



**LUND**  
UNIVERSITY

**Virtual Youtuber Kizuna AI:  
Co-creating human-non-human interaction and celebrity-audience  
relationship**

Author: Xin Zhou

MSc in Media and Communication

Supervisor: Deniz Neriman Duru

Examiner: Tobias Linné

May 2020

## **Abstract**

As an easy entry into public attention, YouTube is encouraging individuals to become petty producers under the name of YouTuber, upload self-generated media texts, accumulate following and popularity before finally finding their way to the professional media world. While YouTuber used to refer to human presence, the virtual YouTuber trend starting in 2016 extends this way of becoming famous to the realm of virtual characters and avatars. Similarly, top virtual YouTubers follow their famous human counterparts to have transmedia presence and solid fan bases. As a new combination of human and machine, virtual YouTubers are bringing novel experiences to audiences and pushing the boundary between the real and the imagined.

This thesis aims to study virtual YouTuber Kizuna AI as a specific and emerging type of internet celebrity containing disputing definitions, meanings and values to her producing company, voice models and audiences on Chinese video platform bilibili.com. By conducting semiological analysis and qualitative text analysis, the thesis displays the producers' efforts in constructing Kizuna AI as an adorable and exotic self-claimed AI YouTuber with distinctive visual features from Japanese culture and connotations of both human and non-human qualities. It unveils the dynamics within Kizuna AI's audience group where they have non-unified definitions over Kizuna AI, and engage Affectively, cognitively, morally with Kizuna AI with different intensities and proximities, and further demonstrates Kizuna AI's whole form as a digital commodity involving interrelated issues of economic logics, labour relations, capital and socio-cultural values.

The thesis finds the existence of push-pull relation between Kizuna AI's producers and audiences where audiences are pulled into the non-human characterisation and appeal of Kizuna AI, while they also actively push back through individual and collective fan practices toward dissatisfactory media contents and production decisions. The research shows that virtual YouTuber is a fluid cultural category where the virtual internet celebrity's definition, characteristics, meanings and values are always in flux and constantly being negotiated and balanced between not only audiences and producers, but also audiences and audiences.

**Keywords:** *VTuber; virtual YouTuber; Kizuna AI; internet celebrity; bilibili; YouTube; non-human; para-social interaction; celebrity-audience relation; audience; media engagement; fan; anti-fan; commodity*

## **Acknowledgements**

I would like to start with the gratitude to my supervisor Deniz Duru who has walked me through the whole process with constructive and convincing guidance in this special corona time and special thanks to Annette Hill who has been inspiring and encouraging me to increase the possibilities of my work.

Thanks also to Michael Rübsamen, Zaki Habibi, Joanna Doona, Hario Priambodho for enlightening me at the early stages of my thesis project, and my MKV classmates, especially my virtual study group Yucong and Ruihao, for mutual support.

Thanks to my parents and friends for generous help, accompany and encouragement along the way. I'm especially thankful to Yijun and Kaining for the games and numerous late-night talks, to Jiabei for the discussions and precious recipes, and Haidong for the debates, the songs and literally everything.

## Table of contents

<b>Abstract</b> .....	<b>2</b>
<b>Acknowledgements</b> .....	<b>4</b>
<b>1. Introduction</b> .....	<b>7</b>
1.1 Kizuna AI and the voice model incident .....	8
1.2 Aims and research questions .....	9
<b>2. Literature Review</b> .....	<b>12</b>
2.1 Identifying "online" celebrity.....	12
2.1.1 Microcelebrity - The start of study on online celebrity .....	13
2.1.2 From influencer to internet celebrity .....	14
2.1.3 Virtual character as internet celebrity .....	16
2.2 Researching Kizuna AI .....	18
2.2.1 Contextualizing Kizuna AI .....	18
2.2.1.1 Japanese culture: kawaii, otaku and moe .....	19
2.2.1.2 Platform: YouTube and bilibili.com.....	20
2.2.2 Theoretical framework.....	21
2.2.2.1 From para-social to multi-social interaction .....	22
2.2.2.2 Push-pull dynamic and spectrum of engagement.....	24
2.2.2.3 Fans, anti-fans and non-fans .....	25
2.2.2.4 Celebrity, commodity and value .....	27
<b>3. Methodology and Methods</b> .....	<b>30</b>
3.1 Social constructionism .....	30
3.2 Qualitative approach and case study .....	30
3.3 Multi-method design .....	31
3.4 Conducting the research: piloting, sampling, data collection and analysis .....	33
3.4.1 The pilot.....	33
3.4.2 Sampling .....	33
3.4.3 Collecting and analysing data .....	34
3.4.3.1 Data collection.....	34
3.4.3.2 Data analysis.....	35
3.5 Ethical issues and limitation of the research .....	36
<b>4. Analysis</b> .....	<b>38</b>
4.1 The appeal in Kizuna AI .....	38
4.1.1 Virtuality and Japanese visual heritage .....	38
4.1.2 An exotic and exceptional AI YouTuber .....	41
4.1.3 Performing intimacy .....	43
4.1.4 Authenticity in real-time relation .....	46
4.2 Engaging with Kizuna AI on bilibili.com .....	48
4.2.1 Building an affective space on imagination and fantasy .....	48
4.2.2 Customizing Kizuna AI's character postulation.....	52
4.2.3 Recognizing and acknowledging the voice model.....	53

4.2.4 Pushing back through moral practices .....	55
4.2.5 Displaying consumer confidence and expectations .....	60
4.2.6 Culture as contestation.....	61
4.3 Kizuna AI as a digital commodity.....	64
4.3.1 Commodifying labour, performance and the virtual YouTuber .....	64
4.3.2 Participating in digital capitalism .....	66
<b>5. Conclusion.....</b>	<b>68</b>
5.1 How is Kizuna AI defined and portrayed in her videos? And how does she perform and invite interaction from the audiences? .....	69
5.2 How do audiences engage with Kizuna AI and each other on bilibili.com? And how do they recognize and engage with voice models? .....	71
5.3 How is Kizuna AI commercialized and consumed as a digital commodity? .....	73
5.4 Moving forward .....	75
<b>6. References .....</b>	<b>76</b>
6.1 Empirical material.....	81
<b>7. Appendix .....</b>	<b>83</b>
7.1 Appendix A – Semiological Analysis.....	83
7.1.1 Scheme.....	83
7.1.2 Example of one semiological analysis .....	84
7.1.3 Results of semiological analysis .....	87
7.2 Appendix B – Qualitative text analysis .....	89
7.2.1 Coding process.....	89
7.2.2 Qualitative text analysis results .....	90

## **1. Introduction**

With the rise of YouTube, individuals get the opportunity to become petty producers who make, upload and share self-generated media texts within their YouTube channels (Van Dijck 2013: 113-114). Currently, YouTube may serve as users' source of information, entertainment and connection, as users can associate, inspire and motivate each other within this huge networking platform (Edosomwan el 2011: 83). By virtue of their innovative and impressive creation, some YouTubers gained numerous views and subscriptions, which eventually turned them into microcelebrities, influencers or internet celebrities with their own fan base (eg. Marwick 2016; Jerslev 2016; Abidin 2018).

YouTubers or narrators in YouTube videos used to be mainly human beings. However, in 2016, there came a "virtual YouTuber" or "VTuber" trend where virtual characters and avatars are taking the lead in YouTube videos and live streaming. The VTuber form usually consists of a 3D virtual model featuring an adorable virtual character who is usually female, and a voice model who provides voice performance and sometimes body movements to the character. Although VTubers are doing quite similar things such as video games, daily vlogs, reaction videos and livestreaming, as their human counterparts, this new combination of human and machine is bringing novel experiences to audiences and pushing further the boundary between the real and the imagined (Michel 2018).

A huge industry also came into being with this VTuber wave. According to a Tokyo-based data research company User Local, the number of VTubers topped 4,000 in 2018 (Nagata 2018). Those VTubers are extending their existence out of YouTube videos into different advertisements, TV programs, games and other diverse media. For instance, Tokyo Metropolitan Television Broadcasting Corporation once had a special production called VIRTUAL BUZZ TALK which was specifically designed for virtual YouTubers. VTubers are also extending their presence to other social media and video platforms out of YouTube. For example, in China where YouTube is not available, many VTubers choose to land on local video

sharing platforms such as bilibili<sup>1</sup> and acfun<sup>2</sup>, while those platforms also react actively to the VTuber trend. For instance, bilibili has added a specific VTuber subcategory on their navigation bar to provide a quicker access to relevant contents.

Originating from an interactive social media platform, VTubers see the importance of audience attention, interaction and contribution to their existence. Audiences as consumers, are indispensable in helping define and shape this new form, and co-creating its meaning and symbolic power with producers and institutions (Hill & Turnbull 2017). Moreover, the issue of hidden labour represented by voice models is evident in the VTuber genre, as many of the voice models are devoted in delivering lively performances to and building connections with audiences while audiences are often loyal to the VTubers not necessarily to them. This thesis is looking at Kizuna AI, one of the pathfinders of virtual YouTubers, for more reliable information about the broader class and the VTuber industry which contains abundant information regarding media production, hidden labour and audience power (Flyvbjerg 2001: 66).

### **1.1 Kizuna AI and the voice model incident**

Self-titled “(probably) the world’s first virtual YouTuber”, Kizuna AI is the one who has popularized the VTuber style and the term “virtual YouTuber” since her debut in 2016 despite that there was earlier use of CG avatar in YouTube vlogs<sup>3</sup>. Nowadays, she is one of the top VTubers who can be viewed as celebrities around the world, with a large fanbase of more than 4 million subscriptions on YouTube and more than 1 million followers on bilibili (Kizuna AI’s official distribution channel in China). Because of her popularity, Kizuna AI has also been the spokesperson for SoftBank and the Japan National Tourism Organization, hosted offline fan meetings, participated in music festivals, appeared in commercials and selected to take part in a talk with Nobel Prize winners from Japan (Michel 2018).

---

<sup>1</sup> <https://www.bilibili.com/>

<sup>2</sup> <https://www.acfun.cn/>

<sup>3</sup> YouTuber Ami Yamato started uploading her vlog featuring a virtual character in 2011  
<https://www.youtube.com/channel/UCdDpQ461uxNA3odAnpuigAg>



Her surname Kizuna means connection and bond in Japanese, while her given name AI represents love and a clever reference to her self-claim as an artificial intelligence system. This claim is not a truth but a branding label utilized to add to the virtual YouTuber's uniqueness and attractiveness, for Kizuna AI is equally dependent on voice model's performance as other VTubers. Under this AI claim, the voice model has remained anonymous for more than three years and hidden across various social media as the producing company chooses not to disclose much information of the producing team or the producing process<sup>4</sup>.

Nonetheless, the issue of hidden labour started to trigger people's attention in the summer of 2019 in the voice model incident where the producing company launched a program to duplicate Kizuna AI by recruiting 3 more voice models (one speaks Chinese and two speak Japanese) and trying to get rid of the original one. A lot of fans got very angry and started their resistance by asking people to "unsubscribe her channel if you love her". They refused to recognize the new voice models as Kizuna AI and named them no.2, no.3 and no.4<sup>5</sup>. After several months' push and pull between audiences and producers, the company has now given the Kizuna AI performed by different voice models separate nicknames and identification details such as hairpins of different colors, so that they are not using exactly the same 3D model as they did at the very beginning. This incident further exposed the interrelations between capital, labour, creative value and audience in popular culture.

## **1.2 Aims and research questions**

As an "extreme" case with richer information than an average case, Kizuna AI activates more actors and basic mechanisms regarding the push-pull dynamic between producers and audiences (Flyvbjerg 2001: 78). The primary aim of the thesis is to investigate Kizuna AI as a specific and emerging type of virtual internet celebrity under Abidin (2018: 15-16)'s definition of internet celebrity as "all media formats (people, products, icons, figures, etc.) that attain prominence and popularity native to the internet", and see what are her definition, meanings

---

<sup>4</sup> Original voice model Nozomi Kasuga remained unstated for 3 years until the company's statement on April 24, 2020, to officially acknowledge her existence. [https://twitter.com/aichan\\_nel/status/1253660164961361925](https://twitter.com/aichan_nel/status/1253660164961361925)

<sup>5</sup> As no.4 speaks another language, wears different clothes and has her own bilibili channel separate from Kizuna AI's Aichannel, more dissatisfaction was expressed toward no.2 and no.3 by Chinese audiences.

and values to her producing company, voice models and audiences, and how those different sides negotiate and balance with each other.

In this research, this aim is split into three parts, and each looks at a different aspect toward Kizuna AI. The first aspect focuses primarily on her videos and aims to examine how she defines herself both visually and verbally in this main form of communication with audiences and how she performs and invites social interaction to build connection with the viewers as a celebrity longing for support. In turn, the second part aims at inspecting audiences' individual and collective reaction to her performances and characterizations from the videos, with a specific emphasis on China's media landscape. Meanwhile, as the aforementioned voice model incident has complicated the targets audiences are engaging with, this examination from the audience side also expands to the reactions to voice models. The third aspect is standing at a higher level to look at Kizuna AI's whole commercialization and consumption process which integrates issues of labour, capital and socio-cultural values. This part aims to view audiences as consumers and Kizuna AI as a digital media commodity and explore their interrelations in a new media context. Those three aspects of the research aim are covered and answered through the following research questions:

- How is Kizuna AI defined and portrayed in her videos? And how does she perform and invite interaction from the audiences?
- How do audiences engage with Kizuna AI and each other on bilibili.com? And how do they recognize and engage with voice models?
- How is Kizuna AI commercialized and consumed as a digital commodity?

As the thesis aims to apply and extend the notion of internet celebrity to a virtual existence, it is contributing to both media and celebrity studies. By exploring the conflicting meaning making practices and power dynamics around Kizuna AI, the study is displaying the novelty brought up by a virtual YouTuber's non-humanness and adding to the existing knowledge of virtual internet celebrity as an emerging cultural category.

The following dissertation consists of four parts: 1) a literature review centring on academic literature of internet celebrity, and the theoretical framework containing concepts of para-social

interaction, push-pull dynamic and spectrum of engagement, fan and anti-fan, and commodity to situate the thesis's intellectual position; 2) methodology and methods section focusing on the thesis's social constructionism methodology, qualitative approach, multi-method design of semiological analysis and qualitative text analysis, the according research process and reflection up it; 3) analysis part which starts with the analysis of the videos' connotation and denotations, followed by the examination on audience engagement and fan practices, and ends with an exploration of Kizuna AI's whole format as a digital commodity; 4) a conclusion which blends all the findings together and proposes further research directions.

## **2. Literature Review**

The literature review consists of two main parts: identifying “online” celebrity and researching Kizuna AI. The first part is a review of the literature and research around Kizuna AI. Situating Kizuna AI in the virtual internet celebrity category, the literature review mainly focuses on the development of internet celebrity and fame online. The second part starts with the contextualization of the case study, and is followed by a four-part theoretical framework as the dissertation’s intellectual position within the existing academic research.

### **2.1 Identifying "online" celebrity**

Fame has long existed as a fluid “human condition” which lies in the historical process where famous figures of different periods disclose much about their society and help elucidate the power dynamics between diverse social forces and values across different times (Marwick 2016: 333; Williamson 2016: 2). Celebrity is seen to be inextricably bound up with the emergence, development and extension of contemporary media over the last two centuries (Williamson 2016: 2; Marshall 2016: 15). They prompt a sense of connection and intimacy through their extraordinary personas that have been established and spread through mass media (Rojek 2001, cited in Marwick 2016: 333).

When media communication changes with the arrival of internet and promising digital technology that has led to expansion of media outlets, increasing access to various information and easier approach to content production for the public, celebrity also evolves with drastic changes taking place in the ubiquitous celebrity culture where celebrity is no longer solely connected to mass and broadcast media, but is situated in a more complex media landscape with a widespread increase in commercialization (Williamson 2016: 160; Marwick 2016: 333). More recently, media technologies and interactive platforms such as Twitter, Facebook and YouTube have made it possible for both famous and ordinary people to “generate vast quantities of personal media, manipulate and distribute this content widely, and reach out to (real or imagined) audiences” (Marwick 2016: 333).

### ***2.1.1 Microcelebrity - The start of study on online celebrity***

*Camgirls* (2008), written by Global studies scholar Theresa Senft, is the first academic book on celebrity on the internet (Abidin 2018: xix). The author documented her ethnography study on a generation of “young American women who acquired internet fame by broadcasting their personal lives via webcams in their bedrooms” (Ibid.). Senft (2008: 25) called this new way of celebrity pursuit “micro-celebrity” where users gather popularity through their online performance with the help of digital media technologies and networking platforms. The camgirls publicize themselves by making their self-presentations into a brand and engaging with audiences through “emotional labour” (ibid.: 8-9). Compared with traditional celebrities or stars who often enjoy far-reaching fame and attention from an extensive range of audiences out of expertise and accomplishments but constantly entail a sense of distance and separation from those audiences, microcelebrities are usually known to only a niche audience with narrower but far deeper popularity in that their fame is built on the feelings of connection and intimacy in the interactive relations with the audiences (Senft 2008, cited in Abidin 2018: 11-12). Furthermore, microcelebrities have greater commitment to their audiences than traditional celebrities, as their fame is “co-constructed through a community of interested viewers on the internet rather than by the mere mechanisms of the traditional entertainment industry” (Abidin 2018: 12).

Alice Marwick (2013) further developed the concept of microcelebrity in camgirls who broadcasted themselves as a hobby through her study on Silicon Valley tech entrepreneurs who used social media for the networking and business purposes in the early 2000s (ibid.). She stressed the significance of authentic and interactive feelings in a microcelebrity’s creation of his/her persona to produce contents or strategically appeal to online audiences regardless of their size or state (Marwick 2013: 114). This ongoing sharing and livestreaming of personal information as self-branding requires a mindset that all the friends and followers online could be one’s potential fan base, which Marwick asserted, can lead to two forms of microcelebrity: ascribed and achieved microcelebrity (Marwick 2013: 15-16, cited in Abidin 2018: 12). Ascribed microcelebrities are those who “constantly posit themselves as being of a higher status than their followers to garner celebrity treatment” while achieved microcelebrities “build their

fame by selectively revealing confidential information to viewers to cultivate feelings of intimacy with them” (Marwick 2013: 116-117, 121-123, 208, cited in Abidin 2018: 12).

Since then, studies on microcelebrity started to thrive and vary across internet users all over the world. Scholars are theorizing and researching microcelebrity as labour (eg. Duffy 2016; Abidin 2016; Lana 2019), branding (eg. Booth & Matic 2011), linguistic practice (eg. Page 2012), and activism (eg. Aziz 2019, Brown & Phifer 2019). Their lens of assessing microcelebrity are also shifting and evolving. Apart from the focus on certain groups of people or particular practices, studies are also investigating community norms (eg. Garcia-Rapp & Roca-Cuberes 2017), and affordances and algorithms of platforms (eg. Gibbs et al 2015).

### ***2.1.2 From influencer to internet celebrity***

Diverse internet celebrity innovations have been inaugurated by this “era of the most commercialized form of microcelebrity” (Abidin 2018: 13). “Influencers” came into existence among all the ramifications as “a full-time vocational job” of being a microcelebrity (ibid.). Abidin (2018: 71-72) refers to influencers as “vocational, sustained, and highly branded social media stars” and “a practice focused on social media-based, multimedia, fame on the internet”, and distinguishes it from the Influencer/mediator in business research. Being a critical form of online celebrity, influencers are able to appeal to, engage with and sustain a considerable following across their social media networks by amplifying attractive messages through personalized content production (ibid.: 71).

There are normalized and conventional tactics and strategies within the “shadow economies” of influencers, such as bot followers, hash tag spams and self-amplification groups to generate more traffic within the attention economy (ibid.: 78-82). Moreover, a transformation from attention economy to affection economy where influencers sell not only products but also the lifestyles they are displaying for emulation is emerging with another shift of focus from quantitative metrics involving the numbers of followers, subscriptions, likes and comments to qualitative impacts such as informal testimonials and networked word-of-mouth to enhance audience engagement (ibid.: 94-97).

Based on the studies of influencers and microcelebrities, Abidin (2018) gave a general definition of the rapidly expanding and transforming internet celebrity as “all media formats (people, products, icons, figures, etc.) that attain prominence and popularity native to the internet” with possible “spillover effects and afterlives” involving “cross-border flows outside of the internet”(15-16), and categorized aforementioned influencers as a specific form of internet celebrity (97). She pointed out the higher visibility of internet celebrity as a structural difference compared with microcelebrity and highlighted the significance of audience in defining an internet celebrity since someone who only upload quality content online but is not noticed or endorsed by others won’t be counted as an internet celebrity (ibid.: 15-16). As “a product of performance and perception”, internet celebrity can be created when deliberately performed by a subject, or may come into being when unknowingly perceived by an audience (ibid.: 19).

Drawing on French sociologist Pierre Bourdieu (1986)’s notion of “capital” which explains “how society comes to exercise judgments in taste, express aesthetic dispositions, and assign status and distinction towards different social behaviors”, Abidin (2018) proposed four most distinctive qualities of internet celebrities: exclusivity, exoticism, exceptionalism and everydayness which connect respectively with economic capital, cultural capital, technical capital and social capital (Bourdieu 1986: 241-258, cited in Abidin 2018: 19; Abidin 2018: 20-33). Exclusivity refers to the rare, privileged and glamorous practices or properties that are not usually seen by the public (Abidin 2018: 20-22). Exoticism relates to “distancing, novel or foreign attributes” that appear when discrepancy exists between the performer’s and audience’s cultural norms or social conventions, and viewers entail perceptions based on their own identity categories (ibid.: 22-23). Exceptionalism highlights the unusual and astonishing qualities in a performance that can either come from elite professionals with expert skills or mundane people’s practices as long as the performance is impressive enough to be admired under the internet’s attention economy (ibid.: 28-29). Everydayness often concerns the regularity and consistency in ordinary daily routine or lifestyle where audiences and performers are likely to cultivate and maintain a sense of community and trust (ibid.: 32-33).

### ***2.1.3 Virtual character as internet celebrity***

Aforementioned discussions on internet celebrity centre on human beings, however, a transgressing category of celebrity where virtual characters take the lead is also emerging with the advent of internet and social media and the development of technology. Based on Abidin(2018)'s definition of internet celebrity, we might temporarily call it "virtual internet celebrity" where different yet overlapping subcategories such as virtual idol, virtual influencer and VTuber exist with similar virtual character traits and distinctive features such as pop singer, cultural icon<sup>6</sup>, Instagramer, vlogger and gamer.

Hatsune Miku might be one of the most studied virtual idols. She first appeared in 2007 as the visual identity of a vocal synthesizer software product of a Japanese company Crypton Future Media (Guga 2015: 37). Thanks to the internet and enormous fan activities, she is then spread widely across the globe and has gained a sizeable fan base both domestically and internationally (Ishita 2016: 151). She is not a human but can sing, dance, perform in commercials, live concepts and world tours through a hologram projection (Zaborowski 2016: 112). According to Guga (2015), her great success lies in the fact that her visual model is open to the public and ready for "further improvements, adjustments, and revisualizations according to one's own needs, desires, preferences, and ideas" (40). In this unconventional mode of audience-entertainer interaction, audiences are "prosumers" who both produce and consume through the process of creating, adjusting and personalizing media contents such as mash-ups and remixes (Black 2012: 216). Fans as prosumers are not disturbed by Miku's virtual body but instead fill it with various identities and identification (Guga 2015: 39). To them, Miku is a "real entity" whose lack of a physical body can actually "enhance the emotional engagements on offer during the audiencehood" (Zaborowski 2016: 112). Additionally, the popularization of Hatsune Miku also ties closely to the cultural meanings behind that prosumer format, as Black (2012: 216) has pointed out "of particular relevance to such a relationship between idol and consumer is the

---

<sup>6</sup> "Lil Miquela, whose persona as a Brazilian-American influencer was carefully constructed to draw from the creativity and cultural capital of Black and Latino youth online" (Marwick 2019: 166)



connection between idols (virtual and otherwise) and Japanese otaku” who are often recognized as enthusiastic consumers and participants of certain culture.

Another famous virtual internet celebrity Lil Miquela or Miquela Sousa, is called “virtual influencer” by Marwick (2018: 161) and “fake influencer” by Hubble (2018: 20). She is not a person, but the creation of a Los Angeles company called Brud (Marwick 2019: 161). In her case, the producing agency is building on a plot similar to cyberpunk films where Miquela is “the self-aware android rebelling against her creators (with a feminist and racially aware twist)” (ibid.:162). There are good reasons for Miquela to be named after “influencer”. The first and foremost is that her social media presence is as economically driven as a human influencer’s and demonstrates the more widely available “potential to convert social capital to economic capital” because of her virtual traits (ibid.: 162). Having more than two million followers<sup>7</sup>, her Instagram account @lilmiquela is an ongoing media flow updating her “life” of flexing with luxury and designer brands, producing music, mingling with other (virtual and human) influencers and celebrities, going to supermarkets and being in a relationship (ibid.: 161).

A few researchers have studied Kizuna AI or VTuber from social, cultural and media perspectives within specific regions in Asia (Cho 2019; Korenaga 2019; Puspitaningrum & Prasetio 2019; Guo 2019). Cho (2019) did the first academic research on VTuber in South Korea to examine audience’s different responses towards VTuber and human YouTubers, and generated basic data for VTuber contents marketing. Korenaga (2019) situated his research in Japan and studied criticism on gender representation of VTuber. He explored “the problem of sociologists criticizing society via the criticism of media images” and discussed “the possibility of practical social criticism via the consideration of media representation”. Puspitaningrum and Prasetio (2019) studied “Kizuna AI’s ‘virtual Youtuber’ phenomenon among Japanese Popular Culture Enthusiasts in Indonesia”, and used phenomenology method to consider “motives and meanings in watching Kizuna Ai’s vlogs”.

---

<sup>7</sup> <https://www.instagram.com/lilmiquela/>

In China's media landscape, the one existing research (Guo 2019) examined Kizuna AI as a virtual vlogger and discussed the audience (especially the youth) interaction happening in the case of Kizuna AI together with other two virtual idols: Hatsune Miku and Luo Tianyi (the first vocaloid<sup>8</sup> character designed for Chinese audience). The author concluded that audiences who worship virtual idol find a way to hide the self within the boundary between the real and the virtual and thus get a sense of achievement and accompany (ibid.). S/he argued further that there is an ontological consistency between the real world and the production and reception of two dimensional/virtual idols (ibid.). However, the production processes of VTubers and virtual idols are totally different. Virtual idols, especially vocaliod-related ones, are de signed as voice synthesizers to trigger user production, while VTubers don't have such united production and consumption. In this specific generic form, producers are still distinct from and sometimes even opposing to audiences. Therefore, there is definitely a need to research VTubers independently from vocaloid idols in China.

## **2.2 Researching Kizuna AI**

The thesis answers the aforementioned need and focuses mainly on Kizuna AI in a Chinese context to add to the knowledge of traits that are unique in a VTuber and the mutually enforcing power from the producers and audiences in the VTuber industry through the multi-method design covering both the production and reception side. The specific locality of the thesis contributes to existing knowledge on Kizuna AI's global media representation and reception because of the region-specific language, platforms and media consumption habits. The dissertation's inclusion of virtual character as the main focus extends the concepts in its theoretical framework to a wider realm of human-machine co-existence and brings in virtuality and non-humanness into the discussion to add to those concepts' existing applicability.

### **2.2.1 Contextualizing Kizuna AI**

As the labour issue has already been stated, here the case contextualization mainly includes two parts: the Japanese cultural attributes that Kizuna AI is born with and is currently dwelling in,

---

<sup>8</sup> Vocaloid is a singing voice synthesizer software. By inputting tones and lyrics, users can generate a synthesized song featuring human voices. Hatsune Miku is one of the most famous vocaloid characters.

and the features of her official distribution platforms: YouTube for global audiences and bilibili.com for Chinese audiences. In her case, the agency, interactivity and participatory engagement from the users, audiences and fans are prominent and key both from a cultural background perspective and a platform aspect.

### *2.2.1.1 Japanese culture: kawaii, otaku and moe*

Originating from Japan, Kizuna AI bears obvious aesthetic styles rooted in kawaii culture and otaku culture in Japan, both of which, especially the latter, contain plentiful fannish<sup>9</sup> cultural practices and passionate and participatory engagement (Ito 2012: Xi). Kawaii<sup>10</sup> culture and aesthetics are considered “a peculiarity of contemporary Japan” which appears in media representations and materializes across products, creative industries and juvenile trends (Pellitteri 2018: 94). Kawaii is a Japanese adjective whose meanings vary upon the context and topic where it appears (ibid.: 96). In contemporary Japanese, this adjective derives from the noun *kawaisa* which means sweetness or nicety, and now obtains “relatively agreed-upon range of meanings that go from ‘cute’ to ‘sweet’, from ‘tender’ to ‘childish’, from ‘innocent’ and ‘gentle’ to ‘honest’ and ‘soft’, and from ‘small’ to ‘lovely’ ” (Lieber Milo 2017, cited in Pellitteri 2018:96; Pellitteri 2018:96). Kawaii culture and aesthetics have been studied as a contemporary phenomenon related to youth subculture and particularly female teenagers (Kinsella 1995, cited in Pellitteri 2018:96).

Originating in the early 1980s, otaku first appeared as “a polite term of address between upper-crust college students who were fans of emergent anime cultures” (Ito 2012: xxi). Later the word has been made into “a stigmatizing label for somebody who is obsessed with anime, manga, and games and out of touch with everyday social reality” and now can also refer to “innovative popular-culture enthusiasts” (ibid.: xxi). Hence, otaku culture inevitably involves a variety of fannish cultural logics, platforms and practices across multiple genres including anime, comics, manga and games where a tendency to passionately participate and engage with

---

<sup>9</sup> Fannish means “of or relating to a fan; characteristic of fandom” (Wiktionary)

<sup>10</sup> Kawaii is often translated to “cute” in English. Since Kawaii covers more meanings than just “cute”, the thesis will use the original word Kawaii instead of its English translation.

popular culture and technology in a networked world is evident (ibid.: Xi).

One of the approaches for a media product to attract and engage with those otaku consumers is to embed moe-elements (moe yōso) and provoke the feeling of moe. Moe is a term said to have emerged in the late 1980s with original reference to “the fictional desire for characters of comics, anime, and games or for pop idols” (Azuma 2009: 47). Being very affective, moe has been described by Toru Honda (2005) as romance inside one’s head (cited in Tsuji 2012: 28). Moe-elements are the components designed to “effectively stimulate the moe of the consumers” (Azuma 2009: 47). Being mainly visual, moe-elements can include cat ears, glasses, business suits and other visual signs, as well as characteristics such as a specific way of speaking, settings and stereo typical narrative development (Ito 2012: Xxii; Azuma 2009: 47). There is also a kind of behaviour “chara-moe” to stand for the feeling of moe or emotional attachment toward characters and their appealing attributes (Azuma 2009: 42; Kijima 2012: 250).

With this origin, audiences and fan practices become one of the key components to complete this virtual YouTuber. This feature also corresponds with the affordance of her interactive distribution platforms: YouTube for global audiences and bilibili.com for Chinese audiences. Audiences and fans are empowered and encouraged by those digital social media to “post, repost, remix, and embed content in their commentary or other para-texts, thus constructing, amending, or reinforcing the boundaries of their fan object”, while being limited and influenced by certain algorithms behind the platforms’ money making mechanisms (Sandvoss, Gray & Harrington 2017: 18). At the same time, a virtual internet celebrity’s practices are equally inspired and restricted by the affordance and norms of the digital media platforms where audiences are often converted into metrics and users are encouraged to compete for attention (Marwick 2019: 163).

#### *2.2.1.2 Platform: YouTube and bilibili.com*

Under the name of YouTuber, Kizuna AI’s ways of producing and publicizing her self-presentations are intimately bound up with the YouTube platform on which she became famous, since it decides the length, frame, format of her videos and the way Kizuna AI is developed into

a celebrity (ibid.: 164). YouTube is known for having users upload “pitted user-generated content” and its novelty of being a social networking platform and enabling two-way communications via the internet under the domination of one-way broadcasting or narrowcasting distribution (Van Dijck 2013: 111, 114-115). According to Van Dijck (2013: 117-118, 121), YouTube bears the myth of being “an unmediated gateway to the professional media world of stars and fame” where amateur performers are introduced and known to the public, and musicians or artists supported by big labels or brands acquire enough followers before they show up in mainstream media. Kizuna AI’s development path makes no difference. Supported by a professional company instead of an amateur individual, she has gained quite plenty of followers and attention on YouTube before having transmedia presence in commercials, mobile games and television programs.

Bilibili.com is an integrated Chinese video platform which contains plenty of user-generated contents covering different genres such as anime, movies, music and dancing (Zhang 2020: 20-21). This interactive video platform is also a place where a lot of users mobilize themselves and manoeuvre around contents on platforms not available in China. It enjoys rapid increase in its popularity, especially among young Internet users: 78% of bilibili users are between 18 and 35 (Bilibili.com 2019). The platform is originally “a video-sharing site for Japanese ACG (animations, comics, and games) content” (Zhang 2020: 20-22). This root in ACG culture makes bilibili users enjoy the sharing and partaking of insider knowledge similar to “otaku’s embrace of social connection” on 2ch(5ch) (Ito 2012: Xxii). As is evident in Kizuna AI’s video comments, those who comment on Kizuna AI’s videos often refer to other ACG characters or VTubers and play with commonly accepted features by fans among each other to indicate and consolidate their identity as an insider.

### ***2.2.2 Theoretical framework***

Realizing the significance of audiences and fans and their practices in co-creating virtual YouTuber as a media form and cultural category, I’m integrating approaches from celebrity, media audience and fan studies in this thesis to examine the complexities and disputes within

this cultural phenomenon and add to the understanding of virtual internet celebrity in existing academic research.

#### *2.2.2.1 From para-social to multi-social interaction*

Para-social relations put forward by psychologists Donald Horton and R. Richard Wohl (1956) is one of the most studied celebrity-audience relationships (Abidin 2018: 8). They called the “seeming face-to-face relationship between spectator and performer” in mass communication a para-social relationship (Horton and Wohl 1956: 215). Within this relationship, the performers in television or radio tailor their performances based on the assumption of audience response and often face the audiences while utilizing talking styles and tactics such as direct address to deliver the feeling of being involved in a personal and private conversation (ibid.: 215). Audiences in turn voluntarily and independently decide their “answering role” to this “illusion of face-to-face relationship with the performer” where effective reciprocity is actually absent (ibid.: 215). Horton and Wohl (1956) named this whole style of conversational give and take “para-social interaction” (215).

Within their mass media context, a new type of performer is created as “a special category of ‘personalities’ whose existence is a function of the media themselves” (ibid.: 216). Those personalities were named “personae(persona)” as a kind of media presence only available for the audiences in para-social relation (ibid.: 216). Personae can construct and maintain an intimacy with audiences who are literally strangers, and this intimacy, according to Horton and Wohl (1956), “is extremely influential with, and satisfying for, the great numbers who willingly receive it and share in it” (216).

Fantasy and imagination instead of obligation or responsibility are what stimulate an audience to remain involved in the para-social relation (ibid.: 215). Audiences think they “know” a persona in a similar way of knowing a real-life friend as they are directly observing and interpreting the persona’s appearance, actions, words under different conditions (ibid.: 216). Furthermore, those audience are specifically invited by the “designed informality” from the persona or program to imagine that they are not merely trapped in passive observation but are

involved in face-to-face interactions (ibid.: 216).

With the advent of new media platforms, communication space and technology, it's necessary to contextualize the "one-sided" and "nondialectical" idea of para-social interaction between media persona and audiences within the new media setting (ibid.; Marshall 2010: 43; Hills 2016: 464). Horton and Wohl's idea of para-social interaction as characteristically "controlled by the performer, and not susceptible of mutual development" is gradually being extended to the new context of social media where celebrities are required by the openness and interactivity of media platforms to perform emotional labour and build intensified connection to their fans rather than establish a cultural notion that the celebrity is distant from and opposite to the fans (Horton and Wohl 1956: 215; Marshall 2010: 44; Marwick and boyd 2011, cited in Hills 2016: 474). Click, Lee and Holladay (2013) interpreted social media's impact on para-social relations from a fan perspective and stated that social media are expanding the possible interactivity between celebrity and fans, and pushing the imaginary relationship to a new, immediate and more real form (376-377).

Matt Hills further proposed a concept of multi-social relation upon the original para-social interaction theory (Hills 2016). He argued that the involvement of fantasy and non-unified selves has decided the already "para-" quality in social relations (ibid.: 472). Within temporary media culture, para-social relationships between celebrity and fans are no longer binary but are multiply performed, experienced and participated within the community of digital fandom (ibid.: 472, 479). It becomes multi-social where fans exhibit, share and discuss the para-social interactions instead of fantasizing alone about the dyadic communication between the celebrity and him/herself (ibid.: 479). The multi-social communication highlights fan-fan interactions in constituting specific fan-cultural discourses and authenticities (ibid.). Situated in a digital social media environment, Kizuna AI is extending the theory of para-social interaction to a new context where multi-social interaction is enabled by reception side's bigger agency and power and co-exists with para-social interaction in the celebrity-audience relation. This co-existence serves as the guideline of the thesis to investigate the construction of intimacy and interaction between a virtual and nonhuman celebrity and her human followers.

#### *2.2.2.2 Push-pull dynamic and spectrum of engagement*

The concept of push-pull dynamics is an integral approach to theorize and analyze both the production practices and audience engagement (Hill 2019: 4). Producers and audiences are mutually enforcing each other in the push and pull relation (ibid.). There can be a push from the production side to allure audiences into contents through distribution or branding strategies, and audiences can also be pulled into the specific storytelling, characterization or narratives of media contents (Hill 2016: 754-755; 2019: 4). Meanwhile, restless audiences, users and consumers are also able to push back through fan activities, disengagement practices or informal viewing (Hill 2016: 755; 2019: 4). The push-pull dynamics is described by Hill (2019) as “complicated power relations in the transactions between media industries and audiences” (4). And it is the specific power relation embodied in the push and pull between producers and audiences that shows their struggle of making sense of the transnational media (ibid.: 5).

Media engagement is another analytical lens to explore the connectedness between the production practices and audience research (ibid.: 6). A closer look at media engagement enables the possibility to discuss industry definitions and rethink about the dominant audience metrics and evaluation methods (ibid.: 6). It also unveils how audiences actually interact with a variety of media by formal or informal means and helps researchers discover more about audiences’ multi-faceted ways to get involved in popular culture (ibid.: 6). Hill (2017)’s idea of a spectrum of engagement is a multilayered framework to apprehend the multidimensionality and dynamics of audience engagement while displaying its different positions and intensities (1, 7). It delineates the integral and multiform traits of engagement and stresses the diversified ways of experiencing digital media production (Hill 2017: 7; 2019: 6). A spectrum includes emotional and critical modes where people switch and shift between positive and negative engagement, or disengagement (Hill 2017:8; 2019:7). She draws attention to the value and meaning of media engagement which is produced from the dynamic relationship between the audiences and the producers (Hill 2019:7).

We could further observe that the soft power of audiences highlighted by Hill(2017, 2019) in



her researches of drama and reality TV also appears or even amplifies with the arise of internet and social media as audiences get bigger agency in archiving, scheduling and sharing media contents that used to be strictly controlled by producers such as broadcasting companies. Using the push-pull dynamic and spectrum of engagement can unveil the power of audiences and provide useful criteria to supplement the research on audience side which hasn't received as much attention as the celebrity side in the para-social interaction theory. The spectrum of engagement is a comprehensive approach to examine audiences' affective reaction to and critical consideration of things happening around Kizuna AI regarding her definition, values and meanings.

### *2.2.2.3 Fans, anti-fans and non-fans*

Fandom is defined by Sandvoss (2005: 8) as “the regular, emotionally involved consumption of a given popular narrative or text in the form of books, television shows, films or music as well as popular texts in a broader sense such as sports teams and popular icons and stars ranging from athletes and musicians to actors”. The fans who practice it are usually associated with emotional input or “affect” (Grossberg 1992, cited in Sandvoss 2005: 6). Despite that they are the most prominent and dominant among a celebrity's followers, fans are not the only group of audiences within the complex celebrity-audience relationship. Plenty of opposite reactions such as indifference, hate and dislike, which are as potentially powerful as like, also coexist on the reception side in other kinds of celebrity followers such as anti-fans and non-fans<sup>11</sup> (Gray 2003; 2005). Gray (2003) defined anti-fans as “those who strongly dislike a given text or genre, considering it inane, stupid, morally bankrupt and/or aesthetic drivel (70)” and referred to non-fans as “viewers or readers who do view or read a text, but not with any intense involvement (74)”.

Similar to other cultural categories, there are no clear boundaries between fan, anti-fan and non-fan and their relevant practices. Anti-fans discuss, behave and perform in an opposed yet somewhat similar way to fans, contesting the idea that they are two pure and polar opposites

---

<sup>11</sup> Gray's research largely lied in television audiences who are viewers, but it's not difficult to detect that many of his findings are also applicable to internet celebrities and their audiences.

(Gray 2005: 840-841, 845). Fans can shift to anti-fans of certain texts when they consider certain media contents or components harmful to the whole fan object (Gray 2003: 73). At the same time, the boundary between non-fans and fans are blurring when many lax and self-deemed fans don't actually bear the discipline fans usually embrace, but adopt more open and ambiguous practices such as watching a program whenever they can or want rather than they must (Gray 2003: 74).

One way to investigate the dynamic and changeable audiencehood involving the fan, the non-fan, and the anti-fan is through examination of audience engagement with media text and para-texts. Para-texts are "semi-textual fragments that surround and position the work" (Gray 2003: 72). They can include "introductory sequences, 'spoilers' and ads, newspaper, magazine or web articles, merchandise and the very buzz or media talk that surrounds any given programme" in television programs (ibid.: 72), and would definitely involve numerous social media messages in the case of an internet celebrity. Such examination highlights the diverse affective relationships audiences have "in different proximities to, understandings of, and engagements with" certain media text (Gray 2005: 842).

One useful approach to look at media text and para-texts is to "hypothesize that all texts have moral, rational-realist, and aesthetic dimensions" (ibid.: 844). In positive engagement, fans can merge all the dimensions and absorb them as one (ibid.). At the same time, anti-fandom may prove the three dimensions' potential incompatibility with each other when the consumer is reluctant or incapable of engaging with a text from all the levels (ibid.). In Barker et al. (2001)'s research on *Crash*, they found viewers refused to or were made unable to look at the film's aesthetics out of concerns of morality or realism (cited in Gray 2005: 844). Similarly, in Gray (2005)'s research on *The Apprentice* viewers, fans were willingly engaging with the reality show as an aesthetic text until program participant Omarosa breached this experience and many fans' ability to relish in its aesthetics (850).

The failure and refusal of experiencing the three dimensions of texts also reflects the expectations in fans, anti-fans or nonfans and values that structure media consumption (Gray

2003: 73). Those expectations could include ideal forms of a media text and the morality or aesthetics it holds as well as negative examples to be avoided and amended (ibid.: 73). Closer examination of such textual dislike could illuminate the existing ideas of text, industry and consumption (Gray 2005: 841). Studying the discussions and interrelations of fan, anti-fan and non-fan provides a means to better explain and understand how the talk and exchange feed back into the text and impact its meanings and contents (Gray 2003: 78). Adopting the notion of fan, anti-fan and non-fan echoes with the dynamics existing within Kizuna AI's audiences. Acknowledging their differences in the level of engagement and proximity towards various fan objects and texts under the same name Kizuna AI helps illustrate the personalization and contestation regarding meanings and textual consumption within this specific virtual internet celebrity category.

#### *2.2.2.4 Celebrity, commodity and value*

There is an inextricable connection between celebrity and capital in the modern age (Horkheimer & Adorno 1999, cited in Senft 2008: 25). Celebrities grow mainly for economic reasons and can be considered one predominant way to sell media commodities (Williamson 2016: 13, 21-22). They have been partaking in the affluence of important media sectors such as print, news, cinema, television and the internet (ibid.: 22). In turn, those media organizations are developing and spreading celebrity-related coverage and contents for their own profit under specific economic circumstances (ibid.:162). More crucially, celebrity culture has played a key role in the pedagogy of consumption where audiences construct the fluid identity or remake the self through purchasing and consuming cultural commodities operated under a celebrity's name (Marshall 2014; Williamson 2016: 162).

Social media platforms provide a new space for "the celebrification of the self" (Williamson 2016: 24). According to Williamson (2016), the basic value of cultural production lies in "sellability" rather than the sharing, circulating or participating of human creativity and meaning creation (160). In this new online ecology, the self is operated as an entrepreneur and to some extent reduced to a branded commodity which is a "thing", across certain social media platforms such as Facebook and Twitter where such self-presentation is being encouraged (ibid.:

160). Social media entrepreneurs situate themselves in the Web 2.0 context and commodify the architecture as well as the intrinsic pursuit for high visibility of major social media platforms to construct their commodity self in the contemporary celebrity culture (ibid.: 161).

This relates to what Marxists mean by commodity fetishism as “a system of production in which relationships between human beings take on the appearance of relationships between things and in which the products of one’s labour no longer belong to the producer” (ibid.:160). In capitalism, products are produced for exchange value and profit rather than use (ibid.: 9-10). When applied to the more complex cultural production in celebrity culture, use values of cultural products and the desire to generate different meanings need to be addressed together with the specificities of particular types or typologies of celebrity in the process of value exchange (ibid.: 11).

Within Williamson (2016)’s discussion of human celebrities and stars, celebrities are commodities in three ways (12). First, those people’s labour is commodified. Then the performances as the result of their labour are commodities to be realized in the exchange act, which they can have control only when they are producers themselves. In a third sense, celebrities and stars are making themselves into commodities, being both the labour and the thing labour produces (Dyer 2004[1986]: 5, cited in Williamson 2016: 12). All the images or media personae they have constructed are vital to their exchange value as commodities, as those performances decide whether they are able to boost the sale and consumption of a film, a television drama or certain material products (Williamson 2016: 12).

Apart from the fundamental economic logics, capitalism is also dependent on specific cultural meanings (ibid.: 13). Similarly, celebrity as a dominant sign of the commodity form and idea of culture for financial benefits simultaneously addresses the desires and cultural values besides the logic of the economic (ibid.: 21). As a part of the hierarchy system and structure, celebrities undoubtedly facilitate the domination of media institutions, offer idealized images of the famous to promote consumerism and neoliberalism, and sometimes can even exclude or marginalize those who don’t fit in certain value system and ideology (ibid.: 26). Within the

contradictory celebrity culture, one is also able to find disputing and competing ideals, values and practices about what it means to be human, and non-dominant ways of being human when human are presented in commodity form while they and their humanity can never be fully contained by the self-as-commodity (ibid.: 26, 163).

Labour issue is very evident in the case of Kizuna AI. The contradiction between labour and capital applies well in the case of Kizuna AI's voice model and Kizuna AI's company which owns the character and controls everything. Based on Williamson(2016)'s critical examination of celebrity culture where human and commodity form permeate and interpenetrate each other, this thesis brings a virtual character into the discussion and adds to the understanding of a more complex format integrating human beings as labour, capitalists and consumers, and virtual and nonhuman character as celebrity and commodity.

### **3. Methodology and Methods**

The methodology and methods part consists of four main sections. The first section includes the thesis's social constructionism methodology and its connection to a qualitative approach and case study. The second part elucidates the thesis's multi-method design of combining semiological analysis and qualitative text analysis. The third section unveils the details of the research process and the last section reflects on the ethical issues and limitation of the research.

#### **3.1 Social constructionism**

The thesis lies in social constructionism as it tries to take “a critical stance toward the taken-for-granted ways of understanding the world” to critically inspect how the new cultural category virtual YouTuber is understood and practiced by different individuals and institutions in China (Burr 2015: 2). Fans and audiences' various responses to the virtual internet celebrity Kizuna AI echo with social constructionism's idea that “each different construction also brings with it, or invites, a different kind of action from human beings” (ibid.: 5). The research is guided by a social constructionist's perspective that people construct the knowledge and understanding of the world between them through the daily interactions in the course of social life, and the view of knowledge constitution as a product of “the social processes and interactions in which people are constantly engaged with each other” (ibid.: 4-5). It has laid emphasis on para-social and multi-social interactions among Kizuna AI, audiences, voice models and producing company to see how they engage and negotiate with each other to construct the meaning of a virtual YouTuber.

#### **3.2 Qualitative approach and case study**

Emphasis on social meaning and texts in social constructionism often leads logically to the use of qualitative methods as the research tools (ibid.: 28). Combining the fact that interpretive approaches can be more precise than quantitative approaches in understanding communication, the thesis adopts a qualitative and interpretive approach to investigate the para-social and multi-social communication and meaning making practices between virtual internet celebrity and her fans and audiences (Kuckartz 2014: 32). The qualitative data generated from such approach,

according to Miles and Huberman (1994), are “well-grounded, rich descriptions and explanation of processes in identifiable local contexts” that are likely to trigger fortuitous findings and novel integrations (1).

As qualitative analysis is “fundamentally case oriented” and qualitative data is often distributed around cases, the thesis centers on a case study of Kizuna AI to explore the “single entity” and “multiple instances” of the virtual YouTuber phenomenon (Bazeley 2013: 5). As the top and first virtual YouTuber who has generated plenty of attention and controversies, Kizuna AI’s case is no longer typical or average (Flyvbjerg 2001: 78). She can be viewed as an extreme case that can disclose more information as she “activates more actors and more basic mechanisms in the situation studied” (ibid.). The case-oriented approach of qualitative analysis underlines the situated interrelatedness of multiple aspects and characteristics within each example of the phenomenon (Bazeley 2013: 5). A case study can produce the kind of context-dependent knowledge that is able to “move from the lower to the higher levels in the learning process” of human-related affairs (Flyvbjerg 2001: 71).

### **3.3 Multi-method design**

Kizuna AI’s bilibili channel is one of the main places where Chinese audiences and fans are pulled into her storytelling and characterization through the videos, and pushed back through their comments, fan creation and collective practices, so I made it my main foothold. As analysis of symbolic forms such as language are essential to social constructionist research methods, I’ve combined two qualitative methods on different symbolic forms: semiological analysis of the video contents and qualitative text analysis of video comments (Burr 2003: 24; Rose 2016; Kuckartz 2014). This multi-method design aims to investigate both the production and reception side of the virtual internet celebrity, while achieving “triangulation” and increasing reliability and validity of the findings (Hansen & Machin 2013: 7). The design of analyzing comments instead of doing in-depth interviews derives from the research aim to examine multi-social interactions between fans and audiences. Within the comment area, I am able to see explicitly not only the opinions and meaning making of individuals, but also the way audiences and fans interact, negotiate and argue with each other when forming collective standing or

constituting different positions toward Kizuna AI, her voice models, and the producing company.

Semiology is an elaborate approach to describe and analyze how signs make sense (Rose 2016: 107). A semiological analysis requires the formation of a well-developed set of concepts to comprehensively document the ways in which meanings are generated (ibid.). By examining denotation and connotation of the signs in the videos, myth can be explored and related to “express the condensed associations of what is represented” (Hansen & Machin 2013: 176). Through exploring myth as “taken-for-granted models of the world that are broadly shared in society to understand how things work”, I get the chance to see the particular interests and power relations of different groups reflected in those pretty arbitrary models, and get closer to an already existing cultural repertoire of the human-non-human relation where connotations are activated (Hansen & Machin 2013: 176; Storey 2015: 126, 129).

The semiological analysis in the thesis mainly covers the visual elements and spoken text (captions) as it is often the text that provides “anchorage” which allows audiences to pin down the meaning of a visual demonstration (Barthes 1977: 38-41; Storey 2015: 129). The texts help viewers recognize the denotations of a visual presentation and in turn limit the possible proliferation of connotative meanings (Storey 2015: 129). Moreover, the written or spoken text in media has an important “relay-function” which amplifies the connotative potential of the image (Barthes 1977: 38-41; Storey 2015: 128). Such function attempts to let audiences believe that “the connotative meaning is actually present at the level of denotation” (Storey 2015: 129). Therefore, the combination of both visual and verbal messages ensures a more holistic and accurate interpretation of Kizuna AI’s videos and the cultural repertoire the producers are trying to add to.

A qualitative content analysis is likely to overcome the weakness of the classical content analysis which lays much emphasis on numerical results and integrate steps of interpretation into the analysis process (Kuckartz 2014:16). Qualitative text analysis enables a “deductive-inductive category construction”, which allows the top down deductive categories to be



acquired from category systems that existed before collecting empirical material, and inductive abstractions to be generated from the open coding process where one investigates, compares, categorizes and develops data into preliminary concepts and dimensions (ibid.: 23, 55, 62). Rooted in Grounded Theory, qualitative text analysis has enabled me to verify my hypotheses and assumptions of Kizuna AI's audience engagement and disclose unexpected findings and patterns that are "lurking within" the comment area (Legewie & Schervier-Legewie 2004: 51, cited in Kuckartz 2014:22).

### **3.4 Conducting the research: piloting, sampling, data collection and analysis**

#### ***3.4.1 The pilot***

Before officially gathering data, I chose the most popular video on Kizuna AI's channel as my pilot to test if my multi-method design can allow me to generate analysable and fit-for-purpose data and assess how long the analysis is going to take (Bazeley 2013: 55). The qualitative text analysis went well and inspired my sampling strategy of choosing most commented videos, for there usually are richer messages in the comments when a video triggers more debate and discussion among audiences. The initial semiological analysis pilot which only included captions and Kizuna AI's actions to invite social interactions didn't go well, so I adjusted the method by integrating Hansen & Machin's ideas on semiotics (2013), while also integrating Gillian Dyer's checklist for exploring what signs of humans might symbolise (Dyer 1982: 96-104, cited in Rose 2016: 115) because my main focus Kizuna AI is a human-shaped female-like protagonist of all the videos. Halliday's (1985) notion of speech acts is also added for more systematic inspection of spoken text in the videos. Then I ran another test round of the refined semiological analysis scheme<sup>12</sup>, and it applied well.

#### ***3.4.2 Sampling***

Kizuna AI's channel integrated more than 1200 videos of various genres (as of February 2020). I started my sampling with an exploratory categorization of all her videos. I first took down the details (including likes, replies, forwards, video length, etc.) of her latest 150 videos (as of Feb

---

<sup>12</sup> See Appendix A – Semiological analysis for the detailed checklist

24, 2020) and came up with 5 categories based on the style of the videos: ASMR, daily (storyline performance<sup>13</sup>), game, music and reaction.

Within this process, I found the voice model incident had caused a lot of conflicts, discussion and even collective “activism” within the comment area regarding the virtual internet celebrity Kizuna AI, the producing company and different voice models. Therefore, I decided to pick the most commented video both before and after the voice model change in June 2019 from each category to observe potential changes or shifts in audiences’ engagement with Kizuna AI. Those videos have the largest potential to showcase the formation of para-social interaction between audiences and Kizuna AI and the multi-social relations established among audiences and fans. During the investigation of the rest 1100 videos, two more categories: real world interaction with human and livestreaming were added. A total of 14 most commented videos were then selected as the samples<sup>14</sup>.

### ***3.4.3 Collecting and analysing data***

#### *3.4.3.1 Data collection*

The main focus is laid on Kizuna AI’s visual appearance, speech and actions in the videos and audience/fan comments around the definitions, meaning and values of Kizuna AI and her voice model(s). Other information such as discussion merely about the game she played was not included in the very beginning. Semiological analysis covered the visual elements and captions of all 14 videos. For qualitative text analysis, the most popular comments and all the replies to those comments are first checked to ensure the depth of the empirical data. Based on the “saturation” rule, I stopped collecting data when I saw “no new properties of the pattern emerge” (Glaser 2001: 191, cited in Charmaz 2006: 113). In order not to leave out important messages in comments which are not that highly ranked, I also looked at the top 50 comments (without the replies to them) of each video to ensure the width of the data. The numbers of comments and replies collected for each video vary from 114 to 809.

---

<sup>13</sup> As Kizuna AI is a virtual character, all the videos featuring her daily life are totally well-planned and scripted stories that are different from other YouTubers’ somewhat documentary daily videos/vlogs.

<sup>14</sup> See reference for video details

### 3.4.3.2 Data analysis

Based on Rose's suggestion on conducting a semiological analysis (Rose 2016: 132), I followed the aforementioned analyzing scheme to first examine each video individually<sup>15</sup>. After deciding "what the signs are" and "what they signify 'in themselves'", I also tried to think about "how they relate to other signs" and "their connections to wider systems of meaning" (ibid.). After completing 14 individual semiological analysis, I went through all the results to come up with analytical categories centering on two focuses: Kizuna AI's characterization and self-definition, and her performed interaction with imaginary fans and audiences to invite and ask them for support and engagement, and further nailed down "the articulation of ideology and mythology" of Kizuna AI as a digital commodity which promotes and sells the value of nonhumanness, virtuality and connectivity of internet culture (ibid.).

The qualitative text analysis follows Kuckartz's instruction (Kuckartz 2014). As the examination of audiences' part is guided by concepts around media engagement, the whole coding process is inspired and thematically led by the model of spectrum of engagement (Hill 2017; 2019). The deductive framework consists of emotional engagement with positive engagement, negative engagement and disengagement as subcategories, critical engagement and context to engagement. Within this framework, I started by inductively reading and interpreting the texts during open coding of the first 3 videos' comments based on my research questions, and then formed them into preliminary analytical categories that are not entirely descriptive but are part of a deeper analysis and higher level of abstraction (Kuckartz 2014: 41). After connecting these categories back to the engagement model, I went back to the comments of the rest 11 videos and coded them according to the existing categories while constantly adding new categories and making adjustments to fit for the purpose of my research<sup>16</sup>. The following analysis part is generated based on the combination of my finalized comment codes, semiological results, and the memos of the thoughts, ideas, assumptions or hypotheses occurred to me during data collection and processing (ibid.: 52).

---

<sup>15</sup> See Appendix A for example of one semiological analysis

<sup>16</sup> See Appendix B – Qualitative text analysis

### **3.5 Ethical issues and limitation of the research**

While collecting data, researchers should be aware of the influence on the subjects of the research (Flick 2007: 73). The comment areas under videos on bilibili.com is a public and open online space. Remaining an academic “lurker”, I didn’t identify my presence to the bilibili users and focused more on the textualized output, patterns and the dynamics within Kizuna AI’s audiences instead of individual demographics (Gray 2005: 847). I’ve anonymized all the commentators in the research and chose not to give them fake names or use their bilibili alias as I don’t have a reliable way to get their demographic information<sup>17</sup> and nicknames alone don’t reflect much. Another unavoidable issue is the innate preconceptions and inherent skills and thinking from the researcher (Bazeley 2013: 4, 27). The way to analyze and interpret data is unavoidably colored by my personal, cultural and social experience and background (ibid.: 4). My personal experience and interests in Japanese anime, manga and game have allowed me to develop certain insider views throughout the research but sometimes they also cause the lack of clarification of what I’ve long taken for granted. Even though I’ve been asking for verification from non-fan friends, my own experience influences the way I obtain and process my data (ibid.: 51).

One major limitation of the research is the language and translation issue which is twofold. On one hand, I am not able to read Japanese which is the main language Kizuna AI uses to perform and present herself in most of the occasions. The dependence on subtitle groups’ translation or official Chinese captions actually distances me from the original media product if translation is viewed as a kind of “misperception” or “an event which involves an element of risk and chance” where translators’ own interpretation and re-creation are involved (Tatsumi 1992: 372, cited in Sanders 2005: 140; Sanders 2005: 140). On the other hand, all the comments presented in this dissertation are translated by me from Chinese to English, which certainly can be viewed as a re-creation process full of my personal understanding and judgement. Another limitation comes from time. During my examination of the channel since its establishment, I often found

---

<sup>17</sup> It’s possible to click into a bilibili user’s public homepage and find more information such as gender. However, plenty of users chose not to disclose their gender on their pages and the disclosure could sometimes be fake.

unavailable comments that had already been deleted, especially in those relatively old videos. This has caused my inability to pin down some of the discussions or controversies during the inspection of fans and audiences' multi-social interactions in the comment area, and I sometimes had to guess and imagine for the whole story. The comment display algorithm of bilibili.com is also a restriction, as I didn't have the capacity to go through all the comments of the selected videos. The platform's choice to highlight certain "popular" comments and their replies has influenced my data collection process and somehow impacted my final results.

## 4. Analysis

The analysis is divided into three parts to cover the three key sides: producers, audiences and voice models who are co-creating the human-non-human experience within Kizuna AI's case. The first section focuses on the appeal producers constructed in Kizuna AI to pull audiences into her characterisations and performances in the videos. The second part turns to audiences' reaction to her appeal in the videos and shows the dynamics within Kizuna AI's audience group where people engage emotionally, critically and morally with her with various intensities and proximities. The third part examines Kizuna AI as a form of digital commodity and brings in the discussion of capital, labour, capitalism and socio-cultural values in cultural production.

### 4.1 The appeal in Kizuna AI

Kizuna AI's appeal in character design comes from her visual appearance which integrates features from cyberculture and Japanese culture, her self-definition as an AI YouTuber which colours her celebrity existence with exoticism and exceptionalism, the intimacy she builds up through her performances in the videos, and the authenticity she constructs in the real-time interactions with the audiences in occasions such as livestreaming.

#### 4.1.1 *Virtuality and Japanese visual heritage*

There are three main components in Kizuna AI's visual character design: pixel vixen, the style of "3-D computer-generated images (CGI) of digital women who are young, slim, fair skinned, wide eyed, and often scantily clad (Matrix 2006: 8)", kawaii aesthetics, "the peculiar expression of cuteness-related aesthetics in Japan (Pellitteri 2018:94)" and moe elements, "traits that invite particular forms of otaku affection and that are recognizable across a wide range of anime and manga characters and narratives (Ito 2012: xxii)".

Pixel vixens are from cyberculture where gender acts as part of the social programming process which is intentionally encoded instead of naturally being there (Matrix 2006: 120). At first sight, Kizuna AI denotes a 3-D computer-generated model in the shape of a human being with healthy skin, nice body proportion and an extremely symmetrical face. Moreover, her virtual body is

displayed with evident features of a young woman: long and shiny hair over the shoulder, a pair of bright eyes on a smooth face without any wrinkles and a shapely body covered under her delicate outfit. All those features are connoting her similarity to human beings instead of a machine, and connection to femininity. Deeming those features attractive and beautiful, she keeps emphasizing her beauty in the media production with statements such as “I am pretty”. What lies behind such confidence and self-statement is an idea from the real world of appreciating characteristics such as white skin, slim body and long legs (which Kizuna AI has) as the commonly acceptable beauty for females. Apart from this beauty myth that has been extended from the power relations and knowledges in real world to the design of virtual characters, Kizuna AI’s visual design is also closely bound up to its geographical spatialities (Haraway 1999: 108, cited in Matrix 2006: 106). Originating from Japan, she bears clear heritage from Japanese visual culture, among which *kawaii* aesthetics and *moe*-elements are the two most prominent Japanese legacies in her visual appeal.

Besides recognizing Kizuna AI’s continuous self-statements such as “It’s me. *Kawaii!*” and “A cat as *kawaii* as I am”. The *kawaii* aesthetics should also be explored in connection to the concept of *shōjo* (literally young woman) in Kizuna AI’s case. Here, *shōjo* refers to “a character type, which may appear in entertaining graphic narratives, non-narrative games, or branding campaigns” as the cute adolescent girl (Berndt 2019:1). Kizuna AI has said in her livestreaming video that she looks like 16 years old on appearance, and concluded audiences’ comments on her characteristics with “I’m *kawaii* in general; by default, I’m a perfect *bishōjo* (literally beautiful young woman)”. She is often with sweet smiles, speaks in an energetic and lively way and actively reacts to whatever she sees: a game, an artwork, a photo or feedback from audiences and fans, connoting a sense of vitality and curiosity that is common in young humans.

Her clothing also displays a common type of *kawaii* fashion, “a style that is deliberately designed to make the wearer appear childlike and demure” (Monden 2019: 211-212). Her clothes are mainly white and pink dotted with only small amount of black, and she is always wearing a conspicuous heart-shaped hairpin in the same kind of pink colour, to connote a general sense of purity and gentleness. Denotations such as the big bow knot and clothes’

necklines are connected to the style of ‘sailor suits’ (girls’ school uniform in Japan), easily reminding people of sweet bishōjo characters in Japanese animation, manga and video games. In music videos when Kizuna AI is portrayed a little differently as a singer, her ensemble integrates more colours and ornaments to connote a sense of distinctive vitality in young females, while also keeping consistent with the major character appearance by practices such as sticking to the same kind of hairpin style.



Normal appearance<sup>18</sup>



Outfit as a singer<sup>19</sup>

There are traces of the very affective moe elements in Kizuna AI’s character design to attract the otakus who are innovative popular-culture enthusiasts passionately participating and engaging with technology in a networked world (Ito 2012: xi, xxi). This moe feeling as “romance inside one’s head’ (Honda 2005, cited in Tsuji 2012: 28) stays quite limitedly within the otaku subculture and needs some insider knowledge to enjoy. One intensively engaged moe element is “alter”, which refers to a drastic change in an anime or game character’s personality or mental state, and the character usually switches to a dark identity. The “alter” element first appeared in Kizuna AI’s video when she was said to had infected with virus and generated another personality of “Black Kizuna AI”. The altered Kizuna AI is listless, pessimistic and takes negative attitudes toward people and things. She was then named “Black AI”, and got her own dark colour outfits in the following videos, connoting a sense of suppression and coldness totally opposite to Kizuna AI’s vitality and optimism. The setting of two contradictory

<sup>18</sup> <https://www.bilibili.com/video/BV1ox411S7Q7>

<sup>19</sup> <https://www.bilibili.com/video/BV1cb411s7U4>



personalities has attracted plenty of audiences and led them to involving in numerous “fannish” practices and creation (Ito 2012: Xi). There are simple affective claims such as “Black AI is my wife”, role-acting under Kizuna AI’s videos using Black’s usual emotionless tone saying “Are you still alive, idiot?” and even debates between passionate fans who hold different opinions on the relationship between Kizuna AI and Black AI. Adopting such moe elements enables producers to offer topics to (otaku) audiences and pull them into the enjoyment of sharing and playing with insider knowledge in Kizuna AI’s video channel (ibid.: Xxii).

#### ***4.1.2 An exotic and exceptional AI YouTuber***

Apart from the character’s visual manifestation, Kizuna AI is further portrayed as an exotic and exceptional internet celebrity by claiming herself as an AI, a machine different from human beings, and defining herself to be (probably) the world’s first virtual YouTuber, a brand-new production format to allure the audience side (Abidin 2018: 22-23, 28-29). Kizuna AI follows the same pattern as other pixel vixens, virtual idols and influencers who make their non-human charm stand out through their character design (Matrix 2006: 113-114; Lam 2016:1114; Marwick 2019: 162). Facing audiences who have “identity category” as human beings, she puts forward the nonhuman self-definition as an AI in her very first several videos and follows this “AI formula” to develop a consistent and standardized performing pattern as an “appropriate production format” (Horton and Wohl 1956: 217). She says “I’m different from you humans” and “I’m a super AI” regularly in her videos, stresses her non-human aging of being three years old (based on an AI assumption), and makes her given name AI (which means love in Japanese) into a careful play around this tasteful AI attribute to “piques the interest of audiences who hold contrasting or different forms of “cultural capital” (Abidin 2018: 22). All of those are connoting the exotic non-human quality of Kizuna AI, different from the aforementioned similarity to human beings or human qualities in her visual design.

The technical capital of “knowledge and mobilization of digital technologies to access information, achieve social mobility, and ensure social inclusivity” is evident in Kizuna AI’s virtual existence. This feature enables her be an internet celebrity exceptional enough to stand out under the attention economy of the internet (ibid.: 28-29). Kizuna AI and all VTubers’ pure

existence as digital data is enough to satisfy the desires of fans who are longing for technology that is not possible in a living celebrity (Black 2012:217). Self-titled “probably the world’s first virtual YouTuber”<sup>20</sup>, Kizuna AI brings about strong interactivity and even sociability based on the digital technology utilized during the deliberately concealed production process. Some netizens<sup>21</sup> speculate the producing team has been using exceptional motion capture techniques during livestreaming and video recording to enable Kizuna AI’s real-time interactive performance. During interviews, Kizuna AI can act like a professional human host who asks prepared questions with further extension into impromptu speeches, except the fact that she is two-dimensional inside a large monitor, integrating opposite connotations of human quality and non-human quality. During livestreaming, audiences and fans are able to see her answering and reacting to comments that are not prepared in advance, and get a sense of liveness and opportunity to participate in this extraordinary and exceptional experience of communicating instantly with a virtual celebrity.



Kizuna AI as a host<sup>22</sup>



Example of livestreaming<sup>23</sup>

The virtuality-reality interrelation and an intension to bridge the gap between human and nonhuman are also embedded in Kizuna AI’s videos. Take her music video AIAIAI<sup>24</sup> where she sings and dances on a screen with human companions who treat her in the same way as they treat other human performers as an example. The human dancers keep the same close distance

<sup>20</sup> <https://www.bilibili.com/video/av17814976>

<sup>21</sup> <https://www.zhihu.com/question/57760908>

<sup>22</sup> <https://www.bilibili.com/video/BV1w441167Wt>

<sup>23</sup> <https://www.bilibili.com/video/BV11W411p73F>

<sup>24</sup> <https://www.bilibili.com/video/BV1cb411s7U4>

with each other as well as Kizuna AI, and actively interact with Kizuna AI through movements such as clapping with her, delivering the connotative meaning that there is not much difference between Kizuna AI and human dancers; they can connect well in the real world. Moreover, the final scene of the video goes back to Kizuna AI's virtual world and denotes human beings on a screen to indicate Kizuna AI's separation from human beings as a virtual existence in another world. It further connotes that the real and the virtual are connected and Kizuna AI is the key to connect them.



Music video AIAIAI

#### ***4.1.3 Performing intimacy***

Constructing a sense of intimacy in the performances has long been the key for traditional media celebrities who prompt a sense of connection through their extraordinary personas that have been established and spread through mass media (Horton and Wohl 1956: 218; Rojek 2001,

cited in Marwick 2016: 333). This cultivation of intimacy is also extended to social media and the internet where microcelebrities use strategic intimacy to appeal to followers and build their fame through the interactive relations with the audiences (Senft 2008; Marwick 2013; Marwick 2016). Similar to her human counterparts, constructing a sense of intimacy remains one of the most prominent approaches for Kizuna AI to invite (para-)social interactions from audiences as a media persona.

She maintains the characteristics to duplicate conversations styles and gestures in an informal face-to-face gathering (Horton and Wohl 1956: 217). Throughout the majority of the videos, Kizuna AI tries to keep constant and direct eye contacts with the imagined audiences in the way human beings do in regular conversations. In the videos, medium shots and close shots are most frequently used to present Kizuna AI from waist up or shoulder up, showing in detail her facial expressions and body movement to construct a feeling of closeness in her audiences and connote the intimate social relation between the celebrity and the viewer.

The settings in some of her videos also help amplify the connotation of intimacy. For instance, one ASMR video displays Kizuna AI in her pink pajamas conducting daily chitchats with her non-specified boyfriend in a tender and supportive tone<sup>25</sup>, while another ASMR video ends with plot that the non-specified audience falls asleep after Kizuna AI's ear treatment<sup>26</sup>. Both of the two scenarios only happen in very intimate relationships. Sound technology has also been fully utilized to manage and create the feeling of intimacy between the persona and the audiences in her media production, especially in the ASMR category to connote physical proximity between the listener and Kizuna AI. In the aforementioned virtual ear picking video, audiences can sense the feeling that Kizuna AI literally is by their side and picking their ears from the technically generated sound effects and the visual demonstration of Kizuna AI moving closer to the audience's ears.

Fantasy and imagination instead of obligation or responsibility are what stimulate an audience

---

<sup>25</sup> <https://www.bilibili.com/video/BV1hE411R7xL>

<sup>26</sup> <https://www.bilibili.com/video/BV1Ex411z7DK>

to remain involved in the para-social relation (ibid.: 215). Within the formation of imaginary relationships with the audiences, celebrities rely heavily on fabricating an artificial public persona (Black 2012: 215). Kizuna AI has the same ability as Hatsune Miku to allow a kind of “far more intimate and all-encompassing” relationship with the consumers, which is not available in a living celebrity, thanks to her lack of a “real life”, private or true self independent of her fans (ibid.: 224-225). Audiences get the chance to assign the attributes they like to their beloved virtual YouTuber without worrying about the discrepancy between their imagination and Kizuna AI’s real identity or private self, as “no part of her exists outside flows of ... digital texts” (ibid.:225). The arbitrary quality on Kizuna AI’s self is differentiating her from human celebrities and developing the idea that one can have greater space for imagination in virtual celebrities compared with human celebrities. The aforementioned comment “Black AI is my wife” can be a short but good verification.

This idea also applies in Kizuna AI’s decontextualized videos and can even be further developed into: one sometimes has to imagine by himself/herself to supplement the missing parts in a decontextualized scenario to fully enjoy the videos of specific plots and themes, or with little background information or introduction. In the aforementioned girlfriend role-play ASMR video, nothing but a still image of Kizuna AI wearing pajamas has been provided without any specification of the time, place or participant of the conversation. Audiences have to give full play to their imagination in unveiling and defining the scenario with the scattered messages provided by Kizuna AI to fully engage with her sentences starting with “you”. They are pulled deeper into the video’s storytelling while trying to situate themselves through imagination. In this video, the only information of the listener is that the person goes to work (instead of being a student etc.). Any audience who fits in with this specification in real life can easily do the imagination and complete the dialogue context. This example further illuminates that the aforementioned fantasy and imagination don’t have to stay only in the one-sided dyadic para-social relation but also can continue to the real life and become the social (Hills 2016: 470).

In another video when Kizuna AI is supposed to reach out to the audience’s ears, she comes

closer to the camera lens and only leaves her elbow visible for the audiences. This openness pushes those audiences who want to engage naturally into the voluntary imagination of picturing how Kizuna AI is helping them with the ear treatment in their own way. All such voluntary imagination and fantasy further extends the intimacy constructed from Kizuna AI's up-close visible performances, and turns Kizuna AI into a more personal existence within an audience's imagination.



Ear picking ASMR<sup>27</sup>



Girlfriend role-play ASMR<sup>28</sup>

#### ***4.1.4 Authenticity in real-time relation***

Authenticity is another key construct social media can create in the online celebrity-fan interactions (Click 2013: 366). According to Elcessor (2012), “illusions of liveness and interactivity in online media are crucial to experiencing online celebrity as uncontrolled, ongoing, immediate, and ‘real’” (52, cited in Hills 2016: 474). Kizuna AI's showcase of unedited and imperfect performance or even mistakes is one kind of the uncontrolled expressions to generate a sense of liveness and authenticity. During one livestreaming, she tried three times to finally successfully pronounce “kyaripamyupamyu”, and reflected by herself that “as it's livestreaming, this cannot be edited. I usually cut this (kind of content) off...please forget it haha”<sup>29</sup>. Consideration of such performances of making mistakes and displaying a “raw” Kizuna AI without any editing as authentic also connotes the human quality in Kizuna AI, for she is not portrayed as a flawless machine.

<sup>27</sup> <https://www.bilibili.com/video/av11949555>

<sup>28</sup> <https://www.bilibili.com/video/av69390978>

<sup>29</sup> <https://www.bilibili.com/video/BV1q4411Z7FE>

Time has also been fully utilized in Kizuna AI's videos to construct a real-time experience for the audiences. The sense of real-time experience first comes from the display of what Kizuna AI "was seeing" at the filming moment, be it a real time game interface or a fan video she was performing reaction to. Her instantaneous reactions to the contents in the videos with various expressions: nervousness, concentration, adoration, dissatisfaction, etc., evident gestures and body movements make audiences who are watching the videos feel they are facing with the same visuals as the virtual YouTuber. In this way, a sense of involvement could be generated as audiences are juxtaposed with the same unpredictability as Kizuna AI was, and probably would start to anticipate or worry about her next choices or movements based on their own assumptions, as if they were experiencing the contents at the same time.

The videos also do well in leaving blank space, to allow voluntary contribution from the audience to fill it up. For example, in the girlfriend ASMR video, Kizuna AI initiated a conversation around today's work. After she asked "What happened? Please tell me no matter how trivial it is", there was complete silence lasting for a couple of seconds for audience to put forward their answers before she replied with "ah...I see". All the waiting and absence of words are pulling audiences into the illusion of co-existing with Kizuna AI and having real-time feedback from her.

The most important real-time experience happens during livestreaming when audiences and Kizuna AI are co-existing within the same time realm and viewers have real time experience with Kizuna AI both within the virtual space and outside in real life. The real-time interactions enabled by Kizuna AI's spontaneous responses to live comments further enhance this experience. Once, she noticed an audience saying his/her livestreaming image stopped moving, and started doing big and continuous body movements right away while telling the audience "if you don't see me move then your internet connection is probably quite bad". It can be possible for such audiences who are replied to experience not only a real-time interaction but a somewhat "real" relation very close to face-to-face exchange (Horton and Wohl 1956: 216).

One way for Kizuna AI to allure audience into the potential of having such real interaction is

through what is similar to Horton and Wohl (1956: 219)'s idea of "coaching of audience attitude". "What to expect of the persona, what attitudes to take toward him, what to 'do' as a participant in the program, is not left entirely to the common experience and intuitions of the audience", but is embedded in the rest of the audience and fans who have been constantly "taught and educated" through watching how the selected audience perform in the fan videos and during real-time responses to support the persona (ibid.: 219). Those "sample" attitudes and way of reactions are then considered the potential way to get in touch with Kizuna AI. The seemingly randomly picked comments help "lure the attention of the audience, and create the easy impression that there is a kind of participation open to them" (ibid.: 218). The responsiveness throughout Kizuna AI's videos and livestreaming pulls audience into a feeling of hope to turn this basically one-sided interaction into a temporarily dialectical one just as the "selected others" have done, and in this way Kizuna AI invites audiences to willingly reply and contribute to her channels.

#### **4.2 Engaging with Kizuna AI on bilibili.com**

The audience group of Kizuna AI are far from monolithic (Gray 2003: 68). The most active are fans who engage intensely with everything around Kizuna AI, both affectively and cognitively (Hill 2017; 2019) and can shift to the opposite of fandom and become anti-fans when they dislike certain media text and choose to treat them morally before aesthetically (Gray 2005). Aside from enthusiastic fans and anti-fans, there are also normal viewers as non-fans who sometimes call themselves "passers-by" and just watch Kizuna AI's videos randomly as they please. This part is unfolding the different levels and modes of engagement from the diverse audience group and their practices on bilibili.com, with primary reference to fans and anti-fans to unveil how their (potential or real) participation as consumers are enhancing the proximity of producers and audiences (Zaborowski 2016: 124).

##### ***4.2.1 Building an affective space on imagination and fantasy***

As a constellation of data, Kizuna AI is both fictional and real. She is "real in terms of having material effects on people's lives and playing a role in the formation of digital lifestyles, and ... fictional insofar as it operates in conjunction with an elaborate fantasy narrative" (Matrix 2006:



106). Under this fantasy narrative, audiences and fans not only experience her as “algorithms and abstract symbols”, but more like user-friendly data that has been “reengaged with personality and the imagination” (Morse 1998: 5, cited in Matrix 2006: 106). They take the “voluntary and independent answering role” to Kizuna AI’s invitation to establish para-social relations in this affective space where they develop connection to the virtual YouTuber and retain control over the content of their own participation (Horton and Wohl 1956: 219).

Affection is one of the most frequently expressed emotions toward Kizuna AI. In this “imaginary social relationship” fans show “attachments to this virtual media persona that is ... in many ways ... parallel to actual social relationships with real ... ‘friends,’ and ‘lovers’” (Caughey 1984: 23, 40, cited in Hills 2016: 468). One felt “empty in the heart” and “the feeling of losing a good friend”, and another got a feeling of “breaking up” when they both decided to unsubscribe Kizuna AI’s channel. Apart from directly saying “I love you Aichan!<sup>30</sup>”, fans left comments such as “Ah, I’m in love”, “the ‘good night’ really moved my heart” under her videos to showcase their fondness toward her gentle voice and adorable performance. They declared “I’m going to die if I can’t have Kizuna AI” and “VTubers are now part of my life” to acknowledge Kizuna AI and VTuber’s significance to their life.

Interestingly, quite a lot of fans would call Kizuna AI “wife” to show a particular kind of love. They not only straightforwardly claimed “Black AI, is my wife” , or “My wife (referring to Kizuna AI) is so good” , but also enjoyed a special kind of interactivity when several people all claimed Kizuna AI to be their wife and replied to each other with sentences like “She is not yours; she is at my place” , “She is on my bed” and “She says she wants to stay in my house”. What one can observe here is a sharing of close relation with Kizuna AI, which corresponds well with the construction of intimacy and arbitrariness in Kizuna AI’s videos. While this could also be further connected to the heritage of otaku culture where making such connection to Kizuna AI serves as a means of circulating insider information and building social interaction

---

<sup>30</sup> Chan (ちゃん) is a suffix in Japanese, expressing that the speaker finds a person endearing.

Aichan was the nickname for Kizuna AI before voice model change, and the now it often refers to the presence of the first voice model.

with other insiders (Ito 2012: xxii).

Despite Kizuna AI's essence as virtual data, fans don't challenge her authenticity but react to the celebrity with various meaning making practices enabled by emotional realism which is "the trait of being deemed by the audiences as true to feeling, rather than true to literal reality" (Ien Ang 1985, cited in Zaborowski 2019:107). There is happiness enabled by merely watching Kizuna AI's videos. One audience commented "it's really happy to watch her videos" and another said "she has brought too much joy and self-healing to me; I could watch her videos from 12 am to 7 am without knowing it". The release of Kizuna AI's music has also brought plenty of delight to her fans. One fan shared that:

"Last year after knowing Kizuna AI's first song Hello Morning was first to be released on NetEase Cloud Music, I was so happy that I stayed with my phone everyday until its release. I remembered the ecstasy when first listening to it and interacting with some people in the song's comment area. To fans, at least to me, this song has very special meaning."

This fan's devotion to Kizuna AI is no less than that to a real human singer, and the same affective emotions and meanings enabled by experiencing the song has been constantly agreed upon and transferred among other fans in their replies such as "We are the same. Each of Aichan's songs is an unforgettable memory for fans".

Fans also sense the feeling of connection in Kizuna AI's videos. To them, Kizuna AI is "a symbol of the industry who has created the trend of VTubers, and more importantly, the source of bonding which connects Kizuners (the nickname for Kizuna AI's fans) all over the world". Moreover, the connection and bonding not only stay within celebrity and fans, but can extend to Kizuna AI and humankind. Here is what one audience passionately wrote under Kizuna AI's music video AIAIAI:

"During 0:42-0:49 of the music video, dancers and Aichan have synchronous body movements. Dancers hide behind the screen while Aichan appears from the edge of it, making people have an association that 'dancers become Aichan'. It is a symbol, an enlightenment – ordinary people can also use Aichan's way to communicate with the world ...Recalling Aichan's progress of the past two years...I do feel she is changing the world in a silent and soft way. For Kizuna AI, the [AI] part feels like an eye-catching headline, but [love] does exist when she continually carries on her

activities for two years with the dream of bonding with humankind...I will always respect Kizuna AI.”

S/he first detected a symbolic connection between Kizuna AI and human dancers in the way they are portrayed based on the video’s aesthetic text, and then turned to very affective interpretation and intense engagement of both the media content and Kizuna AI’s overall performance. The fan alienated Kizuna AI from all mankind, and situated the virtual celebrity in a quite high level where her video can be “enlightening”, she can enjoy “respect” from fans and she is building connection with human beings out of “love”.

As the values of involvement and intimacy are stressed in the videos, audiences are getting accompany and encouragement from the para-social relations with Kizuna AI. They can easily distinguish Kizuna AI’s voice because her voice has accompanied them for plenty of nights. Additionally, her existence and performance offer emotional support to help sooth audiences’ difficulties in real life. One audience shared his/her encounter with Kizuna AI in 2017 in the comment area of Kizuna AI’s first video<sup>31</sup>:

“I was in my 11th grade in 2017 and was too cool to have friends (I actually had friends, but only a few). Under growing studying pressure, I found Aichan at my loneliest time and felt like having one more interesting friend. I couldn’t really play or relax due to the ever-growing pressure and would feel guilty for playing video games even occasionally. But I felt relieved every time I watched her videos, and could obtain an unrestricted kind of happiness...I’ve always found Aichan lively and full of energy. She is even driving me who is often in low spirits to treat life in a pink light.”

To him/her, Kizuna AI’s videos and the moment s/he was watching them served as an escapism from the overwhelming real life. It’s possible to predict that his/her consumption of Kizuna AI and her videos was based on a “lack” of adequate close friends or a positive attitude (Baudrillard 1968; It. trans. 2004: 255, cited in Pellitteri 2018: 100). But more importantly, the person’s interaction with Kizuna AI didn’t stop at a stage where s/he fulfilled completely in imagination the part s/he wasn’t able to do in reality (ibid.: 222). The experience and engagement with Kizuna AI’s fascinating personality and the constructed intimacy were actually changing the

---

<sup>31</sup> <https://www.bilibili.com/video/av9800170>

audience's life, offering reassurance in the hard time and calling out positiveness in the long run.

#### ***4.2.2 Customizing Kizuna AI's character postulation***

The majority of fans have realized that Kizuna AI is not an artificial intelligence that she claims herself to be. Nonetheless, the loyal audiences “accept the gambit offered” by completing Kizuna AI's performance through “enactment” of their role and in this way participate in the para-social and multi-social interactions (ibid.: 219-220). They play with her AI characteristic individually, and further enlarge it to something more than just the moments of video watching through fan discussions in the reciprocal communication enabled by the possibilities of social media (Hills 2016: 472; Gray 2003: 76). For instance, one audience replied “I'm artificial intelligence; I never blush” in the tone of Kizuna AI to another's question of “Don't you blush (for saying ‘I'm a genius’ in the video)?”. While others called the word “ち ゅ め て” that Kizuna AI made up in the lyric of her song, a kind of “AI language” .

After being pulled in the video contents and character design, fans and audiences are also pushing back to further develop and customize the characteristics settled by the production side. Despite Kizuna AI's recurring claim of “I'm a super AI” in her media presence, fans are calling her “Artificial Idiot” for her not that intelligent performance in her videos, for example, the really bad gaming skills and calculating ability. They said “the God gave her an exquisite face and a perfect figure, but took away her IQ”, and more frequently, they combined this newly-created feature with video contents. They named Kizuna AI's interview with Alan Walker<sup>32</sup> “renowned music producer Alan Walker's journey to lower his intelligence”. In one ear picking ASMR<sup>33</sup> video, audiences on bilibili.com reacted with “Aichan picked away your IQ” and told the story that “I've recommended this ASMR video to my classmate. He insists on listening to it every night, and now he has regressed into a monkey”. However, it's worth noting that this “artificial idiot” customization is more a distinctive trait than a negative accusation, as fans have been clarifying themselves: “we are not really calling her stupid when we say artificial

---

<sup>32</sup> <https://www.bilibili.com/video/BV1w441167Wt>

<sup>33</sup> <https://www.bilibili.com/video/av11949555>

idiot”, and showing considerable appreciation to this characteristic: “I love you so much, artificial idiot Kizuna AI” ; “I’ve watched this video repeatedly for dozens of times. I think I fell in love with an artificial idiot.” The normally offensive word “idiot” has been shifted to a distinctive endearment widely accepted and utilized among knowing fans. “Artificial idiot” could even be viewed as a sign which constitute part of the insider knowledge of Kizuna AI and the VTuber industry. Those who know it and are able to engage with it will oftentimes automatically obtain a fan identity in others’ eyes and consequently a closer distance to the virtual YouTuber.

The customization not only stays within the fans/audience community, Kizuna AI is also integrating some of the audience-generated features. For instance, she once mentioned that audiences are calling her “ぼんこつ (a piece of junk)” for she generally tends to fail in what she does, which echoes well with the aforementioned nickname given by Chinese fans to add to her “stupid yet adorable” performances. Fans’ work also makes up a good part of Kizuna AI’s media presence. Apart from the abundant livestreaming messages and video comments, fans are making full use of their fan labour to publish song covers, animated video clips, cosplay photos and many other art forms relevant to Kizuna AI across different social media platforms, freely publicizing their beloved media persona in the way they want. At the same time, they are also submitting those contents to Kizuna AI’s channel as potential materials to be reacted to by the virtual internet celebrity and be absorbed in as part of the “official” text. Apart from such customization and dissemination, a more distinctive “push back” from the audiences started when Kizuna AI’s voice model incident took place, where audiences and the producing company debate and negotiate around “who is Kizuna AI” and the voice model’s role in constructing this virtual internet celebrity.

#### ***4.2.3 Recognizing and acknowledging the voice model***

Realizing that Kizuna AI as an artificial intelligence system is just a marketing strategy, audiences have long recognized the existence of voice model. Even though the company had been deliberately avoiding disclosure of information regarding the producing team, fans spared no effort in gathering information and making predictions before they finally narrowed down

to one specific female voice actress Nozomi Kasuga as Kizuna AI's voice model. There were not a lot of discussions and interactions around this voice model before the voice model incident, as one audience has pointed out that "in the past when there was no risk of displacement, people didn't care about the voice model". Nonetheless plenty of debates and negotiations over her were initiated after fans realized the possibility of voice change and started to reflect on who they were really loyal to.

They started with introducing and acknowledging the existence of this original voice model, writing out her name and explaining explicitly she is the original voice performer for Kizuna AI. Some went further to state that "the Kizuna AI we are following is the voice model Nozomi Kasuga". In order not to get Nozomi into trouble, there were also comments asking people "not to mention voice model's name when publicizing the voice model incident" and "use 'no.1' or 'old AI'" instead as a kind of protection<sup>34</sup>.

Plenty of fans have been recognizing Nozomi's contribution in developing Kizuna AI into an attractive virtual celebrity. According to them, voice model has a bigger responsibility than just "reading out the script lines word by word". In their words, "all of Kizuna AI's features attractive to us (audiences) are added by the girl named Nozomi" who "earnestly brings happiness to us". Some fans have stronger sense of emotional attachment and connection to the voice model when they have a longer history of previous viewership. Their relationship with the celebrity and her voice model has been gradually built and deepened, and they come to believe that they "know the persona more intimately and profoundly than others do" and they understand the character and are able to appreciate the voice model's values and motives (Horton and Wohl 1956: 216). They compared Nozomi with other voice models who "might only treat voice performance as work", and concluded that "she is different because she really loves this job". They also acknowledged the efforts she had paid since Kizuna AI's debut and stated that "the approximately 1 million's subscriptions of AIChannel on bilibili.com are enabled by Nozomi".

---

<sup>34</sup> It should be noted that fans are using the name of no.1 and old AI differently. Some used them to refer to the original voice model, while others were referring to the Kizuna AI who was performed by the original voice model.

At the same time, they expressed their opinions on the interrelation between Kizuna AI and Nozomi. To a lot of fans, Kizuna AI's voice model and her appearance constitute Kizuna AI. They consider voice performing as an occupation to ensoul a character, and in Kizuna AI's case, the voice model Nozomi is the soul of Kizuna AI that is more important than her visual identification. Here is an example of a fan's reflection on the distinctiveness of VTuber and voice model's core value:

“(The relationship between voice model and) VTuber is different from that between *seiyū* (voice actors) and ACG (Anime, Comics, Games) characters, vocaloid and Tachie (character portrait with transparent background). In essence, *seiyū* only performs a character's voice and vocaloid's charm comes more from diverse derivatives or re-creation. However, VTuber is different. We cannot say VTuber is equal to voice model, but the conveyance of Aichan's loveliness is actually the expressions of voice model's pleasure, anger, sorrow and joy. She is the core of Aichan, and the most inseparable part.”

Recognizing Nozomi as the key to Kizuna AI, some fans therefore refused to admit no.2, no.3 and no.4 as Kizuna AI and considered Nozomi's version of Kizuna AI irreplaceable as one fan clearly stated: “it's not that we only like Nozomi; it's because only Nozomi can be Kizuna AI”. Some fans also took voice change incident as an opportunity to think about what they were really into, and plenty of them found they were actually more loyal to the voice model. One said “Before I thought I liked Kizuna AI, but now I realize I like Nozomi” while the other reflected “I will continue to be a fan of Kizuna AI if Nozomi is still here. But if Nozomi is rejected, Kizuna AI will be dead to me”. To those fans, the visual manifestation and style of Kizuna AI alone isn't that important and meaningful without the voice model, and the original voice model Nozomi is who actually “has bonding with” them.

#### ***4.2.4 Pushing back through moral practices***

New voice models no.2 and no.3 debuted in Kizuna AI's channel in June 2019 without any previous official announcement. At first, they often performed together with original AI while using exactly the same virtual appearance and saying the same greeting “Hi, I'm Kizuna AI”. However, in the following days, fans felt that there were fewer and fewer videos featuring original AI. The dissent and denial of the company and the two new voice models finally broke

out in August 2019 when two insulting videos were uploaded on Kizuna AI's YouTube channel. One video<sup>35</sup> referred implicitly to the original voice model, and said bullying words: "Can't you do something about that stupid tone of yours?", "aren't you a bit puffed up about being a teensy bit popular", and "don't be so full of yourself because you are cute". The other video<sup>36</sup> displayed Kizuna AI's Ahegao (orgasm face) several times and called it a grimace, which has been deemed by the audiences as very inappropriate for an innocent and adorable character who was even once a travelling ambassador for Japan National Tourism Organization.



Example of two Kizuna AI<sup>37</sup>



Example of Ahegao

Both Chinese and foreign fans were really angry at those two videos for destroying Kizuna AI's image, defaming original voice model and making audiences uncomfortable. Since then, many dislike or anti-fandom of the new voice models no.2 and no.3 emerged in Kizuna AI's video channels on bilibili.com. It should be noted that the majority of those anti-fans still like the original voice model and her version of Kizuna AI, and engaged passionately with the few videos featuring her. However, when it turned to no.2 and no.3's performance of Kizuna AI, the same character Kizuna AI became the object of dislike. Under Kizuna AI's name, one can find very sophisticated media engagement where a fan can hop between like and dislike in regard to different objects and media texts, making it hard to nail down a person's fan identity as s/he can be a fan of Kizuna AI and Nozomi and at the same time an anti-fan of voice models no.2 and no.3.

<sup>35</sup> <https://www.youtube.com/watch?v=QpwQMrNy8D4&t=3s>

<sup>36</sup> <https://www.youtube.com/watch?v=lAqFiE2fugQ>

<sup>37</sup> <https://www.bilibili.com/video/av62650008>



The dislike toward a text can be as strong as the admiration or other affective emotions, producing equally powerful practices and influence (Gray 2005: 841). Anti-fans driven by this dislike are also well motivated and mobilized as fans (Gray 2005: 842). Even though the two videos were not translated and uploaded on Kizuna AI's bilibili channel, Chinese fans (and anti-fans) mobilized themselves to informally find and watch the videos, and they then managed to upload them on bilibili.com<sup>38</sup>. Additionally, they also wrote articles, made videos and left comments to discuss about the oppression and marginalization of the original voice model based on the evidence they gained formally or informally from various sources. Voice model Nozomi's Twitter account was one most frequently used information source. Fans and anti-fans referred to her post "Old AI said on Twitter that she doesn't want to lose to annoying adults" and combined it with other traces to construct a version of text that they felt was accurate enough for them to go against with (Gray 2003: 71). Here is an example of how a Nozomi fan commented to justify the "cause" of their practices:

"Are you not seeing all the inside stories that have been recently dug out, the original team that have been fired and Nozomi's tweets? We are not surmising a dark event all because we haven't seen old AI in videos for a while. A8 said she was busy and tired, but she turned out to be doing home workout, going to comic convention and playing with fish. It's precisely because we care about what old AI is doing, we know all these..."

One can detect the rationale the fan put forward to explain the marginalization of Nozomi by the company and predict the abundant time spent on collecting and reading relevant media texts and actually writing out this long comment to discuss why a given text has made fans like her/him angry (Gray 2005: 840). From the tone, a reader is likely to have the feeling that the author must feel close to Nozomi, as the author paid so much attention to the whole thing, really cared in the way of treating an intimate and important friend and defended firmly the intention to help her.

After settling down the cause for dislike in their anti-fandom, Nozomi's supporters started to engage with Kizuna AI, especially her videos with no.2 and no.3, as a moral text (Gray 2003:

---

<sup>38</sup> One example of the user-uploaded video

<https://www.bilibili.com/video/BV14t411T7tg?from=search&seid=514031774443047593>

71; Gray 2005). They stayed in the text's moral dimension and refused to further interact with its aesthetics or rationality (Gray 2005: 844, 847). Here are two illustrative comments under voice model no.2's ASMR video<sup>39</sup>:

“Without reference to the stance, it will be nice for no.2 to change her visual appearance and be a new VTuber (the same to no.3 and no.4). Unfortunately, she is entrapped by A8 and performs as Kizuna AI. She can only get a bad reputation due to the fact that she impaired original voice model's interest and was involved in insulting her senior.”

“This voice is really nice. I might support you if you changed the visual image. But since you are using the visuals of Aichan, no matter how fabulous your voice is, I'm going to scold you.”

Similar to the *Crash* “refusers” in Barker et al. (2001)'s research, both of the viewers above chose to consider no.2 as a moral text whose morality they didn't approve and thus “infringed on their enjoyment and ability to revel in the aesthetic” of her voice performance (ibid.: 844). Feeling compelled to consume the moral dimension first, the two anti-fans turned their back on a text they would in another context treat with plenty of enthusiasm and applause and deemed it not worthy of further interpretation or compliment (ibid.: 844, 850, 852). Following this idea, there was also deliberate disengagement with videos featuring voice model no.2 and no.3 while giving full attention to videos featuring Nozomi, for instance, only watching and giving “likes” to Nozomi's videos.

As is illuminated above, Kizuna AI's fans “use texts to perform and construct their identities” and actively enjoy this participatory engagement with the text both individually and collectively (ibid.: 854). When they turn into anti-fans, this ability continues, and anti-fans' campaigns, groups or communities can be as organized as their fan counterparts (Gray 2005: 854; Gray 2003: 71). The greatest and most successful collective campaign was a disengagement activity initiated by those who were against the producing company's attitude toward Kizuna AI and her voice models. The initiators asked all kinds of audiences: fans, non-fans and anti-fans to unsubscribe Kizuna AI's official bilibili channel as a resistance to the decisions of Kizuna AI's producing company Activ8 (A8), and an approach to let A8 know their stance and attitude.

---

<sup>39</sup> <https://www.bilibili.com/video/av69390978>

“Unsubscribe the channel if you love her” and “please unsubscribe until old AI is back” were the most frequently used slogans to ask people “not to stop fighting until A8 officially compromises”.

Around the initial call for unsubscribing practice, fans and anti-fans have also been mobilizing across different media platforms to realize their disengagement plan. They formed different kinds of groups outside bilibili.com and included the link to join those groups in the videos’ comments. Some groups were informational and aimed at telling those who were often new fans or non-fans but wanted to know or follow the development of Kizuna AI’s voice model incident. Some communities aimed at recruiting members into original Kizuna AI’s protection group, or “United Front” as they called it, to support original voice model Nozomi, and others were places for “Activ8’s objectors” to organize collective resistance to the company’s decisions.

Their collective activity not only stayed within the Chinese platform but was extended to a global scope. Within discussions on the Chinese platform about what could be the best way to rescue Kizuna AI, participants were also referring to the practices of oversea fans on YouTube, especially Japanese fans who are local consumers of Kizuna AI. They called on Chinese fans to also unsubscribe Kizuna AI’s YouTube Channels: “If you have subscribes aic (A.I.Channel) and aig (A.I.Games) on YouTube, please unsubscribe. Whether it’s useful or not, at least we show our position through actions”, and tried to make those international social media their additional space of communication. One fan reflected and suggested that “if only bilibili responds to it, while YouTube stays fine, A8 will possibility leave it as bilibili only represents China’s domestic fans; (we) should call upon old AI’s fans on YouTube to act together and let A8 see our power”. While another fan asked Kizuna AI fans to mobilize to Twitter to explain the origin and development of the incident to Japanese audiences and make Japanese audiences understand their action was “for the benefit of Kizuna AI’s fans all over the world”.

Having the idea that “the unintelligent AI is disappearing, and only we can save her”, Kizuna AI’s fans have been turning their group chats, videos, articles, comments and both domestic

and global social media platforms into moral spaces that according to them, are not for profit or interest, but are for “impartiality” and “justice” . Guided by morality and responsibility, they considered this incident “relevant to the future development of the whole VTuber industry, for “if A8 is able to take advantage of old AI who is at the top of VTuber industry, the company will dare to take advantage of anyone”.

#### ***4.2.5 Displaying consumer confidence and expectations***

During the process of constituting collectivities in those moral spaces enabled by social media, fans and anti-fans were reflecting on their role in the voice model incident (Couldry & Hepp 2017:170). The faithful and convincing encouragement happening between fans, non-fans and anti-fans are showing the confidence in their power as audiences, and more importantly as consumers in the attention economy of VTuber industry.

For their primary aim in the moral practices, the resistance advocates thought they needed to explain the situation to all new comers and adhere to their principles of letting the original voice model have a voice. They referred to previous oppression cases in other VTuber projects Game Club Project and hololive to state that “things will only get better when audiences speak up”, and “if fans compromise, it means anyone can be the voice model, and the company will have stronger control over voice models”. They had faith in the power of public opinions and believed “as long as we work hard, we should be able to bring old AI back”.

It’s also quite clear that they already recognized their importance to the company and industry which “are waking up to the commercial need to actively listen and respond to” the reception side (Jenkins 2013: 192, cited in Guga 2015: 40). They believed A8 doesn’t dare to run counter to so many fans of old AI, for if audiences don’t take it, A8 won’t be able to make money. They also gave justification to the aforementioned disengagement with no.2’s and no.3’s videos where they actively shaped media flows and made their presence felt (Jenkins 2013: 192, cited in Guga 2015: 40): “the views and clicks for no.2’s and no.3’s videos are not enough to feed 120 people; we will succeed as long as we persist”. The ability to consume has empowered the audiences, especially the passionate fans who often spend more money and time than other

passer-by viewers, and given them confidence in rejecting the disliked texts and asking for what they want from the production side.

Behind all the aforementioned dislike toward new voice models and moral practices against the producing company, there are always people's expectations of ideal forms of a media text and the morality or aesthetics it holds as well as negative examples to be avoided and amended (Gray 2003:73). Their expectation is also related to the values reflected in media consumption, use and meaning in a VTuber genre (Gray 2005: 841). The complex fandom and anti-fandom within Kizuna AI's case reflects audiences' expectation of respect for themselves and the creators represented by the voice model, which can be illustrated by the following two comment examples:

“It's okay even if you want to change voice model, but please announce in advance (like Kurosaki Rin), and don't do it secretly. Whether it's multiple identity configuration (HimeHina), soul changing (Kurosaki Rin) or different characters within the same channel (Eilene family)...all of these can be done fair and square, but A8 executed it in the most disgusting way”.

“The incident is in essence a great disrespect for content creation and audiences' emotions.....Even if we want Aichan to upgrade her personality, we do hope that “each of her is an independent and full existence. It shouldn't be that the portrait image is an empty shell and the voice models are ...its plug-ins. Within ACG creation, even if you are only treating it as a portrait image instead of a 'living presence' and are finding voice performers on this condition, you should know it's self-defeating to attempt to change the (already existing) setting and direction of a virtual work.”

Through the first fan's reference to and comparison with other VTubers' strategies for coping with voice model change or having multiple voice models, one can easily find an explicit demand from the fan to be treated in a more respectful way by the producing company, and to at least have adequate notice before everything happens. The second fan further showed his/her wish for the creative value behind the hidden labour to be prioritized and recognized. S/he wanted each individual personality (no.2,3 or 4) to be fully shaped and engaged, and hoped that they could all contribute to Kizuna AI as an virtual work, developing the visual portrait more than an image to a 'living' existence integrating the personalities from the human voice models.

#### ***4.2.6 Culture as contestation***

The constitution of solidarity is prominent in the multi-social relations among fan communities in the case of this virtual internet celebrity, yet people's ideas are not necessarily unified (Hills 2016: 472, 479). As Zaborowski (2019:107) has stated, "the online fan discussions concerning which character traits were understood to epitomize the performer could result in disagreements, and as in most fandoms, it could get ugly". Conflicting understandings are evident within different shades of fan experiences and imaginative representations of Kizuna AI (Matrix 2006: 8). Similar to vocaloid fans, VTuber audiences' negotiation around Kizuna AI's characteristics, identity and meaning is also a multidimensional and interactive process beyond individual interpretations (Zaborowski 2019:107).

There are different views on the same object. Take the "artificial idiot" name for example, plenty of fans and audiences are enjoying the conversations and playacting around this endearing characteristic, but others dislike the excessive use of this "essentially vulgar" nickname. The name objectors consider those who "follow the trend" to contribute to an overwhelming presence of "artificial idiot" and "lowering intelligence" in the comment area very "annoying" and "disgusting".

When it comes to the voice model incident and multiple AI configuration, one can see predominant support to original voice model and objection to other voice models as is mentioned before. However, the other voice models do have their own fans and audiences, especially no.4, who speaks Chinese and started her own bilibili channel by the end of 2019. Adequate acceptance of those three newer voice models is also evident in audiences' engagement. To some audiences, "they (all the 4 Kizuna AI) are complete when they are together" and there is nothing wrong with the other three. Moreover, not everyone likes the idea of viewing Kizuna AI and Nozomi as one. To some audiences, "Kizuna AI is Kizuna AI, and Nozomi is Nozomi". As one fan has commented "What I love is Kizuna AI. She is both the old AI and new AI. I don't like equating Seiyū (voice performer) behaviour with Kizuna AI's behaviour. That will make me feel as if I insulted Kizuna AI."

There were also dissenting voices regarding the collective moral practices. Some questioned

the intention and credibility of the anti-activity by arguing:

“you disapprove just because A8’s decisions didn’t meet your expectations. Do you really care about what Kizuna AI no.1 is doing? You haven’t seen her live streaming for a while, so you think she is hidden by the company. Making your own logic workable doesn’t mean this is the case.”

Some were against the way others take to demonstrate their attitude: “many people enjoyed swearing in the comment area; they are not really putting forward requests to the operating company”. While others categorized the practices as fans’ common support of their idol and further commented “this kind of worship makes it easy for people to become fanatical and blind”.

It should be noted that the enthusiasm from fans and anti-fans, especially those who are named “extreme fans”, does at times take up too many spaces and resources on various social media platforms. As we have seen before, the devoted fans and anti-fans not only experienced the texts morally by themselves, but additionally “encouraged an avoidance of the aesthetic text in others too” (Gray 2005: 848). Therefore, in Kizuna AI’s bilibili channel, it’s very easy to find “overloaded expectations of the text” that can predetermine and restrict other viewer’s ability of understanding and experiencing it (Barker et al. 2001: 35, cited in Gray 2005: 844). One can find complaints from non-fan audiences who just hoped to freely enjoy the aesthetics of Kizuna AI’s videos, but were compelled by others to experience the morality first. Moreover, one is likely to be attacked if s/he doesn’t comply to the moral standard set by some extremists, as fans of no.4 have commented “many old AI fans were named as spammers, and many passer-by fans were forced to take sides” and “nobody really cared about the videos, they only cared about the sides others take”.

The disputes happening between audiences and fans echo with the push-pull dynamic between audiences and producing company which has been unfolded throughout the thesis, and further verify culture as contestation. One can see in the earlier analysis that the producers are pulling audiences into the exotic and exceptional characterization and storytelling of Kizuna AI as a virtual celebrity integrating both human and non-human qualities (Hill 2016: 754-755; 2019:

4). In audience engagement, different kinds of restless Chinese audiences are actively pushing back to the characterization or identity configuration given by the producers through individual and collective fan practices and disengagement activities, while mobilizing themselves informally to websites not available in the country to resist certain distribution strategy of the operating side (Hill 2016: 755; 2019: 4). The producing company, and different types of fans have contrasting definitions of who is Kizuna AI and what is Kizuna AI. The VTuber category is produced as similarly “unstable, open to contestation by different groups, dynamically related to other contexts, stabilised via institutions (including business institutions)” as other cultural categories (Highmore 2016: 19-20). Kizuna AI’s characteristics and meaning are dynamic and are constantly being negotiated and balanced between both producers and audiences and audiences and audiences.

### **4.3 Kizuna AI as a digital commodity**

Taking a holistic view on Kizuna AI, one will see inevitable traces of commercialization and commodification that is equally evident in human internet celebrities’ struggle under the attention economy of the internet to generate traffic (Abidin 2018: 78-82). Kizuna AI’s “sellability” as a virtual being embodies even more complex interrelations of labour, capital, capitalism and socio-cultural values (Williamson 2016: 160).

#### ***4.3.1 Commodifying labour, performance and the virtual YouTuber***

Williamson (2016: 21-22)’s argument that “the growth of celebrity has occurred primarily for economic reasons” also works in the case of virtual internet celebrity Kizuna AI. When the producing company is creating her media presence, more emphasis can be found on economic value than creative value, as one audience has commented: “Kizuna AI is a money-making commodity for the company”. The company’s commercialization decisions, for instance the irresponsible change of voice model, reflect the coexisting and interpenetrating “creative activity and exploitation within the context of the emerging online economy” (Andrejevic 2008: 25, cited in Hesmondhalgh 2019:335). Those decisions also mirror the extreme inequalities and injustices that are apparent in contemporary capitalist societies (Hesmondhalgh 2019: 10).



In the commercialization process, the virtual internet celebrity is a commodity in multiple senses. First, the hidden labour represented by the voice models, who had long remained unstated and been forbidden to have any public presence related to the virtual character, is a commodity. The labour is bought from the performer and controlled by capital (Williamson 2016: 12). Hidden labour also don't possess the products of their labour as they do not really have much control in the production process (ibid.). Both their labour and the products of their labour belong to the capitalist, not them (ibid.). The outputs of the voice model's labour such as the songs, videos, interviews and live performances are in this sense commodities tailored for consumption. Moreover, virtual YouTuber Kizuna AI, the character and internet celebrity which can be viewed as the final product constructed by voice model's labour and the outputs of her labour is a commodity.

“Sellability” is deemed as the basic value of cultural production rather than the sharing, circulating or participating of human creativity and meaning creation (ibid.:160). In the digital age when consumption is not confined to tangible materials, such ability is also extended to the capability of attracting people's attention and alluring them into the consumption of intangible (media) contents, echoing well with an internet celebrity's need to exhibit “the economic, technical, cultural, and social skills required to maintain one's visibility on the internet to create a sustained and viable business” (Abidin 2018: 115).

The strategy for establishing attractiveness in Kizuna AI is similar to the “double dialectic of empathy and defamiliarization, human and non-human experientiality” that can be applied to non-human narrators in literary work (Bernaerts et al 2014: 69). It connects closely to the human and non-human connotations that have been unveiled at the first part of the analysis. By commodifying nonhumanness and virtuality, Kizuna AI deliberately distances herself from the audiences, using her identity as an artificial intelligence and virtual character to make audiences and fans defamiliarize with her stories, even though she possesses a young female human being's body and face. Audiences can choose to acknowledge the otherness or exoticism of Kizuna AI as a nonhuman narrator, and in this way have a chance to rethink about “some of their assumptions and expectations about human life and consciousness” to achieve a kind of

“non-human experientiality” (ibid.). At the same time, her human-like behaviour and YouTuber positioning to perform and invite social interactions enable audiences to empathize and project human experience onto the AI system that is not “conventionally expected to have that kind of mental perspective (ibid.).

Kizuna AI is therefore both unnatural and familiar, inaccessible and approachable. The attention and popularity drawn by such existence has made her extremely “sellable” in the attention economy of the internet culture. For instance, she has guzzu (goods) available on Comic Market<sup>40</sup>, live concerts, offline fan meetings and transmedia presence in various advertisements. In this case, it’s possible to further consider her fame and popularity as a kind of commodity that could be utilized by advertisers and brands.

#### ***4.3.2 Participating in digital capitalism***

Kizuna AI’s commodification process echoes with Matrix’s idea of pixel vixens as “technosirens” who can draw both male and female users into the participation of digital capitalism (Matrix 2006: 8). Such participation involves more than tangible materials or digital media contents when the fundamentally economic “value in capitalism... is also reliant on specific cultural meanings” (Williamson 2016: 13). The commodification and consumption of Kizuna AI is “enculturating” (Matrix 2006: 8). A “high-tech cultural imaginary” that advocates a “well-connected, mobile, and flexible digital lifestyle” is constantly being built through the consumption and commodification of Kizuna AI (ibid.: 120). What usually existed and happened between human and human, for instance friendship and intimacy, is now possible in the interaction between digital data (non-human existence) and any person, anytime and anywhere. This possibility further blurs the boundary between the real and the imagined, the human and the non-human.

Upon this new kind of human-non-human interaction, this virtual internet celebrity is further promoting and publicizing the “values of digital capitalism”: virtuality, sociality and

---

<sup>40</sup> Japan's largest indoor public gathering and exhibition of self-published publications (doujinshis) centered around manga, anime, video games, and other related genres <https://www.comiket.co.jp/info-a/WhatsEng201401.pdf>

interactivity, all available and achievable through one's consumption (ibid.: 107). As a hollow and virtual body "containing and reflecting ideas, desires, and creative forces of every single one of us" (Guga 2015: 38), Kizuna AI's lack of a real-world existence is giving her audiences bigger agency than a human internet celebrity who obtains a more real identity in private life. Audiences especially fans are empowered to generate different meanings upon her and her performances thanks to the arbitrariness of a virtual identity and personalize their engagement and making Kizuna AI personal.

One can get accompany and sociality by simply watching her videos and interacting with her on the internet. If a fan wants to have deeper interactivity with or more access to Kizuna AI, s/he can buy her material or virtual fan goods such as physical figures or digital album, spend money going to her live performances to experience her on the spot or watch her performance's live broadcasting on YouTube and repeatedly engage with those diverse forms of fan objects. One can even get a deeper sense of connection to Kizuna AI, simply by spending more (time or money) and consuming more frequently those virtual or material commodities. As a cultural and digital commodity, Kizuna AI endorses consumption as a way to construct the fluid identity or remake the self to a more idealized status (Marshall 2014; Williamson 2016: 162). As a virtual internet celebrity, she now becomes a resource for audiences to construct and deepen their human identity through co-creation of human-non-human interaction (Hills 2002: 166, cited in Hills 2016: 467).

## 5. Conclusion

The thesis centered on the primary aim of situating Kizuna AI as an extension of Abidin (2018)'s definition of internet celebrity and examined her diverse and competing definitions, meanings and values to the producing company, voice models and audiences who are constantly negotiating with and enforcing each other from three aspects: the production, the audience and the generic form as a whole.

Through semiological analysis (Rose 2016) of Kizuna AI's videos, the thesis first looked at the production side and found producers are defining Kizuna AI as an adorable and exotic self-claimed artificial intelligence with distinctive visual features from Japanese culture and connotations of both human and non-human qualities. Extending the traditional celebrity-audience relationship theory – para-social interaction theory (Horton and Wohl 1956) to a social media context, the analysis showed that Kizuna AI is constructing a sense of intimacy and authenticity while performing and inviting audience reaction. The video production pulls people into not only the para-social interaction with the celebrity but also the multi-social interaction with other audiences through their voluntary imagination and fantasy.

The thesis then turned to audience side and integrated media engagement (Hill 2016; 2017; 2019) and fan studies (Gray 2003; 2005) to study Chinese audiences' individual and collective reaction to Kizuna AI's performance and characterization through qualitative text analysis of video comments. Drawing on Hill(2016, 2019)'s notion of spectrum of engagement, and Gray(2003, 2005)'s studies of anti-fan and moral text, the thesis argued that viewers, especially fans and anti-fans are engaging with Kizuna AI both affectively and cognitively, and more importantly, they are also engaging with relevant media texts (especially those with voice models) morally. Based on Hill (2017, 2019)'s push-pull dynamics, the thesis found that audiences are not passive information receivers but are actively rejecting dissatisfactory media texts and selecting objects that they are willing to engage with. They push back to producers' decisions on Kizuna AI through individual and collective practices out of confidence in their power as consumers and expectation of more respect for audiences and content creation.

At last, the thesis looked at virtual YouTuber as a holistic generic form and analyzed Kizuna AI as a digital commodity inevitably involving issues of economic logics, labour relations, capital and socio-cultural values. Building on Williamson (2016)'s discussion of celebrity and commodity fetishism, this part considered audiences as consumers and made explicit the virtual internet celebrity's strategy of selling non-humanness and virtuality in this specific form. The thesis further argued that Kizuna AI is promoting a high-tech cultural imaginary with the allure of realizing human-non-human interaction in the celebrity-audience relationship, and values in digital capitalism to realize an idealized self through digital consumption.

Few studies have looked in detail the visual and verbal appeal in virtual YouTuber's videos, the audience reception or the contestations between producers and audiences in China. In general, the thesis contributes to the knowledge of virtual internet celebrity as a specific and emerging type of celebrity centering on the commodification of non-humanness in virtual characters. The thesis' empirical focus on China adds to the understanding of virtual internet celebrity in a special locality. The multi-method design and analytical lens put together both the production and reception of a virtual internet celebrity, acknowledging the significance of voice models in audiences' meaning making of virtual YouTubers and illustrating the existence of push-pull dynamics between audiences and producers within this cultural category. It also adds to the understanding of audiences' aims and reasons for exercising their agency and consumer power in resisting production decisions in the realm of cultural production.

### **5.1 How is Kizuna AI defined and portrayed in her videos? And how does she perform and invite interaction from the audiences?**

Kizuna AI's portrayal in her videos consists of characteristics from pixel vixen in cyberculture, kawaii aesthetics in Japanese culture and moe elements in otaku culture. At first sight, she is a 3-D computer-generated model in the shape of a human being whose virtual body displays evident features of a young and adorable woman, connoting human quality and femininity. The cuteness in her clothing, appearance, expressions and speaking tone echoes with the concept of shōjo and kawaii that are distinctive in Japan. The affective components in her character

configuration and visual style such as personality “alter” shows its connection to moe elements and otaku culture.

In her videos, she defines herself as artificial intelligence, deliberately distancing and alienating herself from human beings. Developing upon this connoted non-human trait, she expands the descriptions and self-claims regarding her AI identity such as the non-human aging and super calculating ability throughout her video contents, and in this way further enhances herself as an exotic existence. Claiming to be (probably) the world’s first virtual YouTuber, Kizuna AI also displays exceptional human-non-human interactions within her video contents, interviews and livestreaming, showcasing a possibility to blur reality and virtuality while building herself into an exceptional internet celebrity.

Constructing intimacy and showcasing authenticity are the two main approaches for Kizuna AI to perform and invite interactions from her imagined audiences. For most of the time, Kizuna AI is keeping direct eye contact and stays up close and personal to audiences in her videos. Frequently used medium and close shots indicate the close distance and connote an illusion of physical proximity. While decontextualized settings offer a sense of openness and arbitrariness for audiences to voluntarily imagine and assign whatever intimate scenarios they want to the para-social interactions implied by the videos.

Time is fully utilized in Kizuna AI’s case to display authenticity. Her immediate expressions, gestures and body movements to contents displayed in the videos, spontaneous responses to live comments and imperfect performances or even mistakes during live streaming are all showing authenticity and contributing to audiences’ real time experience and real time relationship with Kizuna AI in both the virtual space and the real life. This real time relation could further be interpreted as an indicator of real relation with and involvement in the virtual internet celebrity. The thesis argues that Kizuna AI is educating her audiences and fans by displaying selected audience contributions such as livestreaming comments and fan videos which got reacted to as examples and pulling audience into a feeling of hope to turn this basically one-sided interaction into a temporarily dialectical one as the selected examples have

done.

## **5.2 How do audiences engage with Kizuna AI and each other on bilibili.com? And how do they recognize and engage with voice models?**

The thesis displays that Kizuna AI's audiences are far from monolithic (Gray 2003: 68). The most active are fans, who emotionally, critically and morally react to everything around Kizuna AI and can shift to the opposite of fandom and become anti-fans when dissatisfactory media contents appear. Aside from the passionate fans and anti-fans, there are also normal viewers who are non-fans or, in their words, "passers-by" who watch Kizuna AI's videos randomly and participate in fan practices as they wish. Having different levels and modes of engagement, the fans, anti-fans and non-fans constitute Kizuna AI's diverse audiences on bilibili.com.

From bilibili comments, one can find emotional engagement to be the most prominent mode of engagement in audiences' reaction to Kizuna AI's invitations in her videos. In the imaginary para-social relationship with Kizuna AI, audiences, especially fans, show affection and meaning-making practices that are in many ways "parallel to actual social relationships with real ... 'friends,' and 'lovers'" (Caughey 1984: 23, 40, cited in Hills 2016: 468). They are making the video channel into an affective space where engagement with Kizuna AI can provide happiness, connection to other fans, bonding with this virtual internet celebrity and emotional support such as encouragement and accompany that actually affects real life.

Audiences also draw on cognitive thinking while interacting with Kizuna AI on bilibili.com, and they further develop on her invitations or claims in the videos to customize the character's postulation and help co-create the human-non-human interaction. Realizing Kizuna AI's claim of being artificial intelligence is just a fictional branding label, the majority of fans still "accept the gambit offered" and develop a new nickname "Artificial Idiot" based on her AI claim and the not that intelligent performance across her media production. This customized name becomes a distinctive sign of endearment between Kizuna AI and fans, and is enlarged to "something more than just the moment(s) of viewing" through people's interaction with each other in the multi-social media talk (Gray 2003: 76). Moreover, fans are actively playing with

customizations like this on bilibili.com as an approach to showcase and consolidate their insider identity as knowing fans. At the same time, such customizations don't only stay within the audience side; they are actually impacting how Kizuna AI reflect on her features and definition in her various media production.

The push-pull dynamics (Hill 2016; 2019) and contestation between producers and audiences became more evident when the voice model incident happened and two insulting videos were released on YouTube. Since then voice models who had never been officially acknowledged by the company or allowed to make public connection to Kizuna AI on their personal social media became another key focus of audience engagement on bilibili.com. The majority of positive engagement is with the original voice model Nozomi Kasuga, who has been recognized by plenty of fans as the soul of Kizuna AI and deemed more important and meaningful than the character's visual identification. Negative engagement is mostly around new voice model no.2, no.3 and the producing company Activ 8 (A8), and the dislike further develops into anti-fandom of the new voice models in some loyal supporters of Nozomi. In their engagement with the videos, they are treating the virtual internet celebrity and her new voice models as moral texts, which should first satisfy their moral standard before being further reacted or interpreted on an aesthetic or rational-realistic level. No.2 and no.3's performances have been rejected and disengaged because they don't have the moral quality as the prerequisite.

Deliberate disengagement with Kizuna AI's video channels also appears out of love for Kizuna AI and the original voice model and as a way of resistance to the producing company's practices. From the qualitative text analysis of bilibili comments, one can see fans and anti-fans mobilize across domestic and international social media platforms to carry out various individual and collective activities such as community forming and information publicizing to call for attention from both domestic and foreign audiences, fans and non-fans to the collective disengagement initiative. They turn those communication places into moral spaces where they can support the right of the original voice model, protect their beloved virtual internet celebrity and realize fairness and justice.



The thesis argues that behind audiences' push back to producer decisions, one can find a sense of confidence in them about their own power as consumers. They have recognized their importance in the attention economy of the VTuber industry and are actively utilizing it to make their voice heard to content providers who rely much on them for money-making. Additionally, there also lies audiences' expectation of this specific generic form in their reactions to the voice model incident. What they are expecting or demanding from the producing company is more respect to audience emotions and the creative value generated by the invisible voice models to ensoul virtual characters in media production.

Audiences' engagement with each other on bilibili.com are not always positive and cohesive. Disputes and contestations are also evident between audiences when various kinds of fandom co-exist under Kizuna AI's name with distinctive views on the same object. Aside from fans and supporters of original voice model Nozomi, there are also fans of voice model no.2 and no.3 who hold contradictory stance against the aforementioned disengagement practices and are passionately engaging with their own fan objects. Some audiences are just fond of Kizuna AI's visuals and draw a very clear line to demarcate voice models from Kizuna AI. There are also non-fan audiences who are "passers-by" and only want to randomly enjoy the aesthetics of the videos instead of being compelled to judging from morality of a media text.

The push-pull dynamic (Hill 2016; 2019) between audiences and producers echoes with such disputes and proves culture as contestation. Audiences are pulled into the characterisation and appeal of Kizuna AI constructed by the producers in the videos, while they also actively push back through individual and collective fan practices toward dissatisfactory media contents and production decisions. All the dynamics and contestations within Kizuna AI's case illustrate the fluid quality of VTuber as a cultural category. It is a genre where a virtual internet celebrity's definition, characteristics and meaning are in flux and constantly being negotiated and balanced between not only audiences and producers, but also audiences and audiences.

### **5.3 How is Kizuna AI commercialized and consumed as a digital commodity?**

Judging from the whole format, Kizuna AI is inescapable of the commercialization and

commodification process which interweaves labour, capital and product and primarily aims at an internet celebrity's economic value. Hidden labour represented by voice models, the performances generated from their work, and Kizuna AI who is the final product of this production are all commodities. Kizuna AI bears a distinctive kind of "sellability" (Williamson 2016: 160) of the digital time where attention and popularity enabled by the time and labour audiences spend on the interactive celebrity can further be commodified and benefit brand owners or advertisers.

The highlight on human-non-human interaction is the main strategy for Kizuna AI to produce her "sellability" as a digital commodity (ibid.). As an assumed non-human narrator within her media production, she intentionally distances herself by stressing her virtual and AI traits and makes audiences defamiliarize with her presence and stories to trigger "non-human experientiality" in the audiences. However simultaneously, she is looking, behaving and talking like a human being to invite audiences to empathize and project their human experientiality back on her. She constructs an exotic yet familiar existence by pulling audiences into the consumption of her media contents.

By commodifying and consuming Kizuna AI as a digital product, audiences and producers are also building and promoting an enculturating high-tech cultural imaginary, where what used to be available only within human beings can be extended to the realm of human and machine. It is pushing the boundary between the real and the imagined, the human and the non-human by advocating a sense of virtuality, sociality and interactivity. Under such imaginary, audiences are empowered to adopt a well-connected, mobile and flexible way to generate different meanings and personalize their engagement with the virtual internet celebrity thanks to her virtuality and the lack of a self in reality. In their personalization of the engagement realized through purchasing and consuming goods and services of Kizuna AI (for instance buying her fan goods and figures online or at comic markets, listening to her digital album whenever wherever one wants and watching her live performances on the spot or through live broadcasting on YouTube), audiences are further drawn into digital capitalism where they need to remake, construct and deepen their self-identities based on consumption.

## 5.4 Moving forward

For further studies, interviews can be conducted within Kizuna AI's audience groups to generate deeper understanding of individual audience's emotions and practices toward this virtual internet celebrity as this thesis has a key focus on the textualities in the videos and audience comments and has prioritized discussions, disputes and relations between different audiences over individual thoughts. It would also be fruitful to combine this research with production interviews of voice models or the producing company, as the investigation into the production side in this thesis is primarily textual. Weaving in production interviews is very likely to bring in insider knowledge that is not often available to the public, and thus benefit the study of not only Kizuna AI, but also virtual YouTuber industry. The thesis's primary focus on China can complement the understanding of Kizuna AI's reception across the globe. It can also be a starting point to investigate China's VTuber (or Vup<sup>41</sup>) ecology on bilibili.com or the whole virtual internet celebrity industry.

---

<sup>41</sup> Up is the way bilibili.com names its uploaders or livestreamers. Up to bilibili.com is like YouTuber to YouTube. In China there is also a Vup trend on bilibili.com in a similar form as VTuber.

## 6. References

- Abidin, C. (2016) 'Visibility labour: Engaging with Influencers' fashion brands and #OOTD advertorial campaigns on Instagram', *Media International Australia (8/1/07-current)*, 161(1), pp. 86-100.
- Abidin, C. (2018) *Internet celebrity: understanding fame online*. First edition. Bingley: Emerald Publishing Limited.
- Aziz, F. (2019) 'Performing as a Transgressive Authentic Microcelebrity: The Qandeel Baloch Case' In: Abidin C. & Brown M. (eds) *Microcelebrity Around the Globe: Approaches to Cultures of Internet Fame*. Bingley, UK: Emerald Publishing, pp. 131-144.
- Azuma, H. (2009) *Otaku: Japan's Database Animals*. Translated by Abel J.E. and Kono S. Minneapolis: University of Minnesota Press.
- Bazeley, P. (2013) *Qualitative data analysis: practical strategies*. London: SAGE.
- Bernaerts, L. et al. (2014) 'The Storied Lives of Non-Human Narrators', *Narrative*, 22(1), p. 68. Ohio State University Press.
- Berndt, J. (2019) 'Introduction: Shōjo Mediations'. In: Berndt J., Nagaike K. & Ogi F. (eds) *Shōjo across media: exploring 'girl' practices in contemporary Japan*. Switzerland: Palgrave Macmillan, pp. 1-21.
- Bilibili.com. (2019). 关于我们 - 哔哩哔哩 (About US). [online] Available at: <<https://www.bilibili.com/blackboard/aboutUs.html>> [Accessed 15 May 2020].
- Black, D. (2012) 'The Virtual Idol: Producing and Consuming Digital Femininity'. In: Galbraith P.W. & Karlin J.G. (eds) *Idols and Celebrity in Japanese Media Culture.*, London: Palgrave Macmillan, pp.209-228.
- Booth, N. & Matic, J.A. (2011), 'Mapping and leveraging influencers in social media to shape corporate brand perceptions', *Corporate Communications: An International Journal*, vol. 16, no. 3, pp. 184–191
- Brown, M. & Phifer, H. (2019) 'The Rise of Belle from Tumblr' In: Abidin C. & Brown M. (eds) *Microcelebrity Around the Globe: Approaches to Cultures of Internet Fame*. Bingley, UK: Emerald Publishing, pp. 121-130.
- Burr, V. (2015) *Social constructionism*. 3rd edition. London: Routledge.
- Charmaz, K. (2006) *Constructing grounded theory: a practical guide through qualitative*

*analysis*. London: SAGE.

- Cho, J.K. (조종학) (2019) 차세대 브이튜버가 콘텐츠에 미치는 영향(The impact of next-generation VTuber on content). *한국디자인리서치 (Korean Design Research)*, 4(2), pp. 173-182.
- Click, M., Lee, H. and Holladay, H. (2013) 'Making Monsters: Lady Gaga, Fan Identification, and Social Media', *Popular Music & Society*, 36(3), pp. 360–379.
- Couldry, N. and Hepp, A. (2017) *The mediated construction of reality*. Cambridge, UK : Polity Press.
- Duffy, B. E. (2016) 'The romance of work: Gender and aspirational labour in the digital culture industries', *International Journal of Cultural Studies*, 19(4), pp. 441–457.
- Edosomwan, S. et al. (2011) 'The history of social media and its impact on business', *Journal of Applied Management and Entrepreneurship*, 16, pp79-91.
- Flick, U. (2007) *Designing qualitative research*. Sage Publications (The Sage qualitative research kit).
- Flyvbjerg, B. (2001) *Making social science matter: why social inquiry fails and how it can succeed again*. Cambridge: Cambridge University Press.
- García-Rapp, F. and Roca-Cuberes, C. (2017) 'Being an online celebrity: Norms and expectations of YouTube's beauty community', *First Monday*, 22(7).
- Gibbs, M. et al. (2015) '# Funeral and Instagram: death, social media, and platform vernacular', *Information, Communication & Society*, 18(3), pp. 255–268.
- Gray, J. (2003) 'New audiences, new textualities', *International Journal of Cultural Studies*, 6(1), p. 64. doi: 10.1177/1367877903006001004.
- Gray, J. (2005) 'Antifandom and the Moral Text: Television Without Pity and Textual Dislike', *American Behavioral Scientist*, 48(7), pp. 840–858.
- Gray, J. and Murray, S. (2016) 'Hidden: Studying media dislike and its meaning', *International Journal of Cultural Studies*, 19(4), pp. 357-372.
- Guga, J. (2015) 'Virtual Idol Hatsune Miku'. *Arts and Technology. ArtsIT 2014. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 145. Springer, Cham. doi: 10.1007/978-3-319-18836-2\_5

- Guo, M. (2019) The Production and Receptive Psychology of Virtual Idols. *Media Criticism*. 2019 (0).
- Halliday, M. A. K. (1985) *An Introduction to Functional Grammar*. London: Arnold
- Hansen, A. and Machin, D. (2013) *Media and communication research methods*. Red Globe Press.
- Harrington, C., Sandvoss, C. & Gray, J. (2017) *Fandom: Identities and Communities in A Mediated World, Second Edition*. New York: New York University Press.
- Hesmondhalgh, D. (2019) *The Cultural Industries* (Fourth Edition), London: Sage.
- Highmore, B. (2016) *Culture*. London: Routledge.
- Hill, A. (2016) 'Push–Pull Dynamics', *Television & New Media*, 17(8), pp. 754–768.
- Hill, A. (2017) 'Reality TV Engagement: Producer and Audience Relations for Reality Talent Shows', *Media Industries*, 4(1).
- Hill, A. (2019) *Media experiences: engaging with drama and reality television*. London; New York: Routledge.
- Hill, A. & Turnbull, S. (2017) 'Nordic Noir' In: Pontell H.N. (ed) *Oxford Research Encyclopedia of Criminology and Criminal Justice*. Oxford: Oxford University Press, pp. 1-21.
- Hills, M. (2016) 'From Para-social to Multisocial Interaction: Theorizing Material/Digital Fandom and Celebrity'. In: Marshall P.D. & Redmond S.(eds) *A Companion to Celebrity*. Chichester, UK: John Wiley & Sons, pp. 463-482.
- Horton, D. & Wohl, R. (1956) Mass Communication and Para-Social Interaction, *Psychiatry*, 19:3, pp. 215-229.
- Hubble, C. (2018) 'Miquela Sousa and the rise of fake influencers', *Eureka Street*, 28(13), pp. 20–22.
- Ishita, S. (2015) 'Construction of the Public Memory of Celebrities: Celebrity Museums in Japan'. In: Marshall P.D. & Redmond S.(eds) *A Companion to Celebrity*. Chichester, UK: John Wiley & Sons, pp.135-154.
- Ito, M. (2012) 'Introduction'. In: Ito, M., Okabe, D. and Tsuji, I.(eds) *Fandom Unbound: Otaku Culture in A Connected World*. New Haven: Yale University Press, pp. xi-xxxii.

- Jerslev, A. (2016) 'In the Time of the Microcelebrity: Celebrification and the YouTuber Zoella', *International Journal of Communication*, 10, pp. 5233-5251.
- Kijima, Y. (2012) 'The Fighting Gamer Otaku Community: What Are They "Fighting" About?'. In: Ito, M., Okabe, D. and Tsuji, I.(eds) *Fandom Unbound: Otaku Culture in A Connected World*. New Haven: Yale University Press, pp. 249-274.
- Korenaga, R. (2019) Criticism of Media Representation and Social Criticism in Action: With Respect to Gender Representation. *社会学研究科年報(Annual Report of Graduate School of Sociology)*, 26, 7-18.
- Kuckartz, U. (2014) *Qualitative text analysis: a guide to methods, practice and using software*. Los Angeles: SAGE.
- Lam, K. (2016) 'The Hatsune Miku Phenomenon: More Than a Virtual J-Pop Diva'. *Journal of Popular Culture*, 49(5), p. 1107-1124.
- Lana, L. (2019) 'Catarina, A Virgin for Auction: Microcelebrity in Brazilian Media' In: Abidin C. & Brown M. (eds) *Microcelebrity Around the Globe: Approaches to Cultures of Internet Fame*. Bingley, UK: Emerald Publishing, pp. 107-120.
- Marshall, P. D. (2010) 'The promotion and presentation of the self: celebrity as marker of presentational media', *Celebrity Studies*, 1(1), pp. 35-48.
- Marshall, P. D. (2014) *Celebrity and power: Fame in contemporary culture*. Minneapolis, Minn: University of Minnesota.
- Marshall. P.D. (2016) 'Introduction' In: Marshall P.D. & Redmond S.(eds): *A Companion to Celebrity*. Chichester, UK: John Wiley & Sons, pp.15-19.
- Marwick, A. (2013) *Status Update: Celebrity, Publicity and Branding in the Social Media Age*. New Haven: Yale University Press.
- Marwick, A. (2016) 'You May Know Me from YouTube: (Micro-)Celebrity in Social Media'. In: Marshall P.D. & Redmond S.(eds) *A Companion to Celebrity*. Chichester, UK: John Wiley & Sons, pp.333-350.
- Marwick, A. (2019) 'The Algorithmic Celebrity: The Future of Internet Fame and Microcelebrity Studies' In: Abidin C. & Brown M. (eds) *Microcelebrity Around the Globe: Approaches to Cultures of Internet Fame*. Bingley, UK: Emerald Publishing, pp. 161-169.
- Matrix, S. (2006) *Cyber Pop Digital Lifestyles and Commodity Culture*. New York: Routledge.

- Michel, P. (2018). *Kizuna AI's NHK appearance sparks debate on social media* | The Japan Times. [online] The Japan Times. Available at: <<https://www.japantimes.co.jp/news/2018/10/13/national/media-national/kizuna-ais-nhk-appearance-sparks-debate-social-media/>> [Accessed 1 Nov. 2019].
- Miles, M. B. & Huberman, A. M. (1994) *Qualitative data analysis: an expanded sourcebook*. 2<sup>nd</sup> Edition. Thousand Oaks, CA: Sage.
- Moden, M. (2019) 'A Dream Dress for Girls: Milk, Fashion and Shōjo Identity' In: Berndt J., Nagaike K. & Ogi F. (eds) *Shōjo across media: exploring 'girl' practices in contemporary Japan*. Switzerland: Palgrave Macmillan, pp.209-231
- Nagata, K. (2018). *Japan's latest big thing: 'virtual YouTubers'* | The Japan Times. [online] The Japan Times. Available at: <<https://www.japantimes.co.jp/news/2018/07/17/national/japans-latest-big-thing-virtual-youtubers/#.XbrgBZozbZs>> [Accessed 23 Oct. 2019].
- Page, R. (2012) 'The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags', *Discourse & Communication*, 6(2), p. 181-201.
- Pellitteri, M. (2018) 'Kawaii Aesthetics from Japan to Europe: Theory of the Japanese "Cute" and Transcultural Adoption of Its Styles in Italian and French Comics Production and Commodified Culture Goods', *Arts*, 7(3), p. 24. doi: 10.3390/arts7030024.
- Puspitaningrum, D.R. & Prasetio, A. (2019) 'Kizuna Ai's "Virtual Youtuber" phenomenon among Japanese Popular Culture Enthusiasts in Indonesia'. *MediaTor*, 12 (2), 128-140.
- Roland, B. (1977) *Image, Music, Text*. Translated by Heath S. New York: Hill and Wang.
- Rose, G. (2016) *Visual methodologies: an introduction to researching with visual materials*. 4<sup>th</sup> edition. London: Sage.
- Sanders, L. (2005) 'Visual Ephemerality: Idoru and Evangelion, Popular Visual cultures in Japan' In Heusser M. et al (eds) *On Verbal/Visual Representation*. Amsterdam; Atlanta, Ga. : Rodopi, pp. 137-150.
- Sandvoss, C. (2005) *Fans: the mirror of consumption*. Oxford : Polity.
- Senft, T. (2008) *Camgirls: Celebrity & Community in the Age of Social Networks*. New York: Peter Lang Publishers.
- Storey, J. (2015) *Cultural theory and popular culture: an introduction*. Seventh edition. London: Routledge.



Tsuji, I. (2012) 'Why Study Otaku? A Social History of Imagination' In: Ito, M., Okabe, D. and Tsuji, I.(eds) *Fandom Unbound: Otaku Culture in A Connected World*. New Haven: Yale University Press, pp. 3-29.

Van Dijck, J. (2013) *The Culture of Connectivity: A Critical History of Social Media*. New York: Oxford University Press.

Williamson, M. (2016) *Celebrity: capitalism and the making of fame*. Cambridge: Polity.

Zhang, Y. (2020) 'Adopting Japanese in a Popular Chinese Video-Sharing Website: Heteroglossic and Multilingual Communication by Online Users of Bilibili.com', *International Multilingual Research Journal*, 14(1), pp. 20–40.

Zaborowski, R. (2016) 'Hatsune Miku and Japanese Virtual Idols'. In Rambarran, S. and Whiteley, S. (eds): *The Oxford handbook of music and virtuality*. New York: Oxford University Press, pp. 112-128.

Zaborowski, R. (2019) 'Fans Negotiating Performer Personas: "Melt" by ryo feat. Hatsune Miku', *Suomen Antropologi*, 43(2), pp. 104–108.

## 6.1 Empirical material

### 1. ASMR

AIChannel官方(AIChannel Official) (2019). [video]. **【情景短剧】**年上女友的工作慰问 ( **【Scene play】** *From your older girlfriend to you, who is tired from work*). bilibili. Available: <https://www.bilibili.com/video/av69390978>

AIChannel官方(AIChannel Official) (2017). [video]. **【熟肉】**A.I.Channel#135ASMR 虚拟的耳骚! ( **【Translated】** *A.I.Channel#135ASMR A virtual ear cleaning!*). bilibili. Available: <https://www.bilibili.com/video/av11949555>

### 2. Daily - storyline performance

AIChannel官方(AIChannel Official) (2017). [video]. **【自我介绍】**A.I.Channel#01 初次见面! 我叫绊爱! ( **【Self-Introduction】** *Nice to meet you! I'm Kizuna Ai*). bilibili. Available: <https://www.bilibili.com/video/av9800170>

AIChannel官方(AIChannel Official) (2019). [video]. **【猜谜】**你能猜出来吗? ( **【Quiz】** *Can you answer this?*). bilibili. Available: <https://www.bilibili.com/video/av62650008>

### 3. Gaming

AIChannel官方(AIChannel Official) (2019). [video]. 卡比原来是这种接吻狂魔吗www **【星之卡比新星同盟】** (*Kirby kisses a lot(lol) 【Kirby Star Allies】*). bilibili. Available: <https://www.bilibili.com/video/av62515008>

AIChannel官方(AIChannel Official) (2018). [video]. **【崩坏3rd】**说到AI的AI酱、是我对吧! ? ( **【Houkai 3rd】** *Speaking of AI whose name is Aichan, it should be me, right?* ). bilibili. Available: <https://www.bilibili.com/video/av19949864>

#### 4. Livestreaming

AIChannel官方(AIChannel Official) (2018). [video]. **【LIVE】**我、准备好玩这个了。**【掘地球升】** ( **【LIVE】** *I'm ready to play this* **【Getting Over It】** ). bilibili. Available: <https://www.bilibili.com/video/BV11W411p73F>

AIChannel官方(AIChannel Official) (2019). [video]. **【live】**summer sonic 开幕! 紧急直播! ( **【live】** *Opening of summer sonic! Emergent livestreaming!* ). bilibili. Available: <https://www.bilibili.com/video/av64057945>

#### 5. Music

AIChannel官方(AIChannel Official) (2019). [video]. **【翻唱】**米津玄師「海の幽霊」*Spirits of the Sea* ( **【Cover】** *Kenshi Yonezu's "Spirits of the Sea"* ). bilibili. Available: <https://www.bilibili.com/video/av61931424>

AIChannel官方(AIChannel Official) (2019). [video]. *Kizuna AI - AIAIAI (feat. 中田ヤスタカ)* **【Official Music Video】** (*Kizuna AI - AIAIAI (feat. Yasutaka Nakata)* **【Official Music Video】** ). bilibili. Available: <https://www.bilibili.com/video/av47832767>

#### 6. Reaction

AIChannel官方(AIChannel Official) (2019). [video]. 来和胖达贴贴吧! (*Meeting with pandas!*). bilibili. Available: <https://www.bilibili.com/video/av81308218>

AIChannel官方(AIChannel Official) (2019). [video]. 奇迹的双重「花—Q」!!! **【粉丝投稿介绍】** (*A miraculous synchronization of "F\*\*K You"!!!* **【Everyone's video contributions】** ). bilibili. Available: <https://www.bilibili.com/video/av44370555>

#### 7. Real world interaction with human

AIChannel官方(AIChannel Official) (2019). [video]. **【SUMMER SONIC2019】** 官方艺人采访 vol.04 **【Alan Walker】**(**【SUMMER SONIC2019】** *Official Artist Interview vol.04* **【Alan Walker】** ). bilibili. Available: <https://www.bilibili.com/video/BV1w441167Wt>

AIChannel官方(AIChannel Official) (2017). [video]. **【参加C93?】** 期待已久的“那个”终于要出了! ? **【A.I.Channel】** #245 ( **【Participating in C93?】** *The long-awaited “that” is finally showing up!?* ). bilibili. Available: <https://www.bilibili.com/video/av17814976>

## 7. Appendix

### 7.1 Appendix A – Semiological Analysis

#### 7.1.1 Scheme

Denotation – meaning/form - What a video/sequence depicts?		Connotation - concept
Spoken text+sound	Visuals around Kizuna AI	
<p>Halliday’s (1985) notion of speech acts. Halliday claims that when we speak we do one of four basic things:</p> <p>1) Offer information (e.g. ‘I am busy today’)</p> <p>2) Offer services and goods (e.g. ‘Do you want any help?’)</p> <p>3) Demand information (e.g. ‘Is this broken?’)</p> <p>4) Demand goods and services (e.g. ‘Could you help?’) (Hansen &amp; Machin 2013: 181)</p>	<p><b>representations of bodies:</b></p> <ul style="list-style-type: none"> <li>● age.</li> <li>● gender.</li> <li>● race.</li> <li>● hair.</li> <li>● body. Fat? Thin?</li> <li>● size. Which is bigger?</li> <li>● looks.</li> </ul>	<ul style="list-style-type: none"> <li>● What is the age of the figures in the photograph meant to convey? Innocence? Wisdom? Senility?</li> <li>● Stereotyped images of masculinity and femininity? Race? - cultural constructions of race, gender and beauty?</li> <li>● Stereotyped images of anime/otaku culture?</li> <li>● Hair to signify seductive beauty or narcissism?</li> <li>● Think/fat - represented as undesirable and unattractive?</li> <li>● Which body parts are displayed?</li> <li>● Bigger – more important?</li> </ul>
	<p><b>representations of manner</b></p> <ul style="list-style-type: none"> <li>● expression. (Who is) happy, haughty, sad...? What facial and other expressions are used to convey this?</li> <li>● eye contact. Who is looking at whom (including you) and how? Analyzing Gaze – (not) look at the viewer, look off frame?</li> <li>● pose. Who is standing and who is prone? <ul style="list-style-type: none"> <li>• <i>To what extent do they take up space or not?</i></li> <li>• <i>Is the body open or closed?</i></li> <li>• <i>Is the body aligned and controlled or liberated?</i></li> <li>• <i>Is there an emphasis on relaxation or intensity?</i></li> <li>• <i>Is there a sense of comfort or discomfort?</i></li> </ul> </li> </ul>	

	<ul style="list-style-type: none"> <li>• <i>Is angularity or curvature emphasised?</i></li> <li>• <i>Do they appear to lean/move towards or away from the viewer?</i></li> </ul>	
<b>Sound</b> Environmental Speech Music	<b>representations of activity</b> <ul style="list-style-type: none"> <li>● touch. Who is touching what, with what effects?</li> <li>● body movement. Who is active and who passive?</li> <li>● positional communication. What is the spatial arrangement of the figures?</li> <li>● positioning the viewer <ul style="list-style-type: none"> <li>• Angle of interaction. Vertical angle? Horizontal angle? Oblique angle?</li> <li>• Proximity and interaction <ul style="list-style-type: none"> <li>◆ Distance</li> <li>◆ ‘size of frame’ (Close Shot, Medium Shot, Long Shot, etc.).</li> </ul> </li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>● Who is positioned as superior and who inferior?</li> <li>● Who is intimate with whom and how?</li> <li>● Sense of involvement/detachment?</li> <li>● Sense of physical intimacy?</li> </ul>
	<b>props and settings</b> <ul style="list-style-type: none"> <li>● props. Objects that have particular cultural significance?</li> <li>● settings. Normal? Exotic? Fantasies?</li> <li>● The size of the space in which the participants are positioned.</li> <li>● Color coordination. saturated/flattened color?</li> <li>● Light</li> </ul>	<ul style="list-style-type: none"> <li>● The cultural meaning of particular objects? Spectacles-intelligence, golden light-tranquility...</li> <li>● What effects does its setting have?</li> <li>● Decontextualized to symbolise an idea or a concept?</li> <li>● A feeling of cleanness, order and simplicity?</li> <li>● Emotional intensity/ reserve and moderation?</li> </ul>

### 7.1.2 Example of one semiological analysis

I first coded all the videos separately in Chinese and started using English when coming up with categories and themes. Here is an example of semiological analysis translated from Chinese.

Video name: 【自我介绍】A.I.Channel#01 初次见面！我叫绊爱！ / 【Self-Introduction】

Nice to meet you! I'm Kizuna Ai

Year: 2017

Available: <https://www.bilibili.com/video/av9800170>

Denotation – meaning/form - What a video/sequence depicts?		Connotation - concept
Spoken text+sound	Visuals around Kizuna AI	
<p><b>1) Offer information:</b> [self-definition] two dimensional, virtual, YouTuber, different from human beings, want to be celebrity; [interaction] curious, want to have more communication with audience</p> <p>-Nice to meet you. My name is Kizuna AI. - [You are different from normal youtuber]. If you think this way, you have sharp eyes. -I’m actually two dimensional. -hmm...I’m 3D, so I’m three dimensional? Anyway, I’m virtual. -It’s cool to be called virtual YouTuber, isn’t it? -In this sense, I’m a little different from you guys. -Therefore, I’m very curious about you humans. -I want to know different things and get along with well you. -With this intention, I started YouTube. -And I want to be in a CM someday. -Popular YouTubers can be in CM, right? -I want to be a celebrity like that and I want to have more communication with you. They should be my right directions.</p>	<p><b>representations of bodies:</b></p> <ul style="list-style-type: none"> <li>● Age: have the appearance of a teenage girl</li> <li>● Gender: obvious female features</li> <li>● Race: pale skin, green eyes and brown hair, hard to pin down</li> <li>● Hair: long brown hair with pink highlight; always wear a heart shaped pink hairpin (which is part of the identification of Kizuna AI)</li> <li>● Body: relatively slender and shapely</li> <li>● Size: only upper body (above waist) is shown in the video</li> <li>● Looks: a cute green-eyed girl with brown hair (eyes big enough to remind people of anime style); pink hair pin corresponds to the pink colour of her clothes; white skin; sailor suit style sleeveless shirt; big bow knot in the collar; white oversleeves on the arms decorated with black lace</li> </ul> <p><b>representations of manner and activity</b></p> <ul style="list-style-type: none"> <li>● Expression: cute and bright smiles, confusing facial expressions when questioning whether she is 2D or 3D</li> <li>● Eye contact + analyzing gaze: looking directly at the lens; keeping direct eye contact with audiences in front of the screen; point directly to the lens (screen)</li> <li>● Action (body movement): only waist up is exposed; main</li> </ul>	<ul style="list-style-type: none"> <li>● Bishojo look, wearing and way of speaking - liveliness and youth</li> <li>● Stereotyped images of anime culture? Character design has typically ACG: strange eye color, extremely big eyes, pink in hair – not a human?</li> <li>● Thin/fat - represented as undesirable and unattractive? : slim body with curves; commonly accepted beauty standard – myth: fat body is not as well accepted and recognized as this slim body</li> <li>● Are those looks submissive, coy, confrontational? : consistent appearance – sense of shojo in expressions and clothing</li> </ul> <ul style="list-style-type: none"> <li>● Innocence and friendliness</li> <li>● Free and unfettered feeling</li> <li>● Shows shyness at first sight – sense of shojo – kawaii and authenticity (draw on human experientiality)</li> <li>● What do the poses convey about distance, interaction, social position, etc.?: gestures and movements are used to display her own mood: nervous at the beginning, excited when expecting the future; used to bridge</li> </ul>

<p><b>2) Offer services and goods:</b> futuristic entertainment and enjoyment</p> <p>-In order to let you enjoy things like livestreaming, and somewhat futuristic entertainment like VR, I will keep trying new challenges.</p> <p><b>3) Demand information:</b> ask questions to imaginary audiences</p> <p>-Can you see me? -Can you hear me?</p> <p><b>4) Demand goods and services:</b> ask for support and feedback from audiences</p> <p>-Please kindly give your guidance. -If there is anything I can do or anything you want me to do, please just request it. -Please support my AI Channel.</p>	<p>movements: head and upper body shows from below the camera lens, hands waving to the camera while saying hello; put hands behind the ears, tilt to the side of the screen and ask if you can hear clearly (a feeling of waiting for audience's answer); when speaking of "you", point to the camera/audiences in front of the screen; shake from left to right to (doubt and confusion, and nervousness?); arm stretching from time to time (no fixed meaning); raise both hands and arms when saying the dream is to be in CM (ambitious expectation); step back with hands behind her backs (sense of shojo)</p> <ul style="list-style-type: none"> <li>● Pose <ul style="list-style-type: none"> <li>• <i>Relatively open body posture; at the beginning actions are more cautious and gradually relax; raise arms and stretch openly both arms at the end</i></li> <li>• <i>Shake from side to side</i></li> <li>• <i>Generally relaxing, no emphasis</i></li> <li>• <i>Sense of comfort from casual movements</i></li> <li>• <i>Lean more to the audience side, slightly look up to the audiences</i></li> </ul> </li> <li>● positional communication: Kizuna AI is the main character, taking all focus. She is situated in the center of the screen with blank background. The character takes approximately 1/3 of the whole 16:9 horizontal screen.</li> </ul>	<p>communication between audience and Kizuna AI: inviting audiences to participate in interacting with her. Sense of authenticity, inviting and welcoming invitation, sense of immersion: Kizuna AI is talking to me, she is inviting me to partake</p> <ul style="list-style-type: none"> <li>● Who is intimate with whom and how? : closeness between Kizuna AI and audiences, similar to introducing oneself to human friends, no sense of distance or inequality has been stressed; give audiences a sense of "Kizuna AI want to connect to you", focus on sense of involvement and participation, not control</li> <li>● Sense of involvement/detachment? Kizuna AI's eye contact gives sense of involvement: she is waiting for my answer (gives audiences great space for imagination)</li> </ul>
---	--	--

	<ul style="list-style-type: none"> <li>● positioning the viewer <ul style="list-style-type: none"> <li>• Angle of interaction: in general eye level, occasional head down (Kizuna AI) ; slightly overlooked angle (audiences)</li> <li>• Proximity and interaction: waist-up medium shot; sense of close distance</li> </ul> </li> </ul>	
<p><b>Sound -music</b></p> <p>The entire video is accompanied by soft background music without lyrics. It gives people a joyful and rhythmic sense of feeling to supplement the content of the video and make the content richer (no significant meaning)</p>	<p><b>props and settings</b></p> <ul style="list-style-type: none"> <li>● Settings: no other elements, only a white background – studio? deliberate decontextualization</li> <li>● Color coordination. saturated/flattened color? No color in the background; Kizuna AI has (pure and fresh) colour: pink, brown, black and white; not dazzling, but audiences can focus on her</li> <li>● Light – natural lighting, shiny hair, general studio effect with plenty of light</li> </ul>	<ul style="list-style-type: none"> <li>● Clothing and accessories: kawaii and fresh Japanese style</li> <li>● Setting puts all focus on Kizuna AI: she is the only one who needs attention; she is the only leading role in the video</li> <li>● A feeling of cleanness, order and simplicity? : vitality, happiness and full of energy; friendly and welcoming invitation to connect with her</li> <li>● Emotional intensity/ reserve and moderation?: relaxing; no sense of burden</li> </ul>

### 7.1.3 Results of semiological analysis

Characterization and self-definition	Japanese culture	Kawaii (cute)
		Bishojo (beautiful young female)
		Energetic
		Multiple personalities (Alter as Black AI)
	Cyber culture	3 years old
		Super AI (perfect, genius, smart, intelligent)
		Different from human beings
		Two dimensional
	ぽんこつ (a piece of junk)	
	Multiple Kizuna AI (different voice models)	
(probably) the 1 <sup>st</sup> virtual youtuber		
Performed interaction	Intimacy	

	Authenticity
	Openness
	Arbitrariness
	Sense of involvement
	Expectation and invitation
	Virtual-real interrelation
	Responsiveness



## 7.2 Appendix B – Qualitative text analysis

### 7.2.1 Coding process

The open coding is carried out in Chinese and the categorization is done in English. Repetitive or not meaningful comments have been ruled out during open coding. Each video is separately coded and finally combined to a whole sheet.

D	E	F	G
comment	descriptive code		category
397 确实-因为人家是天才嘛!	爱管自己叫天才		
398 你不脸红吗(← △ )-			
399 回复 @吃炸炸炸炸炸:人家是AI完全不会脸红????	因为是AI所以完全不会脸红		
400 对呀, 对呀(“”)可惜是个人zz...算是说智能啦,( ∇ )	可惜是人工zz/人工智能		
401 AI球			
402 build genius			
403 你他娘的可真是个Genius			
404 苏响AI	苏响AI		
以前净化心灵	以前净化心灵, 现在智减神器		
405 现在智减神器			
406 降智打击	降智打击		
407 然而现在智也降不了[大笑][大笑][大笑]初号啊	现在智也降不了		
408 现在直接换人	现在直接换人		
409 降智神器没有了, 爱管不回来, 永远不回来, A8Biss	降智也没有了, 爱管不回来永远不回来		
想想, A8到时候看到事情真的闹大了, 会不会然后再让老爱回来, 继续赚钱, 还不是照样拿	A8看事情闹大了不会让老爱回来继续赚钱; 还不是照样拿粉丝当工具, 不知道		
粉丝当工具, 谁知道以后还会不会有其它的理由把老爱赶走, 还是永远不回来了, 事情到	以后会不会有其他理由把老爱赶走; 还是永远不回来了, 事情到这一步已经结束了		
410 这步已经结束了, 挺心疼老爱的	挺心疼老爱的		
411 现在莫得降智了 (心情复杂.jpg)			
412 默默地取关, 默默地留下4号机, 默默地一等她[大笑]	取关留下4号机; 等老爱		
回复 @阿莱娜亚 相信老爱会回来的(“; ω; )相信a8不敢和这么多老爱的老粉对着干的(“;	相信老爱会回来; a8不敢和这么多老爱粉丝对着干		
413 ω; )			
414 回复 @惠家小墨 如果老爱回来我估计还是会支持的, 就算a8再高生老爱还是老爱明qwq	老爱回来还是会支持; A8高生		
415 回复 @初雪之言 哇的一声哭出来[2233娘_大笑]			
416 我宁愿降智啊[大笑]	就算降智也喜欢看真正的爱管花q啊qwq		
417 回复 @苏酥酥洛七 就算降智也喜欢看真正的爱管花q啊qwq	就算降智也喜欢看真正的爱管花Q		
418 现在“隐退”幕后(“; ∇; )			
回复 @踩点小王子Drapple 现在也就4号大家还乐意看, 但是2号3号的播放量越来越少, 粉	现在就4号大家还乐意看, 23号的播放量越来越少		
419 样也即将突破80万			
420 回复 @轮的_SL 现在四号都快没人看了[doge]	四号都快没人看了		
421 确实, 你知道吗, 现在的你已经可以自己开演唱会了哦, 曾经只是梦想上cm的你现在已经变	爱管可以自己开演唱会了; 曾经梦想上cm现在已经成为人人知晓的超级ai了		
422 代价就是智商没了( ∇ )	代价就是智商没了		
423 那么, 古尔丹, 代价是什么呢?			
424 回复 @爱管管: 把你的智商作为代价吧[娘音娘_滑稽]	把智商作为代价		
425 回复 @战神小早川 降智打击			
426 看到这条我哭死[委屈][大笑]			
427 现在看扎心了	用自己的一切力量去和别人宣传垃圾a8, 保护一号机		
回复 @本杰明唐格堡 我自己也很扎心(“; ∇; )我现在正在用自己一切力量去和别人宣传垃圾			
428 a8, 保护一号机(“; ∇; )			
我们最能做的只有取关和向更多的人揭露a8的丑恶嘴脸, 真是不甘心啊, 期待那个本应该	取关和向更多人揭露a8的丑恶嘴脸		
3-ASMR-年上女友的工作想问 2-你能猜出来吗 1-video 01 analytical codes			

(open coding)

A	B	C	D	E	F	G	H
category	sub category	sub-sub category	video	descriptive code/comments	preliminary category		
1476 positive engagement	voice model	affection	1-1	1 老婆不回来, 对剩下顶着爱管名字的人没遗憾和感情; 将最后的第100w关注送给老婆正式回归的礼物, 这是老婆三年来努力的回	affection to aichan/original ai		
1477 positive engagement	voice model	affection	1-1	1 挺, Vtuber界的劳模和属于老婆的荣耀;	affection to aichan/original ai		
1478 positive engagement	voice model	affection	1-1	1 嗨嗨嗨 爱管老婆爱了	affection to aichan/original ai		
1479 positive engagement	voice model	affection	1-1	1 好想老婆, 希望她回来能给我看看	affection to aichan/original ai		
1480 positive engagement	voice model	affection	1-1	1 影是为了爱管, 这有老婆的vb没有存在的必要	affection to aichan/original ai		
1481 positive engagement	voice model	affection	1-1	2 右边的老婆爱管已经听不见了, 我发了我婆婆删老婆视频	affection and imagination of original ai		
1482 positive engagement	voice model	affection	1-1	2 右边的老婆爱管, 不过这是上个月23号的存货	affection and imagination of original ai		
1483 positive engagement	voice model	affection	1-1	2 一听说听出来了, 右边是老婆	affection and imagination of original ai		
1484 positive engagement	voice model	affection	1-1	2 虽然关注时间不长, 但还是喜欢老婆	affection of Kizuna AI		
1485 positive engagement	voice model	affection	1-1	2 爱管打打打, 老婆爱管的视频真的挺快乐的	affection and imagination of original ai		
1486 positive engagement	voice model	affection	1-1	2 待老婆回来之时便是我400硬币用光之时	affection and imagination of original ai		
				我们喜欢老婆又不是因为地玩游戏多厉害! 即使她玩最后一台我们也会喜欢老婆			
1487 positive engagement	voice model	affection	1-1	3 因为老婆缺一个, 她的内在是其他机型不能比的! 替代不了的!	affection of original ai, irreplaceable original ai		
1488 positive engagement	voice model	affection	1-1	3 唉, 二号替我老婆问声好	affection of Kizuna AI		
1489 positive engagement	voice model	affection	1-1	5 只喜欢老婆和少儿子部道	affection and imagination		
1490 positive engagement	voice model	affection	1-1	回复 @半半CASH-ERN 现在总是想知道老婆现状, 是不是真的想管管说的			
				5 只是老婆忙, 还是网友传的要被官方替换了, 是真的舍不得老婆啊(“; ∇; )	affection and imagination of original ai		
1491 positive engagement	voice model	affection	1-1	回复 @Rayoth 爱管“分崩”开始的时候, 许多人以为老婆会消失, 然后老婆	affection and imagination of original ai		
1492 positive engagement	voice model	affection	1-1	说会消失她的主意, 她就不会回来的, 听到这我还在感动的... 唉	confusion around voice replacement		
1493 positive engagement	voice model	affection	1-1	5 后事情莫名其妙变成这样了?????????搞不懂, 真心搞不懂... 唉~	affection and imagination of original ai		
1494 positive engagement	voice model	affection	1-1	5 回复 @bbbdidi 这就是老婆, 印在dna的声音, 不会错的	affection and imagination of original ai		
				11 老婆爱管!	affection and imagination of original ai		
				12 我喜欢老婆以前高中时候就喜欢周末看她视频 望的推特搬运我也关注了	affection and imagination of original ai		
				我也喜欢老婆			
1495 positive engagement	voice model	affection	1-1	12 我和你贴贴	affection and imagination of original ai		
				老婆和大熊猫贴贴			
				我喜欢大熊猫 = 我和大熊猫贴贴			
1496 positive engagement	voice model	affection	1-1	12 老婆和你贴贴	affection and imagination of original ai		
1497 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 做表情那期视频油管删教比教数多了	data of video		
				坐等老婆百万粉丝开播的时候, 我只希望新人就能得到这种待遇的人都不			
				是单个人啊! 这一期视频还在哪里直播地骂老婆 - 就是两个a8的傀儡啊~			
1498 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 那有什么可怜楚楚之说	dislike no.2/3		
				舔天油嘴副鸟恶人的视频几乎一片倒的恶评, 基本都有带7号的单词高			
1499 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 风尖嘴, 二三号机来的视频都在油管上踩了一遍都不解气, 明明是老婆	dislike and anger toward no.2/3/4		
				2 好不容易建立起来的羁绊却被两个寄生虫践踏, 真你的啊A8			
1500 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 这种坏自己前程的事是个屁嘛敢主动要求的? 就代替老婆玩粉图的3号, 音	thoughts on relationship between a8 and voice performers		
1501 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 方一点啥个说的! 这道理不是出来被要求的?			
1502 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 这也太恶心的了, 我都想骂人了			
1503 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 占老婆的频道, 用着老婆的脸, 模仿着老婆的声音, 暗讽着老婆的人, 做	dislike and anger toward no.2/3/4		
1504 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 qd跟老婆爱的人说, 今天还要把老婆的歌拜在地上踩上两脚!	dislike no.2/3		
1505 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 在出, 那得出来说嘛, 去吃点信信的			
1506 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 由一个外人用爱管模型模仿高甜内心实在接受不了			
1507 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 我亲眼这俩身在这请罪嘛	dislike and anger toward no.2/3/4		
				我仇恨开始的, 歌曲一首歌河都是爱管写的, 这些不要脸的居然献唱,			
				2 抵制到底			
				我更喜欢老婆能站出来, 这样老婆的粉丝也会支持他们, 就算买了饭砸			
				已经对时好, 抗争成功的活老婆大概会带他们玩, 比现在这样好, 现在的			
1508 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 情况她们已经是帮凶了			
1509 negative engagement	voice model	dissatisfaction and anger with no.2/3/4	1-1	2 同流合污的一个也不能放过			

(final collection)

### 7.2.2 Qualitative text analysis results

theme	category	Sub category	Sub-sub category	Sub-sub-sub category
engagement	Emotional engagement	Positive engagement	Kizuna AI	Affection and imagination
				Expectation (of fan goods)
				Meaning making
				play with ai's characteristics (and video contents)
				retrospect
				Relate to additional information
			Voice model	Acceptance (of no.2/3/4)
				Affection
				Recognition
				Support
		Other fans/audiences	Appreciation	
			Encouragement	
			Understanding	
		human-kizuna ai interaction		
		Negative engagement	Kizuna AI	Give up Kizuna AI
				Sadness
				worry
			company	dissatisfaction and anger with A8
			Other fans/audiences	Collective activity
				Disappointment
	Conflicts between different kinds of fans			
	Disapproval			
	Voice model	dissatisfaction and anger with no.2/3/4		
	Disengagement	Unsubscribe aichannel		
		Disengage with no.2/3/4		
	Critical engagement	Capital & labour	Company-labour relation	
			Capitalism	
			Commercialization of Kizuna AI	
		Fans and audiences	Transnational fans	
			Role of fans/audiences (in voice model incident)	
Importance of fans (to producing company)				
Reflection on VTuber industry				
What/who is Kizuna AI				
Context to engagement	Cultural background	Anime and Otaku culture		
		Bilibili and “danmaku” culture		
		Internet culture		
	Place	Fan creation distribution and sharing		
		Distribution on bilibili.com		
		Support subtitle group		

			YouTube as official distribution platform
			Social media in China