The Virtru Trusted Data Platform:
Delivering Privacy-Preserving Contact Tracing and Analytics

Privacy and Contact Tracing

Contact tracing is a well-understood and important weapon in fighting pandemics. Unfortunately, both legacy and emerging contact tracing efforts have compromised either on data privacy or response effectiveness. The public and policy makers have been presented with two general options: centralized data collection systems with weak privacy assurances or fully decentralized approaches with stronger privacy protections but limited utility.

This is a false choice.

In order to beat COVID-19 decisively and prepare for future fights, privacy and control must enhance sharing with "centralized" organizations, such as healthcare and disaster response organizations. This virtuous circle is possible through **verifiable trust**: proven assurance that our data is only being used for its intended purpose, for a specified period of time. In so doing, individuals will be empowered to share data with centralized organizations in a way that both protects our civil liberties and increases access to vital data.

The good news is that this approach is not a fantastic notion. Verifiable trust is possible with technology available today. Individuals can be granted verifiable control over how their data is used and for how long, and authorized organizations can gain the ability to leverage deeply sensitive information in a controlled and audited fashion. As a result, organizations can respond more quickly armed with deeper insights, while ensuring transparency and individual control over exactly how personal data is used over time.

Current contract tracing approaches fall into two categories:

- **Centralized Data Collection** (strong insight; weak privacy) - The first category relies on "trusted central authorities" that have nearly unfettered access to private data. Users are asked to trust that authorities only use the data for intended purposes and will not intentionally or unintentionally misuse or leak this data. This is the approach taken by such countries as South Korea, Taiwan and China. These approaches have scale and speed, enable correlation of multiple data sources, and often require a high degree of compulsion.

- **Decentralized "Proximity-based" Tracing** (limited insight; stronger privacy) - The second category uses proven cryptographic privacy mechanisms that allow specific questions to be answered in a secure and private manner. Many approaches have focused on leveraging Bluetooth capabilities - for example "contact chirping" - to notify individuals of "proximity" to others who have COVID-19 (e.g., MIT Private Kit; Google/Apple Collaboration). Insights are typically available only to individuals and do not allow health

©2020 Virtru Corporation. All rights reserved.

organizations to answer questions on a broad scale. In addition, there is no opportunity to develop analytic insights that might enhance the effectiveness of these efforts (e.g., addressing false positives) . While these approaches offer strong privacy assurances, they can deeply constrain how data is used to save lives and actually delay the reopening of our economy.

A New Approach: The Virtru Trusted Data Platform

The Virtru Trusted Data Platform (TDP) offers a new approach that starts with the principle that the individual must always be at the center of control. Given the string of intentional and unintentional abuses and misuse of data over the past two decades, individuals are conditioned not to "opt-in" to systems that require trusting third parties to "do the right thing" with our most sensitive data. Compulsory approaches, on the other hand, raise deeply serious civil liberty concerns.

As evidenced by Virtru's commercial adoption the past seven years, giving individuals control over their data **increases sharing**. With confidence that their data is only being used for intended purposes, trust is created.

Effective and comprehensive contact tracing requires leveraging highly sensitive, personal data at scale - including such information as where individuals travel, with whom they may have been in contact, and potentially combining this data with other data sources. Giving individuals verifiable control of their data in a way that also empowers authorized institutions to make use of this data (for intended purposes, for specified periods of time) is critical to the public trust required for contact tracing and surveillance to work at scale.

The Trusted Data Platform is Virtru's secure, open platform, available to anyone, that allows users to seamlessly protect their data. The platform tightly controls how user data can be accessed and consumed, down to exactly what questions can be asked using this data, and by whom and which specific analytic systems. This platform utilizes open technologies across authentication, encryption, data encodings, containerization, and more. And, at the center of it all is the data itself, protected by the open Trusted Data Format (TDF).

To unlock critical analytic-driven workflows, the Virtru TDP provides a transparent and fully auditable means by which vetted analytics with known behaviors can be given access to data through encryption key grants. If users wish to revoke access, they can turn off access to the encryption keys at any time, rendering the content totally inaccessible unless the user decides later to re-authorize it.

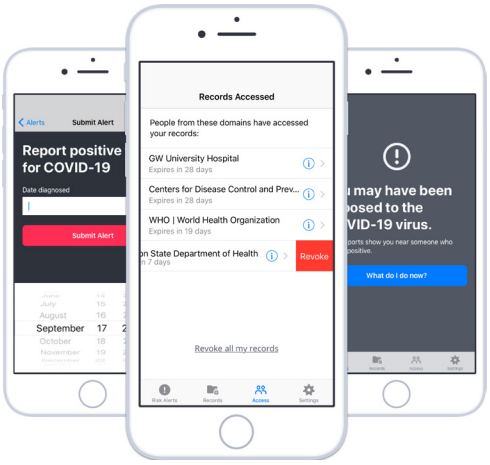
As analytics are approved for use by data owners, value can be unlocked immediately and increasingly over time. Two simple analytic examples available today highlight how both individuals and institutions can be empowered:

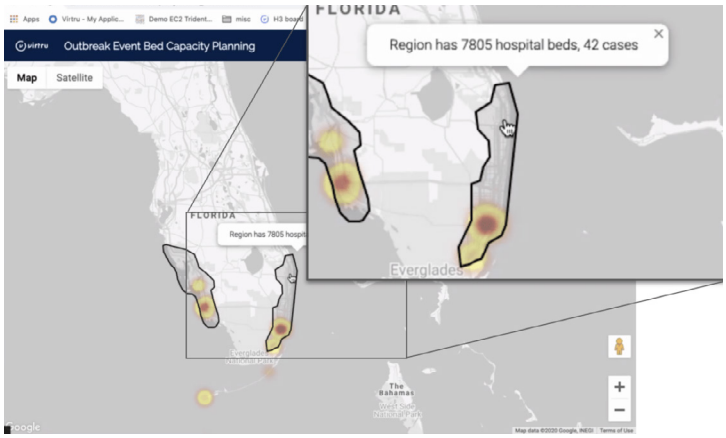
- **Empowering the Individual** - "Proximity / Possible Contact Alerts" - This is the core contact tracing use case, where geolocation and other cell phone data like Bluetooth proximity may be combined and correlated to automatically alert users about possible contact with a COVID-19 positive patient.

- **Empowering Institutions** - "Outbreak Heatmap" - A simple example of enabling institutions while ensuring privacy is demonstrated by granting geolocation permission to an analytic that produces only anonymous heatmaps of active or suspected cases. Because the output is aggregated into case ranges, no personal information is at risk when sharing the map and its reflected data.

©2020 Virtru Corporation. All rights reserved.

Virtru's Contact Tracing Reference App, and the analytic output viewer for disaster response. Note that no PHI is revealed under this particular authorized use.





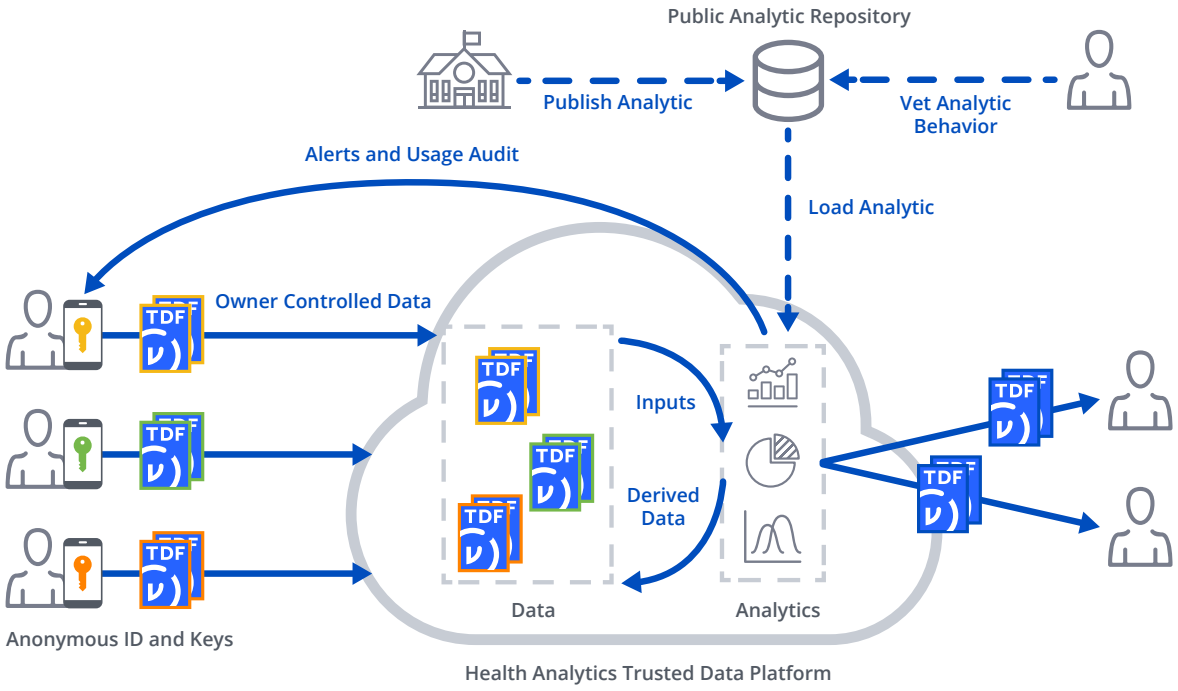
The Virtru Trusted Data Platform Implementation for Contact Tracing

The Virtru TDP stack enables data owners to control usage and access to their data throughout the entire data lifecycle, not only for a specific use case or stage of analysis. This control can start at the moment of data creation. To demonstrate how the TDP can be leveraged to implement a secure contact tracing capability, Virtru has built a full lifecycle reference implementation.

After allowing for geolocations to be collected, the initial operating capability enables proximity analytics to detect various COVID-19 patient contact algorithms and automatically alert users, all while ensuring the privacy of the users involved. It also demonstrates privacy-preserving analytic authorization from third parties, such as anonymized, aggregated geographic case data.

©2020 Virtru Corporation. All rights reserved.

Trusted Data Platform Overview
Data Owners control and audit how their data is used through its lifecycle.



The Mobile App - Data Collection & Protection

For individuals participating in a Contact Tracing effort, data collection and protection start with their existing mobile devices. The Virtru TDP app can immediately protect data as it is collected. For contact tracing uses, such data can include:

- Survey / Interview Data (e.g., symptoms, risk factors)
- Current Location Tracking (e.g., GPS)
- Historic Location Tracking (e.g., Google Location History API)
- Bluetooth Proximity Logs
- Other Structured or Unstructured Medical Data

False Choices and Losing Control

"Donation of data" as described today can be misleading. An individual can authorize his or her data for use for particular purposes. Without some control, "donated data" can be weaponized or otherwise used against that person in ways that can be hard to predict. If the individual wishes to truly make his or her data public, that choice should be made without having to risk subjecting it to misuse by medical researchers.

©2020 Virtru Corporation. All rights reserved.

Trusted Data Format

In order to ensure proper protection, either immediately or prior to transmission, the Virtru TDP app can encrypt data using the Trusted Data Format. TDF goes beyond just encrypting the data. It allows explicit policy requirements to be bound to the data, and digital signatures that can be used to ensure the authenticity and integrity of data.

Access Policy

To establish a TDF access policy, the Virtru TDP can ask the user what permissions they want to grant per data To establish a TDF access policy, the Virtru TDP can ask the user what permissions they want to grant per data type (or data attribute). Virtru's app, for example, asks users for how long they wish their geolocation track data to be accessible, using a range from 28-days to indefinitely. The user can always change these permissions, and they can be updated even for data already transmitted to third parties. The underlying control system follows an Attribute-Based Access Control (ABAC) model. While the model is deeply powerful, users can still enjoy a straightforward, intuitive interface for control and audit.

Identity

To ensure that only the rightful owner is in control of their data, a strong cryptographic identity is generated on-device and stays within the application. Traditionally, this would involve authentication with an authoritative Identity Provider (IDP), which Virtru facilitates by default in other applications. For this use case, however, Virtru has implemented anonymous identity keys much like a cryptocurrency identity. Referred to as a pseudonymous identity, it allows users to avoid using real names or email addresses, but still prove ownership and authority.

Sharing / Transmission

Before transmitting data off-device into a shared database / data lake, the app automatically TDF-protects all data using keys the user controls, enforcing the policies previously set by the user.

TDF Key Management

TDF-protected data is actively managed by a key server, or Key Access Service (KAS). This key server is responsible for enforcing policies set by data owners. Anyone can run such a service, allowing keys to be held in any location, even as the underlying data may converge into a shared analytic environment.

A Host-Anywhere, Zero Knowledge Data Lake

Because all data is encrypted before it arrives at a shared environment, it is useless until a decryption key is provided. Administrators of the environment have technical access to all data deposited there. However, they must request the keys to each owner's data, and these requests must be granted before the data may be used for any purpose. This form of encryption uniquely affords persistent control, unlike most other forms of encryption (e.g., cloud provider at-rest solutions). Since data carries its protections with it, such a data lake can be hosted in an untrusted environment and data owners can be confident their information is safe.

Data Sharing Constraints vs Empowerment

The PACT system relies on only vetted HC professionals submitting data. This is a very fragile approach at scale, and constrains potential value. TDF provenance allows ANYONE to submit data. Analysts can pick and choose what data they want to trust for what purposes. It allows for cryptographically asserted provenance via digital signatures and bound assertions.

©2020 Virtru Corporation. All rights reserved.

Analytics

Only analytics with a trusted identity may be granted access to keys. Our reference implementation has all analytics published to a public repository for maximum transparency, as any party may evaluate them. This includes the proximity detection analytic that produces derived "contact alert" data to alert users (using pseudonymous identity keys to protect said alert messages).

Analytics execute in a Trusted Execution Environment (TEE). TEE's are responsible for many things, including measuring the integrity of analytics and remotely attesting their identity to key servers. They are also responsible for ensuring the analytics are appropriately protected from external tamper or introspection. TDP-enabled TEE's also have strong cryptographic identity. With this identity, analytics may request access to keys from a key server, and data owners can know exactly what analytics asked for their data, and when.

Audit

All use of data is audited via key server, and is made available to data owners to examine. Analysis of usage behavior may inform future changes in policy, which can be leveraged to reduce or even fully revoke all access.

Closing the Loop - Alerts

All "Contact Alert" messages are TDF encrypted with pseudonymous keys provided as part of the original data submissions. By using these anonymous IDs and TDF encryption, mobile users can receive updates without compromising the confidentiality of the underlying content or identity.

Enabling Future Use Cases

By establishing a data lake of highly valuable data under owner control, future research can be accelerated. The Virtru TDP app already contains the mechanisms to handle requests for access to data for new purposes, and granting these requests becomes much less of an issue when data owners are confident in the protections afforded their data. Thus, the Virtru app already incorporates the functionality needed to permit individuals' sensitive data to be used for a number of purposes, including those related to the global effort to combat COVID-19.

If you'd like to learn more about using Virtru's Trusted Data Platform to integrate healthcare analytics into your infrastructure, please contact us at platform@virtru.com.

About Virtru

At Virtru, we empower organizations to unlock the power of data while maintaining control, everywhere it's stored and shared. Our portfolio of solutions and tools—built on our open data protection platform—governs data throughout its lifecycle. More than 5,000 organizations trust Virtru for data security and privacy protection. For more information, visit virtru.com or follow us on Twitter at @virtruprivacy.

©2020 Virtru Corporation. All rights reserved.