KGRI
Keio University Global Research Institute

KGRI Working Papers

No.1

Quest for Broadly Acceptable Architecture for Data Governance

-A Man-Machine Conviviality Approach-

Version1.0

November 2019

Jiro KOKURYO(i), Jonathan CAVE(ii), David FARBER(i), Hiroaki MIYATA(i), Jun MURAI(i), Takehiro OHYA(i), and Tatsuhiko YAMAMOTO(i)

(i) Keio University, Cyber Civilization Research Center (CCRC)

(ii) Alan Turing Institute and University of Warwick (CCRC adviser)

Keio University Global Research Institute

Quest for Broadly Acceptable Architecture for Data Governance
-A Man-Machine Conviviality Approach--

Jiro KOKURYO, Jonathan CAVE, David FARBER, Hiroaki MIYATA, Jun MURAI, Takehiro OHYA, and Tatsuhiko YAMAMOTO,
Keio University, Cyber Civilization Research Center

Abstract:

In order to protect human dignity given the ever-growing power of computers and data aggregation, we must adopt a new philosophy and architecture for the governance of data. We step beyond the conventional assumption of human monopoly of intelligence and envision integrated man-machine agents that will emerge to safeguard personal data on a firm basis of trust. A new "cyber civilization" is dawning, in which humans live with machines in conviviality, and we must develop governance structures that address this reality. A few guiding principles for the design of the architecture are proposed.

1. An Alternative Strategy for the Protection of Human Dignity

The dilemma between beneficial exploitation of data and maintenance of privacy aimed at the protection of human dignity is deepening. This issue has arisen in step with the growing power of computers and networks but also our waning ability to know, let alone understand or control their effects. It would not be an exaggeration to say that policy makers and operators handling data around the world see themselves confronted with a stark choice between 1) abandoning human rights in favor of the power of collective data, 2) protecting individual rights even at the cost of falling behind competing countries in the exploitation of data for machine learning and collective intelligence, or 3) finding new ways to think about and guard human dignity by way of use of collective data.

This is an introductory note which explores ways to realize the third option by appropriately designing the architecture of data governance. The primary argument in this paper is that it should be possible to design an “architecture” for data governance in which citizens can “entrust” their data without fear of their misuse[1]. This idea is actively debated in Japan with a view to

[1] “Misuse” here can take many forms, but here, we broadly assume situations in which people are unfairly discriminated against, or lose autonomy in thoughts and action, based on data collected by others.

substantiate the framework of "Data Free Flow with Trust"(DFFT)(METI, 2019), which the Japanese government advocated at the 2019 G20 Summit as its chair.

Architecture here is defined as institutional/technological artifacts that provide an environment (Matsuo et al. 2017: Lessig, 2006) which nudges (Sunstein, 2014) actors to behave in desirable ways, where desirability in this paper is assessed by the degree to which human dignity is mutually honored. In the sense that it relies on persuasion rather than enforcement, this approach can be categorized as a "soft law" approach. It may, as an example, include the adoption of machine assisted "personal data artificial intelligence (AI) agents," as considered in IEEE P7006" (Merchant, 2019). Personal data AI agents that are given missions, incentives and controllability (the meaning of which includes visibility and explicability) to represent the interest of individuals can collaboratively maximize the social benefits of shared information.

One might note that an underlying perception in our proposal is that the dogma of "self-control of private information" (Waldman, 2018), without the involvement of machines, e.g. algorithmic content mediation, is becoming unrealistic given the flood of information individuals face. A collaborative relation between "trustworthy" systems and individuals is necessary to nurture harmonious man-machine conviviality.

We come to this perception resigned that humans no longer enjoy a monopoly of intellect. We now live in a world where automated intellectual capabilities that are detached from human physical existence are starting to make decisions that humans cannot catch up to. We know that this is threatening the philosophical and institutional foundations of modern western societies which assume human monopoly of intellect and rely on individual (autonomy and) consent as pillars of social order. As much as we would like to dismiss it, a new reality is setting in. Therefore, we should somehow recognize and position machine intelligence, as well as new forms of (machine-mediated) connected human intelligence, in the architecture of governance.

Boldness in changing our world view seems required to realize the obvious societal benefits of making use of data. Let's put aside temporarily questions of legitimacy and agency of commercial enterprises accumulating and using data in their own interests. Though potentially running counter to the interests of consumers and other data subjects, there also seems to be a potential for "common" good to result from data accumulation and use. For example, sharing personal health records (PHR) clearly offers great benefits through diagnosing and curing individual patients, improving medicine more generally and advancing related sciences (World Health Organization, 2019).

Japan, though the pioneer in adopting western individualist and liberal democratic principles in Asia, has also been grappling with the communal traditions of Asian society. For instance, we might recognize that a society based on "altruistic" virtue, where people are rewarded for respecting the dignity of others and for their contribution to society, is in fact well suited to a world of shared data and co-created value. This is why we cannot simply dismiss state governance of data as "Digital Leninism" (Heilman, 2016).

In an even longer historical and cultural context, Japan also has an animistic tradition that considers humans as a "part" of nature, rather than regarding them at a "top" or "central" position in it (Jensen, 2013). This contrasts with the societal hierarchies of western and other civilizations which regard human intellect as a supreme value. The Japanese are also much more receptive to the idea of recognizing the "self" in non-human beings or even inanimate matter and thus able to accommodate changes in nature.

While recognizing such alternative systems, we are also cognizant of the great achievements of democratic systems in delivering stable, open, and fair societies that embrace human dignity. Even in the East, where collective welfare is emphasized, notions of honor and autonomy of individuals are upheld as values to be protected. Clearly, the misuse of private information can result in grave consequences, most notably discrimination offending personal dignity. Used to point at others, surveillance can easily lead to oppression. Therefore, while democracy in East Asia may be seen under threat, personal freedom and dignity must be protected. There must be a way of maintaining the autonomy of individuals while harvesting the value of shared data. In accordance with the data platform "Person-centered Open Platform for wellbeing" originally introduced by the Japanese government, our team at Keio University humbly remains human-centric, not rightfully so but because we selfishly desire to be so.

2. Governance in "Cyber Civilization"

We believe that the root of fundamental change currently taking place lies with the network externalities that underpin the digital economy (Katz and Shapiro, 1985: Schilling, 2017). A force that was not fully recognized in conventional industrial societies is at work, fundamentally changing the way societies function. We define network externality in this paper as the tendency of goods and services used by networked individuals – especially those that help them interact – to become more attractive as others adopt them. This has several consequences, including undersupply of connectivity (because individual adoption or connection decisions may not consider benefits to others) and undersupply of competition (because one or a few interoperable products may tip into dominance regardless of whether they are the 'best').

As the term "network" externality suggests, decisions to join or use a network typically affect other users. For the sake of readers who are not familiar with the term, a simplistic illustration follows – though it should not be over-generalized. A classic telephone (connecting only two people by voice at a time) has no value if there is no-one to call. For this illustration, we make a challengeable assumption that the value of a network is proportional to the number of connections it enables. A second phone allows one connection. A third phone gives three connections; n phones allows $n^2/2$ possible pairwise conversations. One notices here that the network value increases more than proportionally with the number of users.

Such more-than-linear increases in value with the growing 'size' of the market were rare in the "material" world of industrial value creation, where value (to a competitive seller) is linearly proportional to the number of units of output sold (the constant of proportionality being the price). In contrast, the nature of the network has been one of the reasons why monopolies, whether firms or states, tend to emerge. Another reason is that a firm with a larger number of users will be more attractive to new users or those patronizing smaller suppliers, and will tend to grow until it absorbs the whole market. As acronyms representing such dominant firms (e.g. GAFA or FANG) are synonymous with "power," platforms that accumulate data are gaining unprecedented power to control and manipulate society. By offering protocols for connection between (and among) suppliers and users, these platform giants are now capable of gathering data and exercising power on both the supply and demand sides of electronic markets for content, e-services, and so on. (Eisenmann et al, 2006). Users are offered access to products that allow them to connect to more and more peers and merchants; merchants not only gain access to a vast installed base of consumers, but also to information concerning their preferences and behavior. This is not just a matter of quantity; choice among more suppliers on the same platform allows users to more closely match their preferences and increases incentives for suppliers to innovate constantly. Suppliers able to use vast data samples can make higher-quality estimates of consumer preferences than those limited to small and selective samples. Patently, in more and more fundamental ways, the governance mechanism that has controlled the "material" world seems unable to control these new realities adequately.

While we are mindful of the danger of simplistic analogy, we note that similar conclusions apply to exchanges of data. This reflects the recognition that data only has value (meaning[2]) in a given context, i.e., a network of data within a certain structure and in connection with other data. While an isolated datum does not generate much value, a collection of data exhibiting certain patterns

[2] We perceive "meaning" in the context of data as "information" in this paper.

does. This is felt with a keen sense of urgency in healthcare, where data tend to be scattered and incomplete. Partial data collection may be useful, but far less than comprehensive and complete data sets. The whole is clearly greater than the sum of its parts.

A careful analysis of the powers and limitations of network effects should be conducted separately; We note for instance that not all connections create value; some destroy value, while others help one party but hurt others. Moreover, complexity and cost limit value; not all benefits (even potential ones) scale up without limit. At the same time, network effects remains a powerful concept to explain how digital economies may behave differently from the material economies of industrial societies – and why their governance may therefore need to change as well.

We believe that these changes are significant enough to herald the emergence of a new paradigm. We call it Cyber Civilization. Here we define civilization as the pattern of behavior and a norm developed by (1) core technology that drives society, (2) core wealth (or source of power) pursued by society, and (3) institutional structure that governs the way technologies and institutions are used. We call it cyber civilization because it consists of a complex of man and machine interacting to form cybernetic beings. (Wiener, 1950)

From this perspective, we understand that we are in the midst of a shift away from industrial civilization in which energy systems had been the core technology. In this older civilization, money (which was the token in exchange for material industrial output) constituted core wealth. The market-based democratic system provided the institutional foundation, with ownership of material objects by individuals and the exchange of ownership being its key mechanisms. We underline here that ownership of property (the unconditional and exclusive access right to the use and utility of an object) has been sacrosanct in industrial civilization.

Following this line of thought, digital technology forms the core technology in cyber civilization, and data (or more precisely, the flow of data through evolving networks) constitutes its wealth. Governance structure, however, has not yet emerged with any clarity.

Thus, governance structure, or the architecture of cyber civilization, becomes an urgent central question. The General Data Protection Regulation (GDPR) can be seen as an attempt to tackle some of the issues arising in an increasingly data driven society (Haug, 2018). However, its underlying philosophy seems to remain focused on the protection of individualist modernity. We would like to see the latter evolve to address the new realities of the connected world.

We know the required architecture will have to promote the benefit of data sharing beyond borders, but that it will not arise of its own accord. The recent emergence of sharing and circular economy notions at the same time as security precautions are increased seems to be indicative of this. It reflects the nature of cyber civilization in which the value of data increases when connected with other data to form meaning.

At the same time, sharing of data should not encroach on human dignity and free will. Therein lies the challenge as far as governance of cyber civilization is concerned. Mere protection of old norms will fail to take advantage of new opportunities offered by this world in which machines help humans to explore new ways of using data to unlock the value hidden in our interactions. Governance should go beyond simply protecting the modern industrial norm and squarely address the nature of the new civilization.

3. Architecture for Trust

What design principles should characterize the data governance architecture in cyber civilization? Based on the above discussion, we argue the architecture should aim at creating personal data agents (PDA), that employ the power of systems in ways that people can trust. Data agents apply their technical competence to maximize the value of data sharing while protecting the privacy of clients.

The question, then, boils down to the design of an architecture that allows individuals to entrust their data to the agent. Here we propose the following four principles.

(1) Avoiding Conflicts of Interest

A major flaw with the existing major commercial platforms is that they seem to honor the economic interests of paying clients (most notably advertisers) more than the interests of the legitimate holders of control of personal information. The incentive structure involved in "target marketing" business models seems to disregard the welfare of likely victims of personal information mishandling. Thus, avoidance of conflict of interest should be a primary design rule of the architecture.

(2) Transparency

While auditing all machine handling of data may not be possible, the algorithms and data used should be auditable by a qualified third party. Here again, we should anticipate that it will be machines that will be assessing algorithmic biases of other machines. We envision a world in

which checks and balances are provided by multiple "beings," i.e., integrations of humans and machines (i.e., cyborgs). Such renewed forms of transparency may, for example, enable a government or a peer to certify data handling organizations as trustworthy fiduciaries of the citizens.

(3) Substitutability

Predictably, life without personal information agents will become intolerably inconvenient. On the other hand, monopolies may deprive us of our freedom if there are no substitutes and we no longer have any reason to strive to further our interests. We must ensure against such undesirable outcomes. To ensure substitutability, the freedom of individuals to move their data from one agent to another, i.e., data portability, will have to be included in the design rules.

(4) Sustainability

Whatever architecture we adopt has to be sustainable. We must be cognizant of how the enormous power of target marketing to generate profit and provide free services has had users give up their privacy unguardedly. Authors of this paper are sympathetic to profit making services handling personal data, to the extent they provide service purely in the interest of the individuals that use those services. AI agents that connect consumers with desired services should be able to get their rewards. Such rewards will have to be proportional to the economic benefits that the consumers receive, rather than to sales gained by the sponsors.

Substantiating these principles will have to be a continuing, evolutionary and forward-looking exercise, as we can expect technologies to advance very quickly and in ways that reflect current governance. Their enforcement requires multi-stakeholder participation, from commercial entities to governments, to be subject to monitoring.

4. In Conclusion

Currently, tests are being conducted in Japan around the notion of "information bank," to which "fiduciary loans" of information are being made on a technical and institutional trial basis. The notion of personal data AI agent has also been actively debated in this connection. For these notions to take firm root in society, a solid philosophical foundation seems essential. Global discussion is also important to build systems that at a minimum are compatible across national boundaries.

References

Cavoukian, A. P. D. (2013). Privacy by design: Leadership, methods, and results in “European Data Protection: Coming of Age,” pp. 175-202.

Eisenmann, T., et al. (2006). "Strategies for two-sided markets." Harvard Business Review 84(10): 92-101+149.

Haug C.J. (2018). Turning the Tables – The New European General Data Protection Regulation. New England Journal of Medicine; 379(3): 207-9.

Heilmann, Sebastian (2016). ″Leninism Upgraded: Xi Jinping′s Authoritarian Innovations″, China Economic Quarterly, vol 20, no. 4, Gavekal Dragonomics, pp. 15‐22.

Jensen, C. B. and A. Blok (2013). "Techno-animism in Japan: Shinto Cosmograms, Actor-network Theory, and the Enabling Powers of Non-human Agencies." Theory, Culture and Society 30(2): 84-115.

Katz, Michael L. and Carl Shapiro (1985), The American Economic Review
Vol. 75, No. 3, pp. 424-440

Lessig, L. (2006). Code 2.0. Basic Books.

Matsuo, Yo eds. (松尾陽編) (2017). Architecture and Law:「 アーキテクチャと法 法学のアーキテクチュアルな転回? 」Tokyo 東京, Kobundo 弘文堂. (in Japanese)

Merchant, Gary (2019), “‘Soft Law’ Governance of Artificial Intelligence,” AI Pulse,
https://aipulse.org/soft-law-governance-of-artificial-intelligence/?pdf=132, Last accessed Oct. 19,2019.

Ministry of Economy, Trade and Industry of Japan(2019) G20 Ministerial Statement on Trade and Digital Economy. Japan.
https://www.meti.go.jp/press/2019/06/20190610010/20190610010-1.pdf.(accessed October 18,2019).

Ministry of Foreign Affairs (2019). G20 Osaka Leaders’ Declaration,
https://www.g20.org/pdf/documents/en/FINAL_G20_Osaka_Leaders_Declaration.pdf

(accessed June 30, 2019).

Schilling, Melissa A. (2017). Technology Success and Failure in Winner-Take-All Markets: The Impact of Learning Orientation, Timing, and Network Externalities. Academy of Management Journal VOL. 45, NO. 2.

Sunstein, C. R. (2014). "Nudges vs. Shoves." Harvard Law Review 127(6): 210-217.

Waldman, Ari Ezra(2018). Privacy as Trust: Information Privacy for an Information Age, Cambridge University Press.

Wiener, N. (1950). The human use of human beings : cybernetics and society. Boston, Houghton Mifflin.

World Health Organization (2019). WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening.