no. 1 (1 January 2020): 1–14. https://doi.org/10.1177/2056305119897322.

Nielsen, Rasmus Kleis, and Sarah Anne Ganter. *The Power of Platforms: Shaping Media and Society*. 1 online resource vols. New York, NY: Oxford University Press, 2022. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=3224265.

Plantin, Jean-Christophe, Carl Lagoze, Paul N Edwards, and Christian Sandvig. 'Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook'. *New Media & Society* 20, no. 1 (1 January 2018): 293–310. https://doi.org/10.1177/1461444816661553.

Popiel, Pawel. 'Digital Platforms as Policy Actors'. In *Digital Platform Regulation: Global Perspectives on Internet Governance*, edited by Terry Flew and Fiona R. Martin, 131–50. Cham: Palgrave Macmillan, 2022. https://www.amazon.com/Digital-Platform-Regulation-Perspectives-Governance/dp/3030952223.

Roth, Yoel. 'Introducing Our Crisis Misinformation Policy', 19 May 2022.

https://blog.twitter.com/en_us/topics/company/2022/introducing-our-crisis-misinformation-policy.

Sardarizadeh. 'Ukraine War: False TikTok Videos Draw Millions of Views'. *BBC News*, 25 April 2022, sec. Reality Check. https://www.bbc.com/news/60867414.

Scharlach, Rebecca, Blake Hallinan, and Limor Shifman. 'Governing Principles: Articulating Values in Social Media Platform Policies'. *New Media & Society*, 7 March 2023, 1–20. https://doi.org/10.1177/14614448231156580.

TikTok. 'Bringing More Context to Content on TikTok'. Newsroom | TikTok, 4 March 2022. https://newsroom.tiktok.com/en-us/bringing-more-context-to-content-on-tiktok.

———. 'Combating Misinformation'. TikTok, 2023. https://www.tiktok.com/transparency/en/combating-misinformation/.

———. 'Community Guidelines', 2023. https://www.tiktok.com/community-guidelines?lang=en#37.

———. 'European Economic Area (EEA) - Code of Practice on Disinformation'. TikTok, 2023. https://www.tiktok.com/transparency/en/copd-eu/.

Treem, Jeffrey W., Stephanie L. Dailey, Casey S. Pierce, and Diana Biffl. 'What We Are Talking About When We Talk About Social Media: A Framework for Study'. *Sociology Compass* 10, no. 9 (2016): 768–84. https://doi.org/10.1111/soc4.12404.

Twitter. 'Community Guidelines Enforcement Report'. TikTok, 30 June 2022. https://www.tiktok.com/transparency/en-au/community-guidelines-enforcement-2022-1/.

———. 'Crisis Misinformation Policy', May 2022. https://help.twitter.com/en/rules-and-policies/crisis-misinformation.

———. 'How We Address Misinformation on Twitter', 2023. https://help.twitter.com/en/resources/addressing-misleading-info.

———. 'Our Synthetic and Manipulated Media Policy | Twitter Help'. Accessed 23 March 2023. https://help.twitter.com/en/rules-and-policies/manipulated-media.

van, José. 'Governing a Responsible Platform Society'. In *The Platform Society*, edited by José van Dijck, Thomas Poell, and Martijn de Waal, 137–62. Oxford University Press, 2018. https://doi.org/10.1093/oso/9780190889760.003.0008.

7. Appendix

7.1 Documents used for the analysis

TikTok	
Category:	Link:
1. Public values	https://www.tiktok.com/transparency/en/our-commitments/
2. Public values	https://www.tiktok.com/transparency/en/content-moderation/
3. Public values	https://www.tiktok.com/transparency/en/protecting-teens/
4. Public values	https://www.tiktok.com/transparency/en/combatting-hate-violent-extremism/
5. Public values	https://www.tiktok.com/transparency/en/countering-influence-operations/
6. Public values	https://www.tiktok.com/transparency/en/security-privacy/
7. Public values	https://www.tiktok.com/transparency/en/upholding-human-rights/
8. Misinformation	https://www.tiktok.com/transparency/en/combating-misinformation/
9. Misinformation	https://www.tiktok.com/transparency/en/copd-eu/
10. Misinformation	https://newsroom.tiktok.com/en-us/an-update-on-our-work-to-counter-misinformation
11. Misinformation	https://www.tiktok.com/community-guidelines?lang=nl
12. Ukraine war	https://www.tiktok.com/transparency/en-au/community-guidelines-enforcement-2022-1/
13. Ukraine war	https://newsroom.tiktok.com/en-us/bringing-more-context-to-content-on-tiktok

Facebook	
Category:	Link:
1. Public values	https://transparency.fb.com/nl-nl/policies/improving/bringing-local-context
2. Public values	https://transparency.fb.com/nl-nl/policies/improving/prioritizing-content-review
3. Public values	https://transparency.fb.com/nl-nl/features/approach-to-newsworthy-content/
4. Public values	https://transparency.fb.com/nl-nl/enforcement/taking-action/restricting-accounts-by-public-figures/
5. Misinformation	https://about.fb.com/news/2021/03/how-were-tackling-misinformation-across-our-apps/
6. Misinformation	https://about.fb.com/news/2020/03/combating-covid-19-misinformation/
7. Misinformation	https://about.fb.com/news/2022/07/oversight-board-advise-covid-19-misinformation-measures/
8. Misinformation	https://transparency.fb.com/en-gb/features/approach-to-misinformation/
9. Misinformation	https://transparency.fb.com/en-gb/features/approach-to-misinformation
10. Misinformation	https://transparency.fb.com/en-gb/policies/community-standards/misinformation/
11. Misinformation	https://transparency.fb.com/en-gb/policies/community-standards/false-news/
12. Misinformation	https://transparency.fb.com/en-gb/features/approach-to-misinformation
13. Ukraine war	https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/
14. Ukraine war	https://about.fb.com/news/2023/02/why-claims-of-bias-in-our-content-review-process-are-wrong/
15. Ukraine war	https://about.fb.com/news/2022/09/removing-coordinated-inauthentic-behavior-from-china-and-russia/

16. Ukraine war	https://about.fb.com/news/2023/02/protecting-public-debate-in-ukraine-and-iran/
17. Ukraine war	https://about.fb.com/news/2022/02/security-updates-ukraine/

Twitter	
Category:	Link:
1. Public values	https://help.twitter.com/en/rules-and-policies/enforcement-philosophy
2. Public values	https://blog.twitter.com/en_us/topics/company/2022/twitter-2-0-our-continued-commitment-to-the-public-conversation
3. Public values	https://blog.twitter.com/en_us/topics/company/2022/ttr-20
4. Public values	https://help.twitter.com/en/rules-and-policies/defending-and-respecting-our-users-voice
5. Public values	https://transparency.twitter.com/en/about.html
6. Misinformation	https://blog.twitter.com/en_us/topics/company/2020/new-approach-to-synthetic-and-manipulated-media
7. Misinformation	https://help.twitter.com/en/rules-and-policies/twitter-rules
8. Misinformation	https://help.twitter.com/en/rules-and-policies
9. Misinformation	https://blog.twitter.com/official/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.html
10. Misinformation	https://help.twitter.com/en/resources/addressing-misleading-info
11. Misinformation	https://help.twitter.com/nl/rules-and-policies/manipulated-media
12. Ukraine war	https://help.twitter.com/en/rules-and-policies/crisis-misinformation
13. Ukraine war	https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine
14. Ukraine war	https://blog.twitter.com/en_us/topics/company/2022/introducing-our-crisis-misinformation-policy