

KGRI Working Papers

No.2

Joint Statement: "Toward a Wholesome Platform for Speech: Digital Diet Declaration" (ver. 1.0)

Provisionary English translation

Version 1.0

January 2022

Fujio Toriumi

Professor, School of Engineering, The University of Tokyo

Tatsuhiko Yamamoto

Professor, Law School, Keio University

Deputy Director, Keio University Global Research Institute

Keio University Global Research Institute

© Copyright 2022

Fujio Toriumi, Professor, The Graduate School of Engineering, The University of Tokyo and Tatsuhiko Yamamoto, Professor, Law School & Deputy Director of Global Research Institute, Keio University

Joint Statement

Joint Statement: "Toward a Wholesome Platform for Speech: Digital Diet Declaration" (version 1.0)

Provisionary English translation

Fujio Toriumi

(Professor, School of Engineering, The University of Tokyo)

Tatsuhiko Yamamoto

(Professor, Law School, Keio University; and, Deputy Director, Keio University Global Research Institute)

Summary

We believe that fake news (disinformation) and the "infodemic" (too much information during a disease outbreak) represent grave social illnesses that can even threaten people's lives. Underlying this is the fact that the "diets" of information that the internet's development has produced are typified by overconsumption and unbalanced selectivity. Accordingly, we have arrived at a shared awareness that "information health"—that is to say, an awareness as both individuals and as a society of what is an appropriate balance of information—is extremely important. This joint declaration compiles the results of having investigated from an interdisciplinary perspective what sorts of initiatives are ought to be pursued for the various stakeholders involved, including the operators of digital platforms, to achieve this information health.

This statement is absolutely a preliminary version (version 1.0). We hope to solicit opinions from researchers and stakeholders in various fields in the future and revise the content accordingly.

Table of Contents

Preamble	2
I. Current Issues	4
II. The Orientation of the Joint Statement	8
III. The Ground Rules for Users	12
IV. The Ground Rules for Operators	14
V. The Ground Rules for the Government	21
VI. Prospects for the Future	23
* List of Co-Signers	

^{*} Research Assistants

Preamble

Because of the COVID-19 pandemic, it has become clear that the infodemic is not only generating confusion in the information space, but also is an acute issue that hinders infection control measures and threatens human lives. Meanwhile, the digital platforms (DPFs), which are the primary venue for this, have become infrastructure-like that supports our daily lives, providing a place to handle everything from everyday shopping to communicating with friends to reading the news. Their usefulness has improved dramatically. These two dimensions are intricately connected. The information that we receive has increased explosively, and it is being delivered to us selectively by specific algorithms that have been constructed to attract our attention. These, we believe, have caused numerous problems. The balance in the information that we take in has been disrupted, the power of fake news to influence us has increased, and filter bubbles and echo chambers are being created.

The option of what information to take in is something that should be left up to the individual's choice in the first place, and their freedom to be particular about what information they consume should also be respected. Nevertheless, as we have seen in the COVID-19 crisis, the spread of fake news and the infodemic have hindered efforts to control the infection's spread. When it is already causing chaos in society, it becomes a problem for society as a whole. This applies not only to the COVID-19 crisis but also to democracy itself, as seen in the chaos in the 2020 US presidential election.

What the issues in particular should be are, under the "attention economy" business model, the facts that (a) most of us are "put into" the position of consuming information in an unbalanced way, (b) we are "forced" to take in specific types of content in an other-directed fashion because of algorithms that maximize the economic interests of business operators, and (c) most people are "made to not fully aware" of the fact that they have been placed in such an information environment.³ And so, the most crucial issues are (a) having a clear awareness of both what the information environment is, in which we are placed, and what sorts of information we are consuming every day⁴ and (b) providing those people who seek to improve the unbalanced

For reference, see Ministry of Internal Affairs and Communications, *Information and Communications in Japan 2020*, part 1, chapter 2, section 3, (1) The dissemination of misinformation and fake news.https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/rWP2020/2020-index.html>

According to a survey of 8,001 persons in the United Kingdom and the United States conducted in September 2020 by a team of researchers affiliated to Imperial College London and the London School of Hygiene and Tropical Medicine, the percentages of those who planned to get vaccinated after having viewed such erroneous information regarding COVID-19 vaccines as "vaccinations will change your DNA" dropped by 6.2% in the United Kingdom and 6.4% in the United States. See, L., Sahil, et al., "Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA," Nature Human Behaviour volume 5, pages337 – 348, 2021. < https://www.nature.com/articles/s41562-021-01056-1 >

Twitter announced in October 2021 that based on a seven-country survey including the United States, some European countries, and Japan, its own algorithms were four times more likely to present posts from right-wing politicians and media than from those on the left wing. The survey results were presented in the Proceedings of the National Academy of Sciences of the United States of America. See, Huszár, F., et al., "Algorithmic amplification of politics on Twitter," PNAS January 4, 2022 119 (1) e2025334119. https://www.pnas.org/content/119/1/e2025334119>

⁴ Public relations company Edelman Japan in its 2021 Edelman Trust Barometer global survey measured "information hygiene" on the basis of four dimensions: do the people surveyed engage regularly with the news, do they avoid information echo chambers and engage with differing points of view, do they verify information, and do they check information veracity before forwarding the content to others? The percentages of those whose levels of information hygiene was graded as "good" across three or more of these dimensions stood at 26% on average across the 27 countries surveyed; by contrast, for Japan on its own, the figure was 19%. Conversely, although the 27-country average stood for "poor" information hygiene—meaning performs well on one or zero dimensions—stood at 39%, for Japan, the figure was 56%.

diet of information that they are taking in with the ability to take in various information in a more balanced way.

This will create a healthy environment for speech in which we will enjoy the advantages that information communications technology produces, avoid the dangers that fake news brings, and harmonize with the fundamental principles of the Constitution.

Toward that end, what is asked of us now is to create a situation in which we have acquired a certain "immunity" (critical faculty) concerning fake news by achieving a balance in taking in various information—in short, by achieving information health. In turn, it will be necessary for changes to occur among users, DPF operators, and the existing mass media and also for the government to provide various lateral assistance directed toward achieving information health.

These issues are also national ones that threaten the fundamental principles of the Constitution such as respect for the individual and democracy. Still, it would be impetuous for the government to impose strict legal restraints across the board on DPF operators. A "hard law"-based approach would not only possibly inhibit the innovations that are the wellspring for development in digital technology but also contain the danger of leading to censorship and efforts to influence thought by the state. At the current stage, this cannot be said to be the ideal way to consider the issue.

This joint declaration presents an opinion on what each of the actors involved—users, operators, and the government—should undertake so that everyone who wishes for information health in spaces for public discourse can enjoy it.

3

https://www.slideshare.net/EdelmanJapan/2021-246556173/24

I. Current Issues

(1) The Big Bang in Spaces for Public Discourse

Before society entered the digital era, the spaces for public discourse could be more properly termed privileged circles that were relatively closed. The air waves and page spaces had physical restrictions, and the transmission of information was unilateral, slow, and in small volumes. Additionally, the people who delivered the information were basically restricted to the so-called mass media, the professional journalists affiliated with it, and those intellectuals upon whom the mass media focused.

With the spread of the internet, the space for public discourse has grown explosively akin to a "big bang" and plunged into an entirely new, heretofore unseen dimension in which information mingles together, moving in both directions, in real time, and without restriction. At the dawn of the internet, there had been expectations that this big bang would result in an open, democratic world. Nonetheless, presently, what we have is a world that is awash in an uneven jumble of user-generated content (UGC) being posted under both real and assumed names by a diverse range of average users. It is a world in which false information of unknown origin exists side by side with fact-based journalism; such disinformation is disseminated and amplified, aided by a nonexistent fake crowd, such as bots.

Several DPF operators mix professionally generated content (PGC) written and delivered by the mass media and professionals with content written and sent by people who have received no specialized training or do not share at an organizational level ethical codes of conduct. These operators present this information in an order ranked by special algorithms. Consequently, the existing mass media—newspapers and magazines—have also been placed into a situation in which they are compelled to compete fiercely under the logic of the attention economy (see I. (2)), where the momentary impact on readers and users governs viewing numbers.

Recommendation algorithms that use artificial intelligence (AI) and other technologies to provide personalized information have the advantage of making it possible for the users themselves to be able to definitely get articles and content that is of potential interest to them. Conversely, this means that they will consume only content recommended to them by the platform's recommendation algorithm. The result, as has been pointed out, is the creation of a "filter bubble" (or "information cocoon") from within which the individual can only see the information (that the algorithms have judged) they want to. This, for the users, means the development of a "selective information diet."

(2) Attention Economy

Our existences have become ones in which we are ceaselessly bathed in information from the internet and social media even as we produce information ourselves. Accordingly, information is being endlessly and repeatedly consumed and shared; there is no being starved for information.

In a society with such an information surplus, since the attention and amount of time for consumption that we can direct toward it will grow scarce in comparison with the amount of information being supplied, that time and attention have value on the marketplace (attention market) and circulate there. This economic model is generally referred to as the attention economy⁵. With smartphones and other mobile devices having deeply penetrated into our daily lives, our body is being controlled more and more by the attention economy.

Psychology speaks of the existence of two modes in human thinking, (1) the intuitive and automatic "System 1" and (2) "System 2," which supplements System 1 and is characterized by careful deliberation (dual process theory)⁶. In the world of the attention economy, the former—that is to say, stimulating System 1—is said to be the important one. That is because the spinal, reflexive response from the users that is obtained by providing stimulus ⁷has economic value. (This "reflex" is assessed on the basis of such factors as page views, the amount of time spent at a website, and stickiness and then converted into an economic value.)

This new economic model is causing a decline in System 2 thinking, which when left to take its own course provides the basis for discussion and mature deliberation. It carries the danger of exposing the democracy that is premised by System 2 thinking to jeopardy.

(3) Mind Hacking

Profiling generated by using artificial intelligence and the like has made it possible to analyze and predict with an extremely high level of accuracy a user's sentiments and psychographic profile, to say nothing of their political convictions. The work undertaken by British political consultancy Cambridge Analytica in the 2016 US presidential election among other contests is regarded as psychological profiling undertaken through the use of psychographics. Through political microtargeting applying such technologies, it is now getting to the point where an individual's sentiments and thinking can be easily "hacked" and manipulated (mind hacking)⁸. This enables others to psychologically intervene and manipulate individuals in situations that they may not or cannot be aware of. It undermines an individual's self-determination, autonomy, and above all, the democracy that is premised by such abilities.

Moreover, for the sake of their own economic interests, DPF operators have been deploying in various settings the so-called "dark patterns" in UI/UX interactions that improperly incite user decision making.⁹

(4) Filter Bubbles and Echo Chambers

Through profiles and targeting using big data, information deemed as being "profitable" to us is being

⁵ See generally, Wu, Tim. "The attention merchants: The epic scramble to get inside our heads." Vintage, 2017.

For reference, see Daniel Kahneman, Thinking, Fast and Slow (New York: Farrar, Strauss, and Giroux, 2011); trans. Akiko Murai, Fasuto and suro: Anata no ishi wa dono yo ni kimaru ka (Tokyo: Hayakawa Nonfiction bunko, 2014).

⁷ For reference, see Ryosuke Nishida, Joho buso suru seiji (Politics armed with information) (Kadokawa, 2018).

For reference, see Christopher Wylie, Mindf*ck: Cambridge Analytica and the Plot to Break America (Random House/Penguin Random House LLC, 2019); trans. Yo Makino, Maindo hakkingu: Anata no kanjo o shihai shi kodo o ayatsuru sosharu media (Shinchosha, 2020).

⁹ See, KGRI, "Daku patan" ni tsuite 'kangeru: Tokubetsu koen "'Daku patan' no gijutsu to ronriteki kadai" (held 17 June 2021).

filtered out from the sea of information that surrounds us, making it possible for us to view information that has been customized for "personal use."

Nevertheless, this is not all good. As has been observed, there is a tendency among people for their thinking to become more extreme when they gather together and talk only with other people who share their views (group polarization). ¹⁰ "Echo chambers" and "filter bubbles" are the phenomena that are said to be produced by the interplay between this human tendency and the specific character of the internet.

Algorithms analyze user preferences and privilege presenting information that is based on those preferences. The result of the information being displayed that gives precedence only to information beneficial to the user is that they wind up only seeing information *regarded as* what they want to see, as though they have been locked into a bubble of their own ("filter bubble"). Gathered together inside this bubble is a vast collection of thoughts and opinions that resemble one's own, whereas opposing ones are removed (filtered). Because of this, even noticing the very existence of such opposing views is difficult.

The same sort of phenomenon occurs on social media. The result of someone following only other users who have the same interests as themselves is that when the individual expresses a certain opinion, all they hear back are echoes of the same view ("echo chamber"). Because one hears only the same sorts of opinions over and over again due to that echo, the individual ends up believing in them that much more fervently as being correct and unerring (think here of conspiracy theories).

Observers point out that group polarization will accelerate because of filter bubbles and echo chambers.¹¹ People who have made their inclinations more extreme are not accepting of people with different views and reject conversation with them. These two phenomena can induce divisions in society and jeopardize democracy.

(5) Fake News

Considerable amounts of fake news (disinformation) created with the goals of generating advertising revenue or causing people to lose trust in prominent individuals, political groups, corporations, and so forth are circulating in the digital space today. It is being disseminated and amplified, aided by bots and the like. The creation and dissemination of fake news can also be said to be a "natural" phenomenon intricately connected with the real nature of human beings as seen in the so-called "social porn hypothesis" (information that users affiliated with a specific community want to disseminate and share reflexively among themselves) and slacktivism (social movements that are easy to participate in). Moreover, the attention economy has a

Refer to the works of Cass R. Sunstein. For example, #Republic: Divided Democracy in the Age of Social Media (Princeton: Princeton University Press, 2017; trans. DATE Naomi, #REPUBLIC (Tokyo: Keiso Shobo Publishing, 2018).

However, careful empirical research will be crucial. For reference, see Tatsuo Tanaka and Satoshi Hamaya, Netto wa shakai o bundan shinai (The internet does not divide society) (Tokyo: Kadokawa, 2019). For a recent important empirical study, see Daisuke Tsuji, ed., Netto shakai to minshu shugi (The network society and democracy) (Tokyo: Yuhikaku, 2021).

structure that encourages the dissemination of fake news because of its emphasis on stimulation and reflection. For example, most algorithms for digital advertisements calculate their costs on the basis of how many users they have attracted (impressions). This has produced the phenomenon of one after another fake news item being created and disseminated to generate advertising revenue.

Also, as has been noted, fake news is being created and disseminated as part of efforts by other countries to conduct "influencing maneuvers." These maneuvers can jeopardize national security and national sovereignty itself.

(6) Slander and Flaming

Slander has also become a big problem in digital spaces for public discourse today. In recent years, extreme statements and abusive language have been circulating in large numbers on social media and message boards on the internet. It is becoming a social problem. This, too, is connected to the structure of the attention economy, with its emphasis on reflexive stimulus. Suicides have also resulted due to flaming, and words on social media are turning into armaments and lethal weapons that bring people to death. Today, the situation is such that slander is something that is becoming a familiar presence not just to famous people but also for the average citizen; anyone can become its victim or its perpetrator.

(7) The Absence of a "Specific Remedy"

As we have seen thus far, various problems have burst forth in the spaces for public discourse today. That said, the mechanisms by which they are generated are both complex and structural. Accordingly, no single remedy exists that could alter this situation completely.

Also, by continuously being exposed to the chaos of the information environment caused by fake news and the like, there is a danger that we ourselves will find our constitutions weakened by such news. The problems go far beyond becoming involved with disseminating fake news and getting caught up in absurd conspiracy theories. The divisions in society could grow further, our tolerance toward others may be lost, and coexisting with others and maintaining our communities may become more difficult. This is a crisis for democracy and liberty.

In the absence of a remedy, what we believe we can do at present is to seek to obtain information health—that is to say, acquire some immunity toward fake news and the like—by envisioning and being conscious about taking in a balanced information diet.

According to Bradshaw et al. (2021), 81 countries worldwide are engaged in computational propaganda in such forms as fake news, an increase of 11 countries from the year before. *See generally*, Bradshaw, S., Bailey, H.& Howard, P.N., "Industrialized Disinformation 2020 Global Inventory of Organized Social Media Manipulation," University of Oxford, 2021. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/01/CyberTroop-Report-2020-v.2.pdf

II. The Orientation of the Joint Statement

(1) Achieving Information Health

Today, when the flow of information in the space for public discourse has undergone a great transformation, it is indispensable that we achieve a state of affairs in which we have acquired a certain degree of immunity (critical faculties) to fake news—that is to say, information health¹³—by coming into contact with a well-balanced variety of information.

For purposes of our physical well-being, we have learned to ascertain what nutritional elements we need for ourselves and our bodies using the System 2 mode of thinking, and we can unconsciously avoid those foods that are dangerous to us using System 1.

Conversely, at this present moment, when we are at the dawn of the age of information excess, we have not yet learned how to take in from the vast amounts of information out there a well-balanced diet of that information that we need for our information health. Here lies the problem of the attention economy, which takes advantage of this asymmetry and encourages people to consume as much information (related to economic interests) as they can.

What is needed is (1) for users to be able to autonomously and independently choose what information to take in through user education and reforms to their own thinking, efforts on the part of corporations, and appropriate (but indirect) government support and (2) for an environment to be established in which anyone can assimilate basic information in the process of maintaining a democratic society.

This way of thinking has affinities with the right-to-know interpretation under the Constitution. The Supreme Court of Japan has stated:

"Since it is indispensable for the purpose that individuals form and develop their own thoughts and personalities and activate them in the social life, and also it is necessary in order to make truly effective the fundamental principle, namely, guarantee of free communication and exchange of thoughts and information in a democratic society, that individuals have opportunities to come in contact with and take in various opinions, knowledge and information freely, such freedom as to come in contact with and take in information etc. is naturally led by the purposes and effects of the provision [Paragraph 1, Article 21 of the Constitution], so to speak, as one of its derivative principles." Although it may seem as the freedom to take in only certain information ("the freedom to eat selectively") through one's own intent and for one's personal interests is protected under the right to pursue happiness in Article 13, the "freedom to know" that is protected as a

The definition of "information health" that we offer here is most definitely a provisional one. We expect to develop a more suitable definition through interdisciplinary and cross-disciplinary debate.

Repeta jiken hanketsu (Supreme Court decision in the Repeta case) (Supreme Court decision 1989.03.08, Minshu Vol. 43, No. 2, at 89).

Yodogo hajakku kiji massho jiken hanketsu (Supreme Court decision in the case of redacting newspaper articles on the Yodo-go hijacking incident) (Supreme Court decision 1983.06.22, Minshu Vol. 37, No. 5, at 793). Under this decision, the "freedom to read newspapers, books, and the like" that are the media for delivering various opinions are held to have been from Article 19 of the Constitution—which protects freedom of thought and conscience—and as a derivative principle of Article 21, which protects freedom of expression. It also argues that such freedom to read is in accord with the intent of

derivative principle of Paragraph 1, Article 21, has, as a consequence of this decision, been established as the freedom to take in "various opinions" in connection with one's own personal development and sustaining a democratic society. The "freedom to know" in this sense can be thought of as closely related to information health as a condition in which we acquire a degree of immunity (critical faculty) concerning fake news and the like by taking in a well-balanced variety of information.

(2) Announcement of Content Categories ("Calorie Chart")

To achieve information health, users must have the opportunity to make independent and anonymous choices when taking in content and information on a DPF.

Accordingly, the "components" and "ingredients" of that content and information should be made plain to users, along with what sort of balance the DPF is applying to present and distribute various types of content. For their own information health, users should be provided with indicators so they can independently and autonomously decide about the sorts of content and information they are taking in and what DPF to use.

This would resemble the charts that show calories and amounts of nutritional elements that manufacturers are obliged by the Food Labeling Act to place on their food products. We have come to think of such calorie charts as one resource to use when deciding whether or not to take in the product in question. In a digital society, it is crucial that the character and substance of content and information be made apparent and that users be informed of what sorts of information categories those fall within.

Also, in television broadcasting, specific regulations are in place under the principle of consistency among programs in the Broadcasting Act to not allow bias to enter into the creation of content. ¹⁶ These regulations based on the programming consistency principle have been deemed warranted because the airwaves are limited and because broadcast media has a strong social influence. Today, a small number (and de facto limited, in the sense that the options are limited) of DPFs have acquired enormous social influence as the primary channels for information distribution.

Accordingly, concerning some DPFs¹⁷, it would be desirable for them to not only recommend specific information aimed at appealing to a user's interest but also create algorithms that can provide users with the

Article 13, which establishes respect for the individual.

Broadcasting Act Article 106, Paragraph 1, stipulates that "excluding those based on special business plans, the basic broadcaster must establish cultural programs or educational programs and news programs, and entertainment programs and must maintain mutual consistency between the broadcast programs." At the time that broadcasters were relicensed for terrestrial television in 2008, based on inspection standards related to the Radio Act the Ministry of Internal Affairs and Communications, the condition that both the NHK General channel and private broadcasters must allot 10% or more of their weekly programming to educational programs and 20% or more to cultural programs for purposes of maintaining mutual consistency between broadcast programs was added. Under the obligations to publicly announce program categories, which was established in a 2010 revision, every year for six months starting in April, the programs broadcast during the third week of alternating months by the organizations in question would be categorized as "cultural programs," "educational programs," "news programs," and "entertainment programs." After the six months had passed, the organizations would then be obliged to publicly announce the breakdown as soon as possible.

DPF here does not necessarily refer to all DPFs. The DPFs for which "balance" is to be particularly expected refers to those that are especially large and have a high degree of "publicness." The definition will be submitted to future discussion.

capability of coming into well-balanced contact with content across a wide range of genres. Specifically, certain DPFs should—mindful of the example of the principle of consistency between broadcast programs—of their own initiative settle on specific ratios for the content they provide (for example, how they lay out their landing pages according to the balance among the coverage of political news, financial news, entertainment news, and sports news) and then present it in an easy-to-understand form. Doing so will make it easy for users to grasp what content is being provided to them and what information biases exist.

(3) Providing an "Information Checkup"

Users should be provided at regular intervals with the opportunity to conduct an "information checkup," so they can audit the state of their own information health. By getting such an "information checkup" of the individual's own volition, it is expected that this will make it apparent to the user what sorts of information they are coming into contact with and provide them with the motivation to change how they go about taking in information.

Information checkups, as presumed, will examine the diversity of the sources from which information is being taken in (to see if biases are being created) along with the credibility of those sources, and then, the results of these inspections will be presented as objective data. One imagines that software for calculating the situation and extent of being "infected" by filter bubble or echo chamber will also be put to use as part of the information checkup.

Under circumstances in which attention economies are becoming commonplace, such information checkups are recommended for people who feel uneasy about biases in the information they are taking in as well as those who want a balance in their information. Nevertheless, these checkups should not be imposed (the government will not be recognized as the practitioner, either). Also, when it comes to the collection and management of the private data that occur in conducting such information checkups, it will be necessary to be fully aware of user privacy, and the safety controls for that data must be thoroughgoing.

(4) Providing a Digital Diet

In the case that a user senses that there is a problem with their own information health, it would be desirable to provide the ability for them to be able to adjust the degree of personalization on a DPF on their own. In the case that the user becomes aware through the information checkup that there is a problem with how they go about taking in information and want to transform that behavior themselves, it will be necessary to prepare some feature for providing information in response. Also, in the case that the user believes that they are faced with a serious problem, it is possible that they might consult with a specialist like discussing diet management during medical consultation and, by their own decision, get the information that they take in under control.

Certainly, there may also be unsuitable digital diets, as is the case with actual diets. It will be necessary to

-

The specialists conceived of here should be people who are conversant in "information health studies" (see VI.). The question of how such specialists will be trained and certified will be a crucial issue for the future because of the danger of having their thought influenced by these specialists.

especially be careful about deceptive remarks in which such language as "think about this yourself" and "your thinking is biased" is used to lead people into a state of greater unhealth. Also, one must be especially vigilant about "ideological rectification" and "brainwashing" done in the name of a digital diet (naturally, the government will not be permitted to the practitioners here).

Still, further, the purpose of the digital diet is for users to take in a well-balanced variety of information and acquire immunity (critical faculty) toward fake news. It should be noted that this is a different concept from the so-called "digital detox," in which someone cuts themselves off from taking in information.

(5) Searching and Exploring for an Economic Structure to Replace the Attention Economy

Today, the space for public discourse is almost entirely under the sway of an economic model known as the attention economy, and it appears that not only DPF operators but also users, the existing mass media, and advertisers have all been caught up in its vortex. The various issues that we have brought up so far can all be said to be an inevitable result of it.

Accordingly, we should be groping radically toward an economic model of a sort that would replace today's attention economy. Obviously, this would come with considerable difficulties in both theory and practice. Nevertheless, we should not avoid this kind of radical investigation once it becomes evident that the fundamental reason for the problem lies in the economic structure.

This new economic structure that would replace the attention economy is not clear presently. It's a topic that will be explored in a cross-disciplinary debate that includes economic perspectives. What should be done first is to investigate the structure of the subsystems that basically support the framework for the attention economy, where content is assessed on the basis of its quality and not a simple indicator like page views, and prices are paid on the basis of that assessment. Also, one can think of building a subsystem with fixed charges on users on the basis of subscriptions that are prevalent today in every area.

III. The Ground Rules for Users

(1) Understand the Current Information Environment

Users should understand the "attention economy" economic model and be aware of the fact that we are embedded within it. This is the first step toward solving the problem.

(2) Be Aware of Information Health

Users should be aware of their own "information health" like they would be of their physical health. This may be thought of as having the following merits.

(a) Not Being Naive About Fake News

When someone is overly selective in their information diet, their immunity to fake news and the like is reduced and their vulnerability grows. By being aware of one's information health, the individual can take in a well-balanced variety of information that includes opposing opinions and thus acquire immunity to fake news.

(b) Not Becoming One of the "Information Poor"

People who have certain difficulties when it comes to obtaining and using information are sometimes described as being "information poor." By controlling and adjusting oneself the personalization of information being taken in and coming into contact with various information of the sort that was removed by one's filter bubble, it will become possible for that person to also access the basic information they should have been taking in. This can prevent turning someone into one of the information poor. Also, through this, it will prevent the economic and social losses that information disparities deliver, and lead to improvements in one's quality of life.

(c) Improvements in Mental Health

Continuing to take in only specific information makes it more likely for people's thinking to become increasingly fixed, rigid, and incapable of accepting others' ideas. Also, the so-called "Facebook Files" point out the possibility of someone's feelings of self-negation increasing and the mental health of the younger users being worsened by being shown exaggerated body shapes and so forth. ¹⁹ One can expect awareness of "information health" to have an effect that prevents any worsening of this sort in user mental health.

Moreover, even slander directed toward specific persons can be seen as a problem created by selective information diets, including being provided with fragmentary information. Such slander naturally affects the mental health of the victim, but it can also harm that of the perpetrator and their life itself. One can anticipate that by people's information health having been achieved, this will prevent situations that result in their becoming slander perpetrators and victims.

(d) Improving the Health of Democracy

Regardless of Facebook itself being aware of these effects on mental health, clearly, from the internal documents from Frances Haugen, the company had neglected taking countermeasures. https://www.wsj.com/articles/the-facebook-files-11631713039>

Our democracy is premised on citizens (sovereign individuals) engaging in "dialog" based on System 2 thinking (mature deliberation) and engaging in political participation independently and autonomously. Toward that end, it will be necessary to achieve a situation in which they can come into contact with diverse views and opinions and lack bias in the information that is being taken in. The American legal scholar Cass R. Sunstein argued that the preconditions for sustaining a democracy are (1) being exposed to the different outlooks of others and (2) the members of the community having shared experiences.²⁰

If the collapse of democracy should go further, there is the danger of a dramatic drop in the level of the freedoms that we enjoy—for example, we will no longer have control over authority, and our basic human rights will be threatened. Achieving information health and seeking to make our democracy sound is quite crucial when it comes to sustaining our freedom in the medium term to long term.

Hence, as the foregoing indicates, awareness of information health has significant benefits from the perspective of maintaining a free and democratic society.

-

²⁰ Refer to previous footnote 10.

IV. The Ground Rules for Operators

1. The Ground Rules for DPF Operators

(1) The Status of Independent Initiatives by DPF Operators

Even until today, the various platforms have already been pursuing their own initiatives conducive toward achieving information health. In Japan, efforts such as Yahoo News' model for prioritizing constructive comments introduced for its comments' sections²¹ have attracted notice in connection with effectuating a platform for diverse speech. Moreover, the Safer Internet Association (SIA) industrial group is proactively engaged in solving problems that face DPFs.

Below, we aim to present some of the necessary basic principles toward encouraging further initiatives by DPF operators.

(2) Consideration for "Cognitive Liberty"

DPF operators must give full consideration to the cognitive liberty of their users and not apply AI and the like to "hack" the individual's thoughts for purposes that are inconsistent with the user's interests (prohibition of mind hacking).

"Cognitive liberty" is an idea that was first advocated for in the field of neurolaw. It calls for protecting the integrity (impenetrability) of the central nervous system in the body, centered on the brain. The debate focused on such topics as measuring nervous activity through magnetic resonance imaging and electroencephalograms and electromagnetic interventions in the nervous system. Nevertheless, viewing the concept of "cognitive processes" in the broader sense, the problems that elements of the external environment such as the information space and architecture present to an individual's cognition can also be seen falling within the scope of "cognitive liberty." From the perspective of a brain's plasticity, the transformation to the nervous system that arises because of ongoing exposure to mind hacking can also be a problem. DPF operators must not take advantage of the dramatically developing knowledge in cognitive neuroscience in ways that infringe on users' cognitive liberty.

Additionally, the European Union has announced that it is totally prohibiting "the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behavior."²²

²¹ Yahoo Japan Corporation press release, "Yahoo! JAPAN starts providing free access to API for natural language processing model (AI) using deep learning introduced in response to inappropriate comments" September 18, 2020.

https://about.yahoo.co.jp/pr/release/2020/09/18a/

European Commission, "Proposal for a Regulation of the European Parliament and of the Council: Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts," Article 5(a) https://ec.europa.eu/newsroom/dae/document.cfm?doc id=75788>

(3) Filter Bubble Measures

Operators should come up with methods for delivering content and other such schemes that would allow users to come into contact with a well-balanced variety of information.

The provision of personalized content does have advantages both for users, who want to get a hold of only the information they want from vast amounts of information available, and for operators, who want to raise click ratios and increase engagement.²³ Nevertheless, by coming into contact with only information that has been personalized for them by operators, there is the possibility that users will be cut off from information that has been assessed as something that the user concerned "does not want," and that they will be prevented from taking in various information as well as basic information.

Accordingly, the functionality should be created to provide users with various information—such as by randomly or even deliberately providing them with nonpersonalized information—so that they are not trapped by external forces in a filter bubble.

(4) Dealing with Fake News

Operators should work up effective measures for dealing with fake news. With those measures, they must be mindful in particular of the following points.

(a) The Definition of Fake News and Fact-Checking Organizations

The number one issue when it comes to measures for dealing with fake news is how does one define and identify "fake news"? For example, the question of whether something is "fake" is not clear at the moment when that information is released. Furthermore, it should be noted that because of cognitive bias, inconvenient facts that users do not want to accept are seen as overlapping with fake news.

The second issue is that of who determines what fake news is. Presently, there are fact-checking groups that function as the judges of fake news, but the trustworthiness of those groups themselves can be a problem. Guaranteeing that fact-checking itself is sufficiently impartial and transparent through the monitoring of such groups is essential.

In this connection, the International Fact-Checking Network (IFCN) has presented the following five principles for its own network.²⁴

- 1. Nonpartisanship and Fairness
- 2. Transparency of Sources
- 3. Transparency of Funding and Organization
- Transparency of Methodology
- 5. Open and Honest Corrections Policy

According to data Facebook released in 2013, the total number of posts that users received from friends and followers stood at 1,500 per day on average. Nevertheless, algorithms filtered down those posts to display approximately 300 on average. In short, what is displayed is 20% of the total posts. *See*, Backstrom, L., "News Feed FYI: A Window Into News Feed," Facebook, 2013. https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed.

²⁴ See, International Fact-Checking Network, "The commitments of the code of principles."
https://www.ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles>

(b) Effective Display of Fact-Checked Articles

Fact-checked articles are most effective when they are shown to people who have actually seen fake news. DPF operators should build algorithms that privilege showing fact-checked articles on the news item concerned to people who have seen fake news on the topic. In this case, it will be necessary for fact-checked articles to be provided as free content.

Understandably, another aspect of fake news is that it is also disseminated through conversations in the real world. Accordingly, in the case of fake news that has been disseminated widely through the world generally, operators should investigate means for widely presenting the relevant fact-checked articles to people other than their readers.

(c) Showing Media Credibility

One can conceive of constructing a mechanism for letting people know about certain media whose credibility has been evaluated, such as by clearly indicating ownership and how the entity's editorial structure is organized. Nevertheless, such evaluations should be entrusted to neutral third-party organizations and their standards for judgment publicly released. They absolutely should not be selected arbitrarily by DPF operators.²⁵

(d) Architecture-Based Measures

It will be necessary to investigate introducing architecture ("nudges") of the sort that encourages self-control in the easy dissemination of information and operation of the careful System 2 mode of thinking. Examples of this are the measures Twitter has implemented to display the links to articles in retweets that encourage users to view the site in question,²⁶ as well as features that display alerts when someone attempts to make an offensive post and encourage the user to alter or delete the post (these have already been introduced to the English-language versions of Twitter and TikTok).

(e) Rules for Removal

For those instances in which slander or grave disinformation is being displayed or for when a series of posts are being repeated over and over by the same account (including bots), measures for removing it will be necessary, such as having that information deleted and placing restrictions on the account in question.

Nevertheless, these removal measures should be limited to those cases in which the problem cannot be resolved even after less-restrictive means have been attempted such as architecture-based countermeasures (see item (d)) and direct warnings to users.

The Journalism Trust Initiative (JTI)—a project led by Reporters Sans Frontières (RSF) whose objective is to create a health information space—has created a framework for indicating the trustworthiness of media as determined through self-assessments and making it public after it goes through an external audit conducted by a certifying organization. The self-assessments consider such factors as (1) disclosing information about ownership, including the members of its board of directors and revenue sources, and (2) disclosing what its editorial guidelines are, who are the editorial contacts entrusted with this, and what their accountability processes are. The standards for evaluating this credibility were released in 2019 as the CEN Workshop Agreement (CWA), which does not have the binding power of the European Committee for Standardization (CEN). See, CWA NO.17493 https://www.jti-app.com/footer/cwa

²⁶ Twitter Japan, 2020.9.25 < https://twitter.com/TwitterJP/status/1309289971690348545>.

(5) Securing the Freedom to Choose Algorithms

Excluding those exceptional periods (such as elections) discussed in (6), DPF operators should make it possible for users to freely select the degree of personalization in the display algorithms. The freedom to select algorithms in this fashion is crucial to users independently maintaining and managing their own information health. Moreover, when it comes to choosing algorithms, DPF operators must have user-friendly UX and UI designs.

Additionally, as one of the premises of these choices, DPF operators should create conditions in which the users can also view and comprehend the degree of personalization in their current display algorithms.

(6) The Obligation to Switch Algorithms

In the cases in which the following exceptional situations are deemed to be in effect, DPF operators will be obliged to switch their display algorithms.

(a) Elections Mode

There are particular demands for user information health during periods when crucial events for sustaining democracy are underway such as elections for public office and popular referendums on constitutional revision. DPF operators will need to guarantee and privilege a certain space when it comes to information that is necessary for appropriate voting behavior, such as materials related to policies of parties and candidates and the major issues of contention in the election.

The Public Offices Election Act has already laid down various regulations when it comes to electioneering such as restrictions on the use of campaign sound trucks and a ban on door-to-door canvassing. These restrictions are based on the understanding that election campaigns "should be considered not as opportunities in which candidates having diverse arguments can freely compete with each other under minimum necessary constraints but as opportunities in which candidates conduct campaigns in accordance with rules that are set to ensure the fairness of the election."²⁷ Moreover, certain broadcast operators are obligated to transmit campaign broadcasts so that eligible voters can attempt to freely and equally obtain information regarding political parties and candidates.

DPF operators should also separate themselves from the logic of the attention economy and prepare "election mode" algorithms and UI/UX designs that emphasize the fairness of elections for those specific periods when an election or popular referendum on constitutional revision is taking place. Specifically, they will be called upon to give thought to how information about parties and candidates will spread widely and impartially akin to campaign broadcasts; privilege the display of information from media deemed highly credible; implement an architecture that controls the easy spread of information (i.e., restrictions on sharing features such as retweeting); restrict microtargeted advertising and the like with political objectives; and establish and monitor compliance with concrete guidelines for "election modes."

(b) Disasters Mode

²⁷ Kobetsu homon kinshi kitei hanketsu (Supreme Court decision on the provision prohibiting door-to-door visits), Masami Ito concurring opinion (Supreme Court decision 1981.07.21, Keishu Vol. 35, No. 5, at 568).

When a disaster occurs, the display algorithms must be switched and the DPF operator must strive to promptly report accurate information and information that helps toward encouraging appropriate evacuations. Specifically, they must promptly release highly credible information from the government, local public entities, and news organizations and also broadly communicate the information needed for evacuations such as the issuance of evacuation directives and orders as well as information regarding evacuation sites. Also, they will need to prevent the rapid spread of malicious disinformation that piggybacks on the confusion, as well as information of the sort that will invite panic and make suitable evacuation efforts difficult.

(c) Pandemic Mode

In those cases when a severe problem emerges in the area of public health such as the spread of infectious disease, operators will need to switch their display algorithms and strive to promptly deliver accurate information and encourage appropriate behavior. As with the Disaster Mode, it will be necessary to clamp down on the spread of disinformation and information that lacks scientific backing. Conversely, in settings where individual decision making is emphasized such as over getting vaccinations, as with the Elections Mode, the operator must distribute as relevant the information necessary for individual decision making.

(d) Minors Protection Mode

Concerning minors, when it comes to their own efforts to take in information, it is quite possible that they lack adequate capacity to make judgments and will be exposed to various risks. Accordingly, for minors, operators should investigate implementing paternalistic algorithms that consider their information health. Particularly, given the capacity of minors for judgments, operators will need to consider the displaying of sexual content and speculative content.

Conversely, minors' right to know should also be considered. DPF operators should make it possible to switch gradually through various algorithms corresponding to the age and attributes that the minors themselves or their guardians have registered. For example, the algorithms that are used should change for junior and senior high school students.

(7) Algorithm Transparency

DPF operators must use appropriate methods to make the algorithms that determine how content is displayed (for example, the ranking for displaying news articles and user comments) transparent. This will make it possible to make plain what each DPF operator's way of thinking and stance is concerning information health and could stimulate marketplace competition over information health (for example, the DPF operator that has been ascertained as using malicious algorithms that poisonously lure users may be weeded out in the marketplace). Also, it will become possible for the media and journalists to independently decide which DPFs they provide their own articles to.

Moreover, DPF operators should regularly draft reports on the algorithms' rules and operational status and make them available to the public to guarantee transparency regarding how they display content and manage advertisements. These reports should be made in conformity with the frameworks that private groups have normalized and standardized (see V. (5)).

(8) Systems for Taking Responsibility and Governance

(a) Keeping Management and Editorial Separate

The provision and delivery of content by DPF operators are handled in collaboration among various departments, such as the department that develops and builds algorithms, the departments that select articles and advertising, and the department that designs UIs. Accordingly, the places where responsibilities lie can easily become unclear. For this reason, DPF operators should make the persons responsible for editorial content independent from those persons responsible for the management and clarify where responsibilities lie.

Doing so will build a firewall of sorts between management and editorial and make it possible to prevent unwarranted interference by management and external stakeholders in the provision and distribution of content.

(b) Establishing an Ethics Committee

DPF operators should form ethics committees composed of outside experts who will check to confirm that nothing inappropriate is present that infringes on human dignity and respect for the individual when it comes to their collection and analysis of personal data and the "nudges" made by their UI/UX design. Also, such ethics committees should ensure that psychology and cognitive neuroscience are being applied ethically and appropriately.

(c) Establishing a Content Committee

DPF operators should establish internal content committees composed of outside experts. Such content committees would administer and monitor such matters the display or nondisplay of articles and content on the DPF; the soundness of the screening done on sources of content (the so-called media screening); the soundness of the advertising screening done by that screening department; and the suitability of how PR content is presented.

Furthermore, when it comes to how individual matters are handled such as stopping distribution and the results of media screenings, if the operator who is a content source issues a complaint, the content committee should have certain powers to adjudicate the matter.

(d) Obligation to Ascertain Actual Conditions

DPF operators must appropriately grasp the realities of problems that arise in connection with their own platforms. Also, the results of the surveys they have conducted to get a sense of the facts should be regularly made public, as long as they do not infringe on business secrets. Also, concerning the details of the survey results (including raw data), to the extent possible, these should be shared with news organizations and researchers.

Moreover, DPF operators should strive to quickly discover and quickly deal with the risks they face by developing systems for whistle-blowing.

2. The Ground Rules for the Mass Media

(1) The Principles of Mass Media

(a) Guaranteeing the Credibility of Mass Media

Recently, the decline of public trust in the mass media has become a problem. Given that "in our democratic society ... news reports offer important materials for the people to make their judgments in participating in the government," the mass media that are the carriers of those reports must continue to have a presence that the people rely on order to attain a healthy democracy. To gain and restore trust, the mass media should not lose sight of the substance of their coverage, in guaranteeing quality through fact-based reporting and editorial work (peer review).

(b) Securing Distance from the Attention Economy

Within the logic of the attention economy, even among the mass media to get page views, we can see a tendency to emphasize articles with high entertainment value that will catch the eye as well as articles of a highly stimulating nature that will evoke a System 1 response. Consequently, they fall into a vicious cycle that invites a decline in the quality of the articles that get carried and reduces the credibility of the mass media. The mass media must consider their own publicness and, in their coverage of areas of significant public interest, work hard to distance themselves from the logic of the attention economy.

(c) Preventing Feedback Loops

Previously, the dissemination of fake news and the like had been thought of as occurring mainly over social media. Nonetheless, today, it has been observed that reporting by the mass media of fake news on social media has produced a synergistic effect and further accelerated the dissemination of fake news (feedback loop). ²⁹ The mass media must be conscious of how much power they have to influence society and strive to prevent this sort of fake news feedback loop.

(2) The Duty of Journalists

Individual journalists are the core actors in reportage and are an indispensable presence in covering stories. This profession has publicness in the same way that the mass media do as organizations. Journalists should dispassionately examine the attention economy model as creators and distributors of information, and execute their duties as autonomous professionals with publicness.

^{...}

²⁸ Hakata-eki jiken kettei (Ruling upon the case of the so-called Hakata railway station case) (Supreme Court decision 1969.11.26, Keishu Vol. 23, No. 11, at 1490).

²⁹ Claire Wardle of the US nonprofit First Draft News in her "Trumpet of Amplification" graphic showing the process by which fake news spreads puts the professional media at the last step. See, Wardle, C., "5 Lessons for Reporting in an Age of Disinformation," First Draft, 2018. https://firstdraftnews.org/articles/5-lessons-for-reporting-in-an-age-of-disinformation/

V. The Ground Rules for the Government

(1) The Constitutional Obligation to Provide "Lateral" Assistance to Initiatives Aimed at Information Health

The government has to help the people maintain "the minimum standards of wholesome and cultured living" (Constitution Article 25) and ensure their "right to know" (Article 21) by maintaining diversity in the media and helping to distribute varied information. On the basis of these facts, one can presume that it has constitutional duty to create a competitive environment that will support a plurality of media and provide "lateral" assistance to varied initiatives undertaken by DPF operators, all to support and improve the people's information health.

During those special periods when the elections for public office and popular referendums on constitutional revisions that are at the core of democracy are being held, there are particular needs for the people to be able to come into contact with a well-balanced variety of information and opinions and to have immunity (critical faculties) toward fake news. During these special periods and unlike at other times, the government would be called upon to work toward maintaining healthy and deliberative spaces for public discourse by, for example, encouraging DPF operators to give special consideration to the information health of their users.

(2) Prohibition of Direct and Excessive Government Intervention

The government should not intervene directly over the public's information health or impose such health on them. Also, the government must not define what such health means. There are concerns that such direct intervention could invite "information health fascism" and instead distort democracy. This point cannot be stressed enough.

Also, excessive interventions concerning DPF operators will not be permitted. The DPF operators themselves are private businesses that enjoy the freedom of expression (Article 21) and freedom to choose their occupation (Article 22). There is concern that if the government were to domesticate DPF operators through excessive intervention, it could lead in real terms to censorship.

(3) DPF Operator Guarantees of Transparency and Accountability

As has already been noted, DPF operators must guarantee transparency in their businesses and be accountable to the public (see IV.1.(7)). The government should design systems that would effectively guarantee that DPF operators would fulfill these obligations. Basically, although the autonomy of DPF operators should be respected, legal involvement should be considered exceptionally if it is deemed necessary.

(4) Maintaining an Environment for Just and Fair Dealings

To prevent the spaces for public discourse from becoming wholly devoted to the attention economy and creating a growing glut of stimulative content designed to elicit reflex responses, content provided by

newspapers and other existing media in which people can trust must be distributed substantially on DPFs and be put into a situation where users can take it in. Toward that end, the existing media must have a certain degree of bargaining strength concerning DPF operators. In cases where it is acknowledged that in reality, an existing media organization does not have the bargaining strength, the government must arrange an environment for just and fair dealings between DPF operators and the media.

(5) Carrying Out Information Literacy Teaching and Awareness Raising Activities

As III.(2) indicates, to achieve information health, users must be information literate. Accordingly, the government must proactively engage in information literacy education in primary and secondary education.

Furthermore, the government should encourage initiatives to enhance information literacy among adults.

VI. Prospects for the Future

The contribution of specialist computer scientists is vital toward constructing the algorithms that bring about information health. Computer scientists should direct their efforts not toward creating systems connected toward buttressing the attention economy but rather those that will be conducive toward creating information health, such as the "information checkup" and "digital diet."

Also, specialists from other fields (for example, legal studies, economics, social psychology, and psychiatry) will need to participate in interdisciplinary and cross-disciplinary discussions directed toward making informed health a suitable reality.

In the future, it is also presumed that based on cooperation among these specialists, a field of "information health studies" will be established to create an information environment that is healthy both for individual autonomy and democracy. In terms of the scholarly issues concerning "information health studies," for the moment, the following items may be raised:

- (a) Refine the definitions of "information health" and "information ill-health"
- (b) Clarify what the "nutritional elements" are in the information
- (c) Normalize and standardize what and how "nutritional elements" are to be displayed
- (d) Identify the specific evils produced by "information ill-health" and the specific merits that "information health" produces
- (e) Elucidate the appropriate methods for providing information for "information health"
- (f) Develop specific methods for conducting "information checkups"
- (g) Explore specific methods for ameliorating "information ill-health"
- (h) Construct an advertising system that accords with "information health"
- (i) Classify and sort the diverse range of DPF operators, and examine which of those are DPF operators who should have special responsibility for "information health"
- (j) Examine economic models to replace the attention economy

*List of Co-Signers (as of January 6, 2022)

Haluna Kawashima (Project Associate Professor, Keio University Global Research Institute)

Tonghwi Soh (Smart News Media Research Institute and Attorney)

Teppei Koguchi (Professor, College of Information, Academic Institute, Shizuoka University)

Kazuhiro Taira (Professor, College of Arts and Sciences, J. F. Oberlin University)

Katsue Nagakura (Journalist, Nikkei Business Publications)

Yusuke Narita (Assistant Professor, Department of Economics, Yale University)

Hiroyuki Fujishiro (Professor, Faculty of Social Sciences, Hosei University)

Asako Miura (Professor, Graduate School of Human Sciences, Osaka University)

Eijiro Mizutani (Associate Professor, Faculty of Sociology, Kansai University)

Shinichi Yamaguchi (Associate Professor, Center for Global Communications [GLOCOM], International University of Japan)

*Research Assistants

Masatoshi Kokubo (Doctoral Program, Graduate School of Law, Keio University, and a member of Keio University Global Research Institute)

Haruki Kadotani, Nozomi Jinnai, and Ryotaro Soma (4th-year students, Department of Law, Faculty of Law, Keio University)