

Snapshot

Is Regulatory Framework the Only Solution for GenAI in Elections?

Balancing Regulation and Freedom of Expression in the Republic of Korea

Minjun Hong

SAIL Snapshots is a peer-reviewed article series featuring key research findings from Safer Internet Lab researchers and its associates. The views presented are exclusively those of the authors and do not reflect the official stance of SAIL, CSIS, Google, or any other affiliated organisation.

Is Regulatory Framework the Only Solution for GenAI in Elections?

Balancing Regulation and Freedom of Expression in the Republic of Korea

Minjun Hong¹

Editor: Noory Okthariza²

INTRODUCTION

With the acceleration of digital transformation, serious concerns have arisen regarding disinformation related to elections (Go et al., 2021). In particular, generative AI (GenAI) has made it easy to produce convincingly realistic disinformation, further threatening the legitimacy of democracy (Oh, 2024). During the 2022 presidential election in South Korea, avatars of presidential candidates created by Deepfake were employed in online campaigns. Although there was considerable controversy during the elections, a larger controversy arose in the 2022 local elections (Eom et al., 2022). A fake video of President Yoon voicing his endorsement of a Namhae County Mayor candidate circulated on social media. This caused viewers to incorrectly perceive that the president had not maintained political neutrality (Lee H, 2024).

The incident sparked controversy and demonstrated that deepfake could be employed to affect elections. Subsequently, a discussion arose concerning the need for regulations related to deepfakes for

¹ Seoul National University

² Researcher, Department of Politics and Social Change, CSIS Indonesia

elections, leading the National Assembly to establish a regulatory framework in December 2023 (Chang, 2024). Despite legal efforts, deepfake videos were created and circulated to criticize or mock candidates in the 2024 National Assembly election (Lee S, 2024). Furthermore, even deepfake videos intended to share accurate information are prohibited starting 90 days before the election, drawing criticism of restricting freedom of expression. However, given the rapid advancement in GenAI, imposing indiscriminate restrictions is neither appropriate nor effective. Instead, it is essential to establish an appropriate regulatory framework that considers how to responsibly employ technology. Accordingly, this study explores the regulation of the GenAI for elections and stakeholders' efforts in South Korea and presents implications.

REGULATORY PREPAREDNESS, GAPS AND CURRENT INTERVENTIONS AND LIMITATIONS

In South Korea, policies related to elections, AI, and disinformation have been implemented, with a focus on deepfakes. In January 2022, the National Election Commission issued an operational standard for the regulation of deepfake videos, stating that, “while election campaigns can utilize deepfake videos in legally permissible ways, failure to indicate that the content is a deepfake video would be a violation of Article 250 (Publication of False Information) of the Public Official Election Act” (National Election Commission of the Republic of Korea, 2022). Regarding this standard, Huh (2022) argued that it could have a chilling effect on freedom of speech, as it constitutes a form of prior censorship.

Despite operational standards, deepfakes continued to be produced with the intent of spreading disinformation for political and electoral purposes. Thus, in December 2023, the Public Official Election Act was amended to include Article 82-8 (Regulation of election campaigns using artificial intelligence-based Deepfake

videos, etc.). It prohibits the production, editing, distribution, screening, or posting of deepfake videos for election campaign purposes from 90 days before election day. It also mandates the labeling of AI-generated materials (Korean Law Information Center; National Election Commission of the Republic of Korea, 2024). Those who violate the Act may be subject to prison sentences of up to seven years or fines of 50 million won (US\$ 36,000). (Freedom House, 2024). Following the amendment to the Act, the National Election Commission launched an AI identification and monitoring team to address disinformation (National Election Commission of the Republic of Korea, 2024). Although the effectiveness of the amendment is unclear, the 2024 general election in South Korea concluded with relatively little disruption (Oh, 2024).

However, the amendment of the Act has been criticized for excessively restricting freedom of expression. That is, even AI-generated deepfake videos intended to convey accurate information are banned starting 90 days before the election (Chang, 2024). Furthermore, there is a lack of exemptions for immunity, despite the high level of penalties. However, in reality, the National Election Commission detected 129 deepfake online campaign videos between late January and the end of February 2024, and it did not refer them to the police. Instead, they requested their removal, treating them as personal expressions. This highlights the challenge of enforcing the law. Thus, some criticized the legislation's over-regulation and antagonistic stance toward new technology (Park, 2024).

In addition to the efforts to establish a regulatory framework through legislation and compliance, various initiatives have been undertaken. In May 2024, the Ministry of Science and ICT announced an Implementation Plan for Establishing a New Digital Order. One of the strategies is to establish a comprehensive

response system to combat deepfake disinformation news, encouraging the voluntary participation of platform companies. Furthermore, the plan includes strengthening the management of deepfake and enhancing capabilities for detecting deepfake disinformation (Ministry of Science and ICT, 2024a). A notable aspect of this plan is the public-private collaborative response to address deepfake disinformation. The emphasis on supporting the private sector and fostering collaboration stems from the challenge of managing this issue through government efforts alone, which requires active platform company involvement. Additionally, direct government regulation has been highly contentious, as seen in criticism of past attempts to regulate fake news. Some have argued that government intervention in the press could act as de facto censorship, raising constitutional concerns (Ahn, 2023).

Platform companies in the private sector have attempted to address the issue of AI-driven disinformation. For example, Naver Corporation³, an online platform and search engine, issued guidelines on deepfake content to comply with the amended Public Official Election Act. A warning label clarifying the limitations of GenAI is displayed when GenAI tools are used for election-related information. Additionally, AI filtering technology is employed to block harmful deepfake content in real-time (GOH, 2024). Kakao Corporation, which operates a popular mobile messaging app⁴ and portal service, monitors malicious deepfake content in its services. Users who distribute false videos on KakaoTalk may face a permanent service ban. In addition, Kakao announced the

³ In 2023, Naver held the largest share of monthly active users among search engines in South Korea, with approximately 58 per cent (Jobst, 2024a). KakaoTalk, operated by Kakao Corporation, remains the most popular mobile messaging app in South Korea (Jobst, 2024b).

⁴ In 2023, Naver held the largest share of monthly active users among search engines in South Korea, with approximately 58 per cent (Jobst, 2024a).

application of invisible watermark technology in its AI image-generation model to prevent the misuse of deepfakes in elections (OH & Hwang, 2024).

Collaboration among stakeholders can be categorized into government- and private sector-led initiatives. Regarding government-led initiatives, the Korea Communications Commission has led the ‘Self-Regulation Promotion Council’ since 2023. This council includes not only government organizations, but also domestic and international portals and platform companies. Another example is the ‘Public-Private Council for User Protection of AI Service’, established in 2024 to analyze AI technology and policy trends and discuss policy measures to protect AI service users. Experts from various fields, including academia, industry, legal circles, and civic groups, participate in these efforts (Korea Communications Commission, 2024). In addition to government-led initiatives, major companies (e.g. Naver, Kakao, SK Communications) jointly signed a declaration in March 2024 to prevent malicious election-related deepfakes, committing to several key actions (e.g., swift response, collaboration with external experts, etc.). Furthermore, the Korea Internet Self-Governance Organization (KISO), whose members include Korean tech companies, has been implementing various strategies to combat fake news. Since December 2023, it has enforced a “Policy on Election-Related Internet Information Service Standards” in preparation for the general election. However, these policies do not apply to international tech companies such as Google, which are not KISO members (Lee & Lee, 2024).

Public perceptions of government policies are challenging to pinpoint, but a survey by the Ministry of Science and ICT⁵ provides

⁵ The Ministry of Science and ICT conducted a survey through their website in June and July 2024, with 765 people responding.

a good overview of the public perception of AI. Among the respondents, 57.3% (438 respondents) stated that the potential benefits of AI outweigh its risks, whereas 181 respondents viewed the risks and benefits as balanced. Regarding regulation versus innovation, 55.2% prioritized innovation, while 122 respondents considered both to be equally important. Regarding potential AI risks, 7.7% mentioned the spread of fake news through generative AI. For key government policies, 34.6% highlighted establishing AI ethics standards and legislation (Ministry of Science and ICT, 2024b). These findings suggest that public perception recognizes the benefits of AI technology and emphasizes fostering innovation rather than imposing indiscriminate regulations.

CONCLUSION

South Korea may lack a comprehensive regulatory framework through legislation if only the number of laws are considered. However, the content of these laws reveals that their responses are extensive and stringent. Notably, the government enforced a complete ban on deepfake videos starting 90 days before an election. This measure may be intended to prevent the possibility that disinformation created through deepfake technology spreads shortly before the election, remains unverified, and consequently has a severe impact on the electoral process. However, it has faced criticism for blocking the possibility of appropriate usage in elections.

This process can be considered as part of an effort to establish an appropriate level of regulation for rapid digital transformation. Given the government's limited adaptation and predictability of emerging technologies, it has understandably adopted a conservative approach, as elections play a pivotal role in democracy. Recently, however, the government has begun to lead public-private partnerships, creating collaborative coalitions to

explore appropriate regulatory levels. The private sector and civil society are also contributing to this endeavor. For instance, in April 2024, the AI Future Forum was launched with scientific and technological civic groups and over 700 AI experts to provide policy recommendations for effective and safe AI use, recognizing current GenAI usage as insufficient (Im, 2024).

Through these efforts, South Korea will identify an optimal level of regulation that safeguards freedom of expression through the appropriate use of AI and prevents potential harm to democracy posed by AI-generated disinformation. This approach is expected to allow GenAI to play a constructive role in future elections by effectively delivering candidates' policies.

REFERENCES

- Ahn, Y. (2023, September 20). Korean media regulator's "one strike" policy for "fake news" denounced as censorship. *Hankyoreh*. Retrieved from https://english.hani.co.kr/arti/english_edition/e_national/1109359.html
- Chang, H. (2024). AI deepfake jegakmul-i seonge-e kkinun yeonghyang [The impact of AI deepfake productions on elections]. *KIPA Regulatory Trends*, 46.
- Eom, K., Kang, I., & Kim, D. (2022). Ingonjiineunggisulbaljeon-e ttareun seongeundong jedo gaeseon bang-an yeongu: deepfake, metabus, bigdata hwal-yong seongeundong-eul jungsim-euro [A study on improvements to election campaign systems according to the advancement of AI technology: Focusing on deepfake, metaverse, and big data utilization]. *2022 National Election Commission Research Report*.

Freedom House. (2024). Key Developments, June 1, 2023 – May 31, 2024. *South Korea: Freedom on the Net 2024 Country Report*. Retrieved from <https://freedomhouse.org/country/south-korea/freedom-net/2024>

Go, S. G., Lee, S., Kim, H., & Song, J. (2021). Online gajanyuseu gwanyeon seongeongwanrijeido gaeseonbang-an-e daehan yeongu [A study on improving election management systems related to online fake news]. *Korea Election Association Research Report*.

GOH, M. (2024, February 21). [Exclusive] ‘This is Harmful Deepfake Content’... Naver to Apply Warning Labels. *MAIL BUSINESS NEWSPAPER*. Retrieved from <https://www.mk.co.kr/en/it/10947993>

Huh, S. (2022). Deepfake Videos on YouTube and Publication of False Statement of Facts. *Journal of Media and Defamation Law*. 8(1). 1-47.

Im, J. (2024, April 3). “Saengseonghyeong AI hwal-yong segye 20wi bak” ... jeonmunga 700yeo-myeong simindan-tewa gyeolhab [“Generative AI usage outside the top 20 in the world” ... over 700 experts join with civic groups]. Hankyoreh. Retrieved from https://www.hani.co.kr/arti/economy/economy_general/1134987.html

Jobst, N. (2024a). Leading search engines in South Korea in 2023, by share of monthly active users. *Statista*. Retrieved from <https://www.statista.com/statistics/1219834/south-korea-leading-search-engines-by-mau/>

Jobst, N. (2024b). Kakaotalk number of monthly active users worldwide 2013-2024, by quarter. *Statista*. Retrieved from <https://www.statista.com/statistics/278846/kakaotalk-monthly-active-users-mau/>

Korea Communications Commision. (2024, July 19). Bangtongwi, Ingonjiineun service iyongja boho mingwan hyeobuihoe chulbeom [KCC launches the ‘Public-Private Council for User Protection of AI Services’]. Press Releases. Retrieved from <https://kcc.go.kr/user.do?boardId=1113&page=A05030000&dc=K00000200&boardSeq=62556&mode=view>

Korean Law Information Center.
<https://www.law.go.kr/LSW/eng/engMain.do?eventGubun=060124>

Lee, H. (2024, February 20). April 10 elections under threat from AI deepfake manipulation. *The KoreaTimes*. Retrieved from https://www.koreatimes.co.kr/www/nation/2024/10/113_369059.html

Lee, K. T., & Lee, J. (2024, February 27). Google's absence in Korea's internet safety initiative raises concerns over fake news. *THE CHOSUN Daily*. Retrieved from <https://www.chosun.com/english/companies-en/2024/01/23/TARTQ5FQWBCHJKJRMXCT4CJ4I>

Lee, S. (2024, May 13). AI and Elections: Lessons From South Korea. *DIPLOMAT MEDIA*. Retrieved from <https://thedi diplomat.com/2024/05/ai-and-elections-lessons-from-south-korea/>

Ministry of Science and ICT. (2024a, May 20). Daehanminguk-i saeroun dijiteol jilsuh jeongnip-ui maseuteopeulraen-eul

gongkaehabnida [South Korea unveils a master plan for establishing a new digital order]. Press Release. Retrieved from

<https://www.msit.go.kr/bbs/view.do?sCode=user&mId=113&mPid=238&pageIndex=&bbsSeqNo=94&nttSeqNo=3184512&searchOpt=ALL&searchTxt>

Ministry of Science and ICT. (2024b, August 7). Gukmin 10myeong jung 6myeong, “Ingongjineung gisul ijeomi wihyeobboda keuda” [6 out of 10 citizens believe the benefits of AI technology outweigh its threats]. Press Release. Retrieved from

<https://www2.korea.kr/briefing/pressReleaseView.do?newsId=156644751&pWise=sub&pWiseSub=C4>

National Election Commission of the Republic of Korea. (2022, January 11). Deepfake yeongsang gwanryeon beobgyu unyong gijun [Operational Standards for the Regulation of Deepfake Videos]. Press Releases. Retrieved from <https://www.nec.go.kr/site/nec/ex/bbs/View.do?cbldx=1084&bcldx=156972>

National Election Commission of the Republic of Korea. (2024, February 6). 90 Days Until National Assembly Elections: Election Campaigns Using Deepfake Restricted. Press Releases. Retrieved from <https://www.nec.go.kr/site/eng/ex/bbs/View.do?cbldx=1270&bcldx=226657>

OH, C & HWANG, S. (2024, March 13). LG Joins Global Efforts to ‘Prevent Deepfake Content’. *MAIL BUSINESS NEWSPAPER*. Retrieved from <https://www.mk.co.kr/en/it/10963541>

Oh, I. S. (2024). How to Response the Distortion of democratic legitimacy by Deep Fake. *INSS Strategy Report*. 260.

Park, S. (2024, February 27). 'Deepfake' seongeundong, choedae jingyeok 7nyeon... gwaing ibbeop jikjeokdo ["Deepfake" election campaign, up to 7 years in prison... criticism of excessive legislation]. *Journalists Association of Korea*. Retrieved from http://m.journalist.or.kr/m/m_article.html?no=55321



Safer Internet Lab

 saferinternetlab.org

 Jl. Tanah Abang III no 23-27
Gambir, Jakarta Pusat, 10160

Find Us On



CSIS Indonesia | Safer Internet Lab