THOMSON REUTERS
Westlaw®

Trust me!: How to use trust-by-design to build resilient tech in times of crisis

By Gabrielle Paris Gagnon, Esq., and Vanessa Henri, Esq., *Fasken*, and Abhishek Gupta, *Montreal AI Ethics Institute*

JULY 19, 2020

Nations across the world have started to deploy their own contact- and proximity tracing apps that claim to be able to balance the privacy and security of users' data while helping to combat the spread of COVID-19, but do users trust them?

The efficacy of such applications depends, among other things, on high adoption and consistent use rates, but this will be made difficult if users do not trust the tracing apps. Trust is a defining factor in the adoption of emerging technologies, and tracing apps are not an exception.

In this article, we argue that trust-based design is critical to the development of technologies and use of data during crisis such as the COVID-19 pandemic. Trust helps to maintain social cohesion by hindering misinformation and allowing for a collective response.

> While the right to privacy implies the confidentiality, integrity, and availability of the personal information processed by technologies, the right to be protected from discrimination requires big data to be subject to algorithms that are fair.

In this article, we discuss how systems built using trust-by-design are more resilient, how such technologies can support an effective response to the COVID-19 pandemic, and by providing best practices that software developers should implement when designing technologies.

In this article, we use the definition of trust provided by Dr. Ari Ezra Waldman in *Privacy as Trust - Information Privacy for an Information Age*, that:

Trust is a resource of social capital between or among two or more parties concerning the expectation that others will behave according to accepted norms. It mitigates the vulnerability and power imbalance inherent in disclosure, allowing sharing to occur in the first place.[1]

Trust therefore presupposes adherence to social and legal norms, which would include respect of individual rights, such as the right to privacy and right against discrimination.

While the right to privacy implies the confidentiality, integrity, and availability of the personal information processed by technologies, the right to be protected from discrimination requires big data to be subject to algorithms that are fair.

But, as the last decade has been marked with numerous data breach and misuses, public trust in technological solutions is not naturally fostered. It must be an integral part of the design process.

We believe value-based design is a critical part of fostering trust which deserves more attention from both a legal and technological standpoint.

Despite the importance of developing technologies where the user will trust that her data is being processed responsibly and ethically, user trust has traditionally been relegated to a secondary consideration in technology design.

In fact, legislations seldom regulate design, focusing excessively on consent and control, and thereby pushing accountability to users.

While the European *General Data Protection Regulation* includes a provision for privacy-by-design, other legislations continue to use the Fair Information Principles, which date from the 1970s and were focused on electronic databases as opposed to user-facing applications.

As for security-by-design, legislations such as the *California IoT Security Law* are a good step in regulating information security in consumer applications.

Nonetheless, there remain a few obligations to consider elements of trust at the design stage, and none to consider the notion of trust as a critical foundation of the decision-making process for technology development.

Whereas regulators are focused on collection, use, and disclosure of data, the truth is that these data actions fail to consider the overall design impact on consumers. We must integrate trust-based

Thomson Reuters is a commercial publisher of content that is general and educational in nature, may not reflect all recent legal developments and may not apply to the specific facts and circumstances of individual transactions and cases. Users should consult with qualified legal counsel before acting on any information published by Thomson Reuters online or in print. Thomson Reuters, its affiliates and their editorial staff are not a law firm, do not represent or advise clients in any matter and are not bound by the professional responsibilities and duties of a legal practitioner. Nothing in this publication should be construed as legal advice or creating an attorney-client relationship. The views expressed in this publication by any contributor are not necessarily those of the publisher.

principles in the design of the applications from the outset. Design is a powerful process which "allocates power and is inherently political."[2]

Design decisions should be made early on by the developers to "anticipate human limitations and predictable errors."[3] Failing to do so can ultimately erode trust towards the application. Developers must therefore design their applications with the explicit aim of fostering user trust.

Public backlash against COVID-19 case tracking can undermine public health initiatives aimed at limiting the spread of the virus.

> We believe value-based design is a critical part of fostering trust which deserves more attention from both a legal and technological standpoint.

Leaving trust out of the design of such policy can rapidly lead to distrust. In Israel, the emergency regulations initially authorized the Israel Security Agency to collect technological data from the telecommunications companies, such as GPS location, upon the request of the Ministry of Health following the identification of a person as a carrier of the SARS-CoV-2 virus.

The whereabouts of this person could then be traced for 14 days using GPS, and all individuals in contact with the "infected" person could then be notified by SMS.[4] The use of GPS tracking meant that individuals exact locations could be identified.

The data was also shared with local authorities to ensure quarantine orders. This program led to widespread criticism because of fears about widespread surveillance and eventually ended by the Israeli Supreme Court on April 26, 2020, which concluded that the tracing application violated the privacy of users.

This illustrates that effective tracing policy must not only be built on widespread trust within the community but must also comply with existing legal privacy protections.[5]

The inclusion of trust concerns within the design of governmental policies for COVID-19 case tracking helps to develop trust among users and thereby facilitates the effective tracking of the virus. In contrast to the Israeli example, Germany took a very different approach to crafting its case tracking application.

The application emerged as part of a public hackathon and was therefore created by citizens as opposed to secretive security agencies as in Israel. The application that was developed from this initiative was designed to comply with existing European privacy laws, the GDPR.

The application also performs tracking based on Bluetooth technology that, unlike with the Israeli Security Agency's GPS tracking, only identifies which individuals have been meters apart.

Trust concerns may conflict with other imperatives, such as effective contact tracing, and must be balanced against these other considerations. Despite the efforts made both at the policy and technological level, the German example has not been without criticism.

A group of 300 scientists across the globe criticized[6] the centralized architecture of the application, where anonymized data are automatically uploaded and stored on remote servers.

They argued that a centralized protocol is prone to abuse and can be repurposed for mass surveillance, advocating instead for a decentralized protocol where the data remains on the users' phones and is not automatically shared with the authorities.

Germany, which initially backed a centralized architecture, publicly announced that the country would adopt a decentralized approach to digital contact tracing solutions. Clearly in this case there is a trade-off that must be balanced between the efficacy of the application and the desire to protect user privacy.

The different paths and opinions raised regarding tracing apps at a policy and technological level worldwide show how much design is critical to building systems that can foster adoption of technologies, and resilience in a time of crisis.[7]

At the same time, it points out the importance of transparency and public discussion around technologies, as it helps to improve the design.

If such a discussion were to occur around more technologies, we believe it would increase the overall resilience of technologies through better cybersecurity. It would also lead to more inclusive and environment-friendly technologies.

Having seen that user trust can be fostered at the policy level, it is important to consider also what practical decisions software developers, designers, and organizations can take to design applications that promote this aim.

Trust can be fostered at the design level with a set of best practices. Briefly, these best practices involve data governance structures, encoded privacy preservation mechanisms and software practices enabling reliability and predictably.

Trust can be built through governance structures. Data trusts[8] provide such a mechanism. They act as an independent third-party that allows access to data based on established privacy and other principles, releasing only necessary pieces of personal data to the developers of different tracing apps.

On the user side, they engender trust by collecting limited amounts of information from the users and uphold their commitments to their users through the fiduciary mechanism where they are legally bound by a fiduciary duty to act in the interest of the users who trust them with their data.

Part of the reason why the public trust towards data-driven solutions eroded though the years is that once data has been misused without people's knowledge or consent, it's hard to hold the wrongdoer accountable.

The data trust mechanism aims to solve that problem by making the trustees accountable towards the beneficiaries.

Beyond governance, trust also requires privacy preservation and information security practices to be deeply integrated. Recent work has highlighted the need to move from theory to practice, especially as it relates to building trustworthy AI systems. An approach that utilizes verifiable claims[9] presents a promising avenue.

It is an opportunity to evoke higher levels of trust from users since it allows to audit the claims that are made by the designers and developers of the system regarding the trustworthiness of it.

Effective tracing policy must not only be built on widespread trust within the community but must also comply with existing legal privacy protections.

Also, using techniques like differential privacy (DP),[10] which adds statistical noise to the data to protect individual data point's privacy while still providing useful information for downstream analysis engender trust from users because of protection against data disclosures, especially given that it provides mathematical guarantees, analogous to the concepts of formal verification techniques in software design.[11]

Technological restrictions can also be implemented in order to limit the use that the system can make with users' data.

APIs can be designed to improve user trust by controlling the inputs and outputs to the application. Techniques like use-based privacy,[12] concretely implemented in a framework like Ancile,[13] encode purpose-limitation and privacy requirements into the software so that misuse of data is minimized.

Therefore, developers who use this API will be restricted in what they can do with user data. For users, knowing that privacy preservation mechanisms are enshrined immutably in the API the developers use is likely to decrease their fear that their data may be misused.

Managing the development of the system using software practices enabling reliability and predictably, like MLOps[14] serves to enhance trust towards them. It ensures reproducibility and benchmarking for comparison of the performance and operations of the system in a standardized manner.

For example, reliability of the system can be encoded through the use of high-availability (HA),[15] fault tolerance,[16] graceful failures,[17] and other techniques, and enforced through the use of Service Level Agreements (SLAs) from a legal perspective. In the context of inherently probabilistic systems like machine learning, providing guarantees and setting behavior boundaries to protect against unexpected operations under out-of-distribution data that the system has not encountered before[18] allows for higher levels of trust in the system.

Akin to the use of privacy impact assessments (PIAs), we advocate for the use of trust impact assessments (TIAs) which encompass PIAs and other techniques as mentioned above to provide an indication to users around the behavioral expectations from the system.

Specifically, building on pieces of work like datasheets for datasets,[19] model cards for model reporting,[20] nutrition labels for datasets,[21] and FactSheets[22] as a way of increasing trust in suppliers' declarations of conformity can enable designers and developers of software systems to demonstrate trustworthiness of their systems.

When it comes to tracing solutions, applying some of these mechanisms will evoke higher levels of trust from the public as they will have a higher degree of visibility on how the solution has been designed, what its capabilities are, what its limitations are, how their data will be used, etc.

Structures which have been built with trust-by-design principles are more likely to be resilient when stress-tested by a crisis, thus more efficient in solving the problems they were designed for. For instance, our legal system has been designed with trust as a cornerstone which allowed it to thrive and evolve overtime.

Concepts such as treason law, fiduciary duty, trustees and the protection of the right to a lawyer in a criminal process to balance unequal powers are all examples of how legal systems have been built with trust-by-design.

Software developers using value-based design and trust as a decision-making framework are more likely to develop technologies that will be adopted widely, serve us and foster social cohesiveness when it is the most needed.

Several notions emerge as intrinsic to trust, such as integrity, predictability and reliability, and the respect of individual rights.

These notions are guiding principles that should steer our development of technologies to ensure they support human development, and not hinder it. It emerges as the foundational principle which will make people inclined to trust the institutions and their solutions.

Notes

[1] Ari Ezra Waldman, Privacy as Trust- Information Privacy for an Information Age, Cambridge University Press, 2018, page 4

[2] Woodrom Hartzog, Privacy's Blueprint — The Battle of Control the Design of New Technologies, Harvard University Press, 2018, p. 51.

[3] Woodrom Hartzog, Privacy's Blueprint — The Battle of Control the Design of New Technologies, Harvard University Press, 2018, p. 50.

[4] https://tcrn.ch/2NMW05a

[5] https://bit.ly/3iu2pk6

[6] https://bit.ly/2NNa6DU

[7] https://bit.ly/2CPUnl1

[8] https://bit.ly/2ZpZuQL

[9] Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., ... & Maharaj, T. (2020). Toward trustworthy AI development: mechanisms for supporting verifiable claims. arXiv preprint arXiv:2004.07213.

[10] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3-4), 211-407.

[11] Kishi, T., & Noda, N. (2006). Formal verification and software product lines. Communications of the ACM, 49(12), 73-77.

[12] Dagher, G. G., Mohler, J., Milojkovic, M., & Marella, P. B. (2018). Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. Sustainable cities and society, 39, 283-297.

[13] https://bit.ly/2CZIpWb

[14] Fursin, G., Guillou, H., & Essayan, N. (2020). CodeReef: an open platform for portable MLOps, reusable automation actions and reproducible benchmarking. arXiv preprint arXiv:2001.07935.

[15] Addo, I. D., Ahamed, S. I., & Chu, W. C. (2014, June). A reference architecture for high-availability automatic failover between PaaS cloud providers. In 2014 International Conference on Trustworthy Systems and their Applications (pp. 14-21). IEEE.

[16] Guerraoui, R., & Schiper, A. (1996, June). Fault-tolerance by replication in distributed systems. In International conference on reliable software technologies (pp. 38-57). Springer, Berlin, Heidelberg.

[17] Shelton, C., & Koopman, P. (2003). Using architectural properties to model and measure graceful degradation. In Architecting dependable systems (pp. 267-289). Springer, Berlin, Heidelberg.

[18] https://bit.ly/2YN6Xdh

[19] Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2018). Datasheets for datasets. arXiv preprint arXiv:1803.09010.

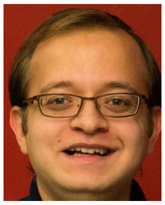
[20] Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019, January). Model cards for model reporting. In Proceedings of the conference on fairness, accountability, and transparency (pp. 220-229).

[21] Holland, S., Hosny, A., Newman, S., Joseph, J., & Chmielinski, K. (2018). The dataset nutrition label: A framework to drive higher data quality standards. arXiv preprint arXiv:1805.03677.

[22] Arnold, M., Bellamy, R. K., Hind, M., Houde, S., Mehta, S., Mojsilović, A., ... & Reimer, D. (2019). FactSheets: Increasing trust in AI services through supplier's declarations of conformity. IBM Journal of Research and Development, 63(4/5), 6-1.

This article appeared on the Westlaw Practitioner Insights Commentaries webpage on July 19, 2020.

ABOUT THE AUTHORS



Gabrielle Paris Gagnon (L) is an emerging tech lawyer at **Fasken** specializing in venture capital financing. She advises and represents tech companies in all phases of their growth, from the idea phase to well-established tech companies. She can be reached at gparis@fasken.com. **Vanessa Henri** (C) is an emerging tech, data governance and cybersecurity lawyer at Fasken. She is a certified data protection officer and an ISO 27701 lead implementor with experience in implementing data protection management system and negotiating tech agreements for emerging growth companies. She can be reached at vhenri@fasken.com. **Abhishek Gupta** (R) is the founder of **Montreal AI Ethics Institute** and a machine learning engineer at Microsoft where he serves on the CSE Responsible AI Board. His research focuses on applied technical and policy methods to address ethical, safety and inclusivity concerns in using AI in different domains. He can be reached at abhishek@montrealethics.ai. This article reflects the situation at the time it was written based on the rapidly changing nature of the COVID-19 pandemic.

Thomson Reuters develops and delivers intelligent information and solutions for professionals, connecting and empowering global markets. We enable professionals to make the decisions that matter most, all powered by the world's most trusted news organization.