# Designing Operable Privacy Standards for Data Sharing during Public Health Emergencies

## Workshop Proceedings

Prepared by **Third Plateau Social Impact Strategies**
for workshops held October 13 & 14, 2022.

COVID-19 Illustration: CDC


CCST
CALIFORNIA COUNCIL ON SCIENCE & TECHNOLOGY

CCST
CALIFORNIA COUNCIL ON SCIENCE & TECHNOLOGY

**PARTNER INSTITUTIONS**

California State University
California Community Colleges
California Institute of Technology
Stanford University
University of California
University of Southern California
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
NASA Ames Research Center
NASA Jet Propulsion Laboratory
Sandia National Laboratories
SLAC National Accelerator Laboratory

# Designing Operable Privacy Standards for Data Sharing during Public Health Emergencies

## Workshop Proceedings

Workshop held October 13 & 14, 2022

Prepared by Third Plateau Social Impact Strategies

Third Plateau

## Table of Contents

## Executive Summary

On October 13th and 14th, the California Council on Science and Technology (CCST) welcomed 26 public health and data privacy experts to a virtual workshop on *Designing Operable Privacy Standards for Data Sharing during Public Health Emergencies*. Participants were asked to reflect on the challenges and risks associated with collecting, sharing, and/or accessing health data and to conceptualize possible policy solutions to known barriers.

Participants reflected on four use cases that were developed by CCST's COVID-19 Steering Committee in consultation with relevant stakeholders. These four use cases exemplified scenarios where privacy protections may inadvertently limit the capacity for effective public health interventions. These use cases served as a heuristic to explore common barriers and challenges. Six themes emerged from these discussions:

1. Data collection practices are inefficient and inconsistent.
2. Data collection and reporting technology and infrastructure are outdated.
3. Data systems were not designed with public health in mind.
4. Data policies are inconsistently interpreted and applied.
5. Privacy remains a concern.
6. Community voice should be centered in solutions.

Participants deliberated over potential solutions to these barriers, and proposed the following:

1. Bring impacted communities to the table.
2. Extend privacy legislation to all health-related data collection, including entities not subject to HIPAA.
3. Invest in workforce development to build fieldwide capacity to collect, manage, and utilize data.
4. Modernize privacy laws to account for current realities.
5. Standardize systems and increase interoperability.
6. Build a universally accessible and interoperable data sharing system.

* More details on the proposed solutions are included within (pages 14 – 17).

These workshop proceedings serve as an archive of the discussions. As such, the recommendations, summaries, and statements within have been validated by workshop participants, including CCST's COVID-19 Steering Committee, but have otherwise not been formally peer-reviewed, nor expanded upon. Policy recommendations derived from this workshop will be further developed with the COVID-19 Steering Committee and released later this year. These forthcoming recommendations will be peer-reviewed and informed by engagement with a broader selection of relevant stakeholders.

## Workshop Context

As part of its Disaster Resilience Initiative in 2021, the California Council on Science and Technology (CCST) convened experts around the question of how best to prepare California for the next public health crisis. In June 2022, CCST hosted a two-day convening that brought together 35 leading public health experts to envision a more holistic and inclusive public health system that would be better equipped to respond to emerging public health threats. One major theme that emerged during these discussions were the challenges to sharing critical data and information within and between the public health system, medical providers, and researchers. This theme inspired this subsequent workshop held in October 2022.

## Workshop Objectives

On October 13th and 14th, CCST invited 26 public health and data privacy experts to a virtual workshop to explore solutions to ensure health data are reliable, secure, and informative at all levels of the public health system. See Appendix A for a list of the institutional affiliations of workshop participants. The workshop asked participants to consider:

- *How should privacy standards be adapted during public health emergencies to better facilitate information sharing?*
- *How do we ensure that key stakeholders readily have access to the data necessary for decision-making without compromising the privacy of individuals?*
- *How can we alleviate confusion about what data sharing is permissible?*

The overarching goal for the workshop was to generate actionable recommendations that CCST will deliver to policymakers in California. To achieve this goal, the workshop had three objectives:

1. Explore use cases to identify key data sharing opportunities, challenges, barriers, and risks;
2. Develop guidelines for which data can be shared, with whom, and under what circumstances; and
3. Define (or make recommendations regarding) a level of re-identification risk that is acceptable under various circumstances.

## Workshop Design

The two-day workshop began with opening remarks from Amber Mace, CEO of CCST and Michael Kleeman, Chair of CCST's COVID-19 Steering Committee.

For much of the workshop, participants engaged in small-group discussions facilitated by Third Plateau, a social impact strategy firm. The participants were organized into two breakout groups, pre-sorted to ensure diverse perspectives and backgrounds would be

present in each group. In their groups, participants explored several use cases to discuss health data sharing challenges, barriers, opportunities, and risks. After discussing additional barriers to data sharing, they completed a prioritization exercise to identify the top three to four barriers that, if addressed, would have the greatest positive impact on information sharing and collaboration. Participants then completed a rapid brainstorming activity to generate ideas for how to overcome those barriers. Participants then collectively identified what they deemed to be the most impactful solutions.

Day two opened with a panel on *Actualizing Recommendations for Public Health*. The panelists discussed strategies for structuring and contextualizing recommendations to build political and administrative buy-in. After the panel, participants returned to their breakout groups to review their priority solutions and discuss feasibility (i.e., initial expectations related to cost, political will, and time). Participants then selected solutions to be further elaborated in smaller workgroups. A total of six small workgroups developed recommendations in support of 6 priority solutions. Finally, each workgroup briefly presented their recommendation to the full group for feedback and questions.

## Workshop Outcomes and Key Takeaways

### Use Cases: Exploring Challenges, Barriers, Opportunities, and Risks

Prior to the workshop and with input from relevant stakeholders, CCST developed four use cases to discuss during the workshop (see Appendix B). The diverse scenarios provided concrete examples for which the groups would consider current and potential challenges, barriers, opportunities, and risks of data sharing. Each breakout group sequentially reviewed and discussed two use cases. Discussion questions included:

Within the context of this use case…

- Where does data flow freely? What current policies allow for the transfer of data in this situation?
- Where is the transfer of data stalled? What are the barriers?
- What current policies do not allow for the transfer of data in this situation? What policies or laws address data sharing in this situation, but are often misunderstood or misinterpreted?
- What would have to happen to allow for data sharing? How might we adjust or amend current policies, to allow for data sharing in this situation?
- What are the risks of increased data sharing? What level of risk is acceptable?

Several themes emerged through the course of these conversations, including:

- **Data collection practices are inefficient and inconsistent.** At the local level, data need to be collected consistently. There's a need to invest in data systems, personnel, and training to support interoperability for real-time, actionable information.
- **Data collection and reporting technology and infrastructure are outdated.** Critical timelines are often delayed due to outdated technology and infrastructure, such as the common practice of sharing information via fax machines. More advanced technology and better designed data collection systems would allow for more seamless and timely data sharing.
- **Data systems were not designed with public health in mind.** Information systems were designed for billing and patient care, rather than to support public health research and response.
- **Data policies are inconsistently interpreted and applied.** Privacy principles are not universally applied or understood by all stakeholders involved in data collection, reporting, and sharing.
- **Privacy remains a concern.** As technology advances, risks of reidentification increase. Caution should be practiced when establishing new data sharing policies and standards. Data practitioners should not collect or share unnecessary data but should consistently consider the minimum data needed to serve the intended purpose.
- **Community voice should be centered in solutions.** An equity lens should be integrated into design solutions. Community-based organizations and community members should be engaged in the development of data policies, practices, and infrastructure to help make decisions about how the data from their communities is being collected and used.

### Key Barriers to Data Sharing

Building on the barriers identified during discussion of the use cases, participants reflected on their broader knowledge and experience to identify additional barriers to effective, efficient, and secure data sharing. A complete list of barriers identified across these discussions can be found in Appendix C. Once these barriers were enumerated, participants were asked to vote on the most significant barriers to health data sharing. Participants identified the following five barriers as the most significant:

- Insufficient public health data system infrastructure, including a lack of standardization and interoperability across data collection, storage, and sharing;
- Inadequate capacity for data quality, management, and analysis;
- Variability in interpretation and compliance with privacy laws and guidelines;
- Consumer protection concerns; and
- Lack of clarity on which data are shared and for what purpose.

### Potential Solutions to Top Barriers

Participants identified many potential solutions to each barrier. After reviewing all solutions, they voted on those that would best support effective, efficient, and secure

data sharing. Solutions for each barrier are grouped into themes. Those deemed most impactful (with at least three votes) are indicated below with an asterisk (*) and colored in green.

Barrier: **Insufficient public health data system infrastructure, including a lack of standardization and interoperability across data collection, storage, and sharing**
Potential Solutions:

- Design a practical, efficient, and secure interoperable data system
    - *Build a unified and interoperable infrastructure across community locations with a common technology backbone, with security/privacy features that comply with federal and state regulations.
    - *Design the next generation data system, driven by user-centered design and spearheaded by new partners and players (as opposed to incumbent stakeholders).
    - *Integrate electronic medical record (EMR) Application Programming Interfaces (API).
    - Design for the smallest players and leverage hyper-secure web-based technologies to protect them.
    - Define basic functions that support interoperability (e.g., standardized export/import formats) and require all data systems to include these components.
    - Incorporate deidentification functionality, information about legal constraints on data sharing, and data sharing mechanisms into electronic health record (EHR) software.
    - Take a more comprehensive approach that would be more adaptable towards future situations, rather than reactive to specific existing public health needs or crises.
- Strengthen operational processes
    - Create dedicated Chief Information Officer roles.
    - Streamline communication and chains-of-command within existing infrastructure.
    - Regulate EHR providers to drive more responsible, functional, and economic behavior.
- Gather community and stakeholder input
    - *Establish a public health "congress" of experts who can define a universal structure for EHR fields/coding that is useful and practical for public health use.
    - *With community input and education, standardize a limited data set of reportable EHR elements to be shared systematically and electronically.
    - In consultation with stakeholders (e.g., point-of-care, laboratories, state and local public health departments), identify and codify key data elements.
- Secure funding
    - *Secure public funding without strings and risks attached to privatization.
    - Secure funding to augment public health data systems and enhance

interoperability with commonly used clinical data systems while integrating other sources of useful data (e.g., geographic information systems) and analytic approaches (e.g., artificial intelligence).
    - Get financing or information technology expertise/software/hardware from private entities whose businesses depend on the aggregation of deidentified data.
    - Fund the qualified health information organizations designated by the Data Exchange Framework (DxF) that serve as data intermediaries of health and human services, community organizations, and health plans.
    - Secure state general fund dollars for all required DxF Data Sharing Agreement safety net signatories, leveraging federal Medicaid dollars as a match.
- Influence and leverage external partners and systems for support, perspectives, and execution
    - Outsource infrastructure to trusted third parties (i.e., health information exchanges (HIEs)).
    - Gather political support for additional funding and clout to address infrastructure and training needs.
    - Encourage innovation and entrepreneurship in the EMR domain and support startups that are working on EMR APIs.
    - Ensure public health is included in use cases for Data Exchange Framework.
    - Leverage/comment on federal standards and regulations that are driving standardization of clinical data among health information technology vendors, EHRs, providers, and HIEs.
    - Break down monopolies in the EMR domain.

Barrier: **Inadequate capacity for data quality, management, and analysis**
Potential Solutions:

- Strengthen education, training, and leadership
    - *Secure sustained and significant investment in leadership and human resources to manage data systems.
    - *Train personnel and leadership in data analysis, data science, data curation, data visualization, and best practices on how to tailor data sharing based on the audience.
    - Invest in health informatics that teach data science and public health.
    - Incorporate data science into educational curriculum, long-term.
    - Build closer alignment with schools of public health and public health practitioners to ensure needed data skills are being taught.
    - Develop and train providers and systems on current and emerging standards for coding to improve data quality.
- Develop simple and clear guidelines
    - *In partnership with experts and communities, develop simple language guidelines for solicitation of personally identifiable information (PII) that clearly

explain why certain data elements are collected and used.

- Secure funding
    - Quantify costs for data system modernization in local, regional, and statewide terms (e.g., assess the cost for a local health jurisdiction to update infrastructure or the cost for the California Department of Public Health to update).
    - Collaboratively (and continuously) advocate for funds
- Strengthen collaborations to improve guidelines
    - Increase collaboration between library science and public health experts.
    - Foster support from the corporate sector (Amazon, Google, Microsoft), without granting indiscriminate access to data.

Barrier: **Variability in the interpretation of privacy laws and guidelines**
Potential solutions:

- Modernize, amend, expand, and clarify laws and policies
    - *Modernize privacy laws to account for current realities (risks, new technology, etc.).
    - *Provide more structured guidance about the minimum known info to be shared that is allowable under a privacy scheme.
    - *Standardize definitions and education on new and existing privacy laws.
    - Modernize definitions and protections in IPA to align more closely with the more stringent protections in Health Insurance Portability and Accountability Act (HIPAA)/Confidentiality of Medical Information Act (CMIA).
    - Expand IPA to apply to local government.
    - Adjust informed consent policies
        - Require opt-in consent from individuals before sharing their information.
        - Create opportunities to ask about willingness to share data through applications.
        - Give people control over when, why, and with whom their data is shared.
    - Apply CMIA to medical data collected by public and private entities regardless of whether that entity is a healthcare provider.
    - Establish clear principles on what needs protecting and what needs to be enabled, to guide policies and practices.
    - Create a hotline to identify privacy breaches.
- Educate and align to increase understanding and consistency
    - *Standardize definitions and education on existing/new privacy laws and policies.
    - *Increase education, training, and onboarding for public health, clinical, and administrative staff working at private and public entities.
    - Align standards and data formats.
    - Educate people on the 21st Century Cures Act, Information Blocking Rules, particularly agencies not accountable to HIPAA and/or the Information Practices Act.

- Regulation and compliance
    - *Increase reporting, monitoring, and oversight to ensure compliance with best practices.
    - *Require entities to follow the most privacy protective law that applies to any of the entities involved (e.g., verily would have to follow IPA/CMIA not the California Consumer Privacy Act when screening people for COVID test appointments).
    - Require public health agencies to identify specific objectives in data requests to facilitate sharing only the data that is necessary to accomplish that goal.
    - Set requirements for community advisory boards within public health space.
    - Establish minimum data security standards for local agencies.
- Implement, support, and improve quality
    - *Build infrastructure, including resources and policy enshrined in statute for essential governmental public health services.
    - Create incentives for quality improvement within organizations regarding data sharing, such as state and private grants and facilitated partnerships with organizations that are experienced in quality improvement.
    - Provide technical support for healthcare organizations sharing data and public health organizations receiving data.
    - Provide specific technical examples for individual use cases of data sharing.
    - Implement "white hat" security testing at institutions that collect, use, share health related data.

Barrier: **Consumer protection concerns**
Potential solutions:

- Community engagement, education, and advocacy
    - *Bring impacted communities to the table in meaningful ways for thinking through potential risks and solutions.
    - Give people control over when, why, and with whom their data is shared.
    - Educate and provide clear, succinct warnings for patients about when their data is shared beyond healthcare.
    - Increase transparency and disclosure about existing de-identification and re-identification practices and problems.
    - Enable disclosures of de-identified data so that the public interest world can try to assess efficacy.
    - Coordinate efforts with other groups working on this issue (e.g., the California Commission on Asian and Pacific Islander American Affairs (CAPIAA) has contacted CDPH regarding AB 1726 (Bonta, 2016) implementation).
- Modernize and adapt laws and policies
    - *Increase legal protections and enforcement mechanisms (e.g., Paperwork Reduction Act).
    - *Modernize privacy laws to account for current realities (e.g., risks, new

technology).
    - *Advocate and enact privacy legislation that extends to all health-related data, not just entities covered under HIPAA. This would include social media, credit cards used to pay for medications, and location tracking.
    - *Create legislation for accountability and liability for public and private industry, including meaningful consequences for inappropriate sharing or using of information.
    - *Update standards to require data collectors and aggregators to demonstrate good practices for privacy protection.
    - Set strict, enforceable limits on what the entity receiving the data can do with it (uses, sharing/disclosing/selling, retention limits, etc.).
    - Expand or amend HIPAA to address aggregated or population level data.
    - Require opt-in consent from the person before sharing their information.
    - Remove the Right to Truth-in-Evidence clause from the California Constitution.
    - Legally restrict the fusion of healthcare data or sale of prescription or other data.
    - Prohibit pay-for-privacy schemes.
    - Eliminate loopholes in existing privacy laws.
    - Set laws to ensure both healthcare and non-healthcare organizations cannot bring harm to consumers by using protected health information.
    - Establish audit data requirements for non-covered entities and organizations with protected health information.
    - Information Privacy Act
        - Apply IPA to local agencies.
        - Amend IPA to limit data collection/sharing to what is strictly necessary for public health goals and inform patients of those protections.
        - Amend IPA to eliminate the non-consensual sharing of personal information (PI) with law enforcement in the absence of a court order/warrant.
        - Reduce the circumstances that allow sharing of personal information with law enforcement (IPA, CMIA, etc.).
        - Modernize IPA definition of PI to account for the possibility of reidentification.
        - Establish data minimization requirements within IPA to prohibit use of data beyond specified purpose.
- Strengthen Systems
    - *Leverage computational approaches for automated use control, including secure multiparty computation, confidential computing, etc.
    - Leverage the use of differential privacy for the 2020 U.S. Census to more broadly apply differential privacy, and related approaches, to key public health-related applications.
    - Conduct more public research into re-identification based on what information is already available.
    - Apply the Local Science Engagement Network (AAAS model).

- Collect only minimum necessary data
    - Apply the principles of least privilege and need to know.
    - Partition the data that are shared and minimize fields.
    - Require public health agencies to identify specific objectives in data requests to limit shared data to what is necessary to accomplish that goal.
    - Think of data as a toxic asset: collect as little as necessary to do the job, keep it only for long as necessary to finish the job, and properly destroy it when it is no longer needed.

Barrier: **Lack of clarity on which data can be shared and for what purpose**
Potential solutions:

- Redesign and improve systems
    - *Replace legacy systems with new, more flexible systems that enable secure data sharing.
    - *Incentivize up-to-date and accurate data collection and dissemination.
    - Design a new system with the different user frames explicitly identified.
    - Provide the funding to an independent party to drive system design and build. Include all relevant players for user inputs.
    - Build on AB 133 (Committee on Budget - Health 2021), broadening the population to all Californians in and out of the safety net.
- Monitor the purposes of data sharing to support privacy protection
    - *Improve accurate recordkeeping around data-sharing and its purpose.
    - Gather metadata behind data elements to delineate purpose and confidentiality.
    - Identify practices of concern, particularly related to highly sensitive or stigmatizing information and aggregation of health-related data with other types of information, which could result in greater ease of re-identification.
    - Institutionalize oversight of data sharing.
    - Shift the framework of data to patient-centered timelines instead of disease-specific reporting, to identify what data is shared and why, across an individuals' timeline.

## Panel Discussion: Actualizing Recommendations for Public Health

On day 2, CCST board member Bruce Darling moderated a panel, titled "Actualizing Recommendations for Public Health." The panel featured Susan Bonilla, CEO of the California Pharmacists Association and former California State Assemblymember; Debra Cooper, Chief Consultant for the California State Assembly Committee on Human Services; and Susan Fanelli, Chief Deputy Director of Health Quality and Emergency Response at the California Department of Public Health. Panelists drew on their individual knowledge and experience while sharing advice on how to successfully advance policy recommendations. Several key themes included:

- **Present a compelling rationale.** To make a strong case for a policy change, share quantitative data that provide concrete evidence for the problem, the need, and the proposed solution. Alongside the data, share human stories to paint a more complete and relatable picture of the impact this solution would have. While the data provides rationale, the stories are what legislators and constituents will remember.
- **Build relationships and activate your champions.** Engage a diverse group of stakeholders to support your initiative. This should include champions within the legislature and administration who can articulate and advocate for the need and recommended solution, navigating potential opposition as it arises. External support should also be built among community-based organizations, community advocates, healthcare, and other industry stakeholders.
- **Advance timely solutions and articulate short-term payoff.** Urgent needs and solutions are more likely to be prioritized and adopted than slower, long-term, complex, and preventive investments. To compete against other legislative priorities, we must articulate the short-term benefits of long-term solutions. When public health solutions gain increased attention and spotlight, such as during the COVID-19 pandemic, there is an opportunity to leverage the moment to build and sustain momentum for public health investments and initiatives.
- **Start small, if needed.** If you don't have the data to show evidence for the effectiveness of the solution, start with a pilot or with initial steps to gather evidence and develop a rationale. A low-cost proposal with clear checkpoints built in to pause, evaluate, and adjust, can make a policy more viable.

### Feasibility of Potential Solutions

Upon returning to their breakout groups, participants reviewed the most impactful solutions identified on Day 1 and were provided the opportunity to add new ideas. With takeaways from the panel in mind, participants then sorted the ideas based on feasibility.

**Red light** indicates the ideas expected to be the most challenging to advance (e.g., because they would be resource intensive, take 10+ years to advance, have significant political pushback, confront oppositional laws/policies). **Green light** indicates the ideas expected to be most easily adopted and implemented (e.g., those that require little to no resources, take 1-2 years, have political/institutional support, or are similar to previously successful efforts). **Yellow light** indicates ideas likely requiring effort and resources that fall somewhere in the middle.

| Red Light | Yellow Light | Green Light |
|---|---|---|
| • Address, clarify, and regulate data sharing | • Require entities to follow the most protective privacy law | • Bring impacted communities to the table |

| | | |
|---|---|---|
| among all employees within hospitals and hospital systems<br>• Establish a public health "congress" of experts who can define a universal structure for EHR fields and coding that is useful and practical for public health use<br>• Update standards to require data collectors and aggregators to demonstrate their practices for privacy protection<br>• Advocate and enact privacy legislation that extends to all health-related data, not just entities covered under HIPAA | that applies to any of the entities involved<br>• With community input, standardize a limited data set of reportable EHR elements to be shared systematically and electronically.<br>• Increase legal protections and enforcement mechanisms (e.g., PRA)<br>• Secure sustained and significant investment in leadership and HR to manage data systems<br>• Build a unified infrastructure for multiple use cases under HHS with a common technology backbone for community locations<br>• Secure public funding to strengthen data infrastructure<br>• Rethink the PII issues from the ground up<br>• Align data sharing norms and regulations within California Government to address the different rules for different diseases and conditions<br>• Consider flexibility for importing nontraditional data sets (wastewater, etc.)<br>• Replace legacy systems with new, more flexible systems that enable secure data sharing<br>• Improve recordkeeping around data sharing and its purpose<br>• Incentivize up-to-date and accurate data collection and dissemination<br>• Conduct monitoring and oversight to support | in meaningful ways for thinking through the risks and potential solutions to data sharing<br>• Improve interoperability and communication (e.g., EMR API integration)<br>• Standardize definitions and education on new and existing privacy laws<br>• In partnership with experts and communities, develop simple language guidelines for solicitation of PII that clearly explain why certain data elements are collected and used.<br>• Through a user-centered design process, develop a central, secure, interoperable technology platform for local communities, that supports appropriate data sharing<br>• Update the Information Practices Act (IPA) to address concerns about non-health entities under a government contract accessing health data<br>• Advance legislation for accountability to consumer protection<br>• Invest in data science pathways in workforce development and education<br>• Modernize privacy laws to account for current |

| | | |
|---|---|---|
| | reporting and compliance<br>• Identify and include the broadest definition of data creators and users when creating privacy requirements | realities (risks, new tech, etc.) |

## Recommendations

Participants self-organized into smaller workgroups that would each focus on a single proposed solution with the goal of developing more detailed recommendations for implementation. Each group completed a worksheet to identify 1) the problem they're working to address; 2) the specific actions they are recommending; 3) the benefits this recommendation would provide; 4) the stakeholders that should be involved; and 5) other considerations, including potential risks, funding needs, metrics, and barriers. The following summarizes participants' recommendations.

1. **Bring impacted communities to the table.**

   Addressing health inequities requires the ability to disaggregate health data by demographics and geography. While these types of data are valuable and in demand, CDPH often cannot share them because of privacy guidelines. Public health would benefit from a strategy to capture (and share) such data while being sensitive to community concerns and without exposing communities to risk.

   Participants recommend that California Health and Human Services engages impacted communities to co-design solutions that strengthen the collection and reporting of disaggregated data, while protecting the privacy of individuals. An iterative, community-informed practice will ensure data collection policies and systems are responsive, responsible, and refined over time. Meaningful community engagement offers an avenue for proactively addressing potential misunderstandings and publicly raised criticism. Specifically, participants recommend the following steps:

   a. Improve community engagement practices and strategies. This may include advisory boards with representatives from historically underserved populations (e.g., Citizen Advisory Boards). Seek meaningful, ongoing engagement with community stakeholders.
   b. Establish standardized data categories (e.g., race, ethnicity, income, disability, age, gender) and levels of analysis (e.g., individual, county, zip code, census tract) and ensure consistent use across agencies.
   c. Invite expert input from independent population scientists for objective input related to analysis and reidentification risk.
   d. Solicit statistical evaluation of risks at various levels of analysis/granulation and

identify safe protection thresholds. Consider updating the data de-identification guidelines (DDG), based on the current realities and risks.

2. **Extend privacy legislation to all health-related data collection, including entities not subject to HIPAA.**

   Sensitive health data are collected, used, and shared by numerous entities not covered by HIPAA. Vulnerable communities and individuals may be seriously harmed by consumer-provided health data disclosures by private companies (for example, if such services are currently criminalized in other states, as is the case with abortion). While extending privacy legislation to all health-related data collection (regardless of HIPAA coverage) would be beneficial, participants recognize that a more feasible first step may be required. As this first step to expanding protections to consumer health data, participants suggest requiring a separate opt-in step for consumers, wherein permission is granted for the collection, use, and storage of sensitive health-related information by a non-HIPAA-covered entity (as opposed to opt-out systems where consumers must take an additional, often obscured extra step to indicate that they would *not* like their data to be shared). Such legislation could build on the work of the California Privacy Protection Act and the California Privacy Rights Act.

   To advance this recommendation, participants suggest engaging and building support among privacy advocates, abortion rights groups, LGBTQ+ groups, civil liberties groups, and tech companies. Participants acknowledged that this legislation could have unanticipated impacts on research, to the extent that public entities and researchers rely on such data for analysis. Participants anticipate opposition by groups that benefit from the "sale" of health data and information. Funding would be required to support enforcement.

3. **Invest in workforce development to build fieldwide capacity to collect, manage, and utilize data.**

   There is a paucity of qualified public health experts trained in data quality, management, and analysis. To build current and future data competencies, participants recommend investing in cross-cutting and equity-informed curricula, training, and support across both educational and workforce systems. Over the long-term, such investments will be rewarded by more sustainable educational and job pathways, more equitable access to educational and workforce opportunities, and increased public knowledge and trust in data science. Participants recommend the following action steps:

   a. Within educational systems:
      i. Expand educational pathways and access to data-related courses and programs (IT, data science, epidemiology, public health, law) at

community colleges and higher education institutions.
    ii. Integrate data science and competencies into cross-cutting fields and majors
    iii. Introduce data competencies to elementary and secondary students, as integrated into core curriculum, to inspire the next generation of data scientists and collaborators.
b. Within the workforce:
    i. Expand continuing education and training for existing workers and personnel, to expand skillsets and stay abreast of evolving technologies.
    ii. Create new and more accessible hiring pathways, to support more equitable hiring practices and opportunities.

Advancing these recommendations will require sustainable funding, political and public support, and buy-in from educational systems.

4. **Modernize privacy laws to account for current realities.**

New advances in technology, like artificial intelligence and machine learning, have resulted in incredible advances in the ability for someone to reidentify "de-identified" data. To preserve privacy amidst evolving technology and increasingly salient reidentification risks, participants recommend that the California Legislature works to expand the scope of privacy laws that apply to state and local agencies. To protect privacy and close loopholes, while also maintaining interoperability, participants have the following recommendations:

a. Collect the minimum information necessary for the particular purpose for which the data is collected.
b. Legally restrict the use of data to only that which the data was intended.
c. Consider automated use controls—technical safeguards that would make it impossible to use data beyond its intended purpose.
d. Set restrictions and limits on data sharing with law enforcement. Consider making the Information Privacy Act more similar to the CMIA in its limitations to sharing data with law enforcement.
e. Expand the IPA to local entities to create uniform privacy protections and reduce confusion.
f. Redefine "personally identifiable information" to support differential privacy.
g. Consider requiring trusted execution environments (TEEs) and confidential computing (CC) as methods to remove cloud providers from the chain of trust. Usually data must be decrypted for analysis; TEEs and CC allow analysis even while data are encrypted.

Advancing these recommendations will require collaboration from impacted community members, community-based organizations, public health experts, and privacy experts.

5. **Standardize data systems and increase interoperability.**

Incompatible data systems create challenges for data sharing between and among public health agencies and medical providers. Participants recommend strengthening and standardizing data infrastructure to ensure seamless and timely access to data for an informed, appropriate, and time-sensitive policy response to pressing local and state public health needs. The data infrastructure should serve equity and outcomes work done by both public health and other health system partners. More specifically, participants recommend the following actions:

a. Redesign the data infrastructure and systems to have
    i. Data integration with user friendly visual tools.
    ii. Interoperability between EHR data systems and public health data systems (potentially HIE networks/organizations).
    iii. Consolidated data (clinical outcomes, equity data, etc.)
b. Train human resources for effective and secure data exchange (local and state).
c. Establish interoperability with public health partners and healthcare delivery systems at the local, state, and national level.
d. Consolidate public health data collection (e.g., consolidate three questions of race to one) to better serve the needs of public health programs and policies.
e. Participate in national standard data consolidation efforts, such as United States Core Data for Interoperability.
f. Leverage public funds in healthcare to modernize data infrastructure (e.g., MediCal, CoveredCalifornia, and CalPERS)
g. Standardize protocols, data collection, data dictionaries, and data transfer mechanisms for both state and local levels.

Advancing these efforts will require education, national coordination and collaboration, and sustainable funding. The successful implementation of this recommendation will require engagement with public health experts, information technology experts, policy advisors, and legal advisors, as well as clear guidance from the Center for Disease Control and the Office of the National Coordinator for Health Information Technology about data sharing permissions and expectations.

6. **Build a universally accessible and interoperable data sharing system.**

In the face of widespread emergencies and disasters laid bare by the COVID-19 pandemic, wildfires, and continuing infectious disease outbreaks and emergencies, California lacks the ability to share real-time, actionable data to drive decision-making to prevent, protect, and enhance the health of all Californians. Participants recommend building a universally accessible and interoperable data sharing system. Such a system would address the lack of a basic foundational infrastructure for technology and data sharing. Timely access to critical health data for all community services and partners

that impact public health (including education, social services, the criminal justice system, and others) will lead to better health outcomes for all Californians.

In pursuit of this interoperable data infrastructure, participants specifically recommend the following action steps:

a. Standardize technology operations.
b. Mandate or incentivize data sharing.
c. Fund basic technology infrastructure.
d. Assess equity regarding the capacity of all stakeholders to comply with data sharing best practices and requirements.

Advancing this recommendation will require funding for both building the technical infrastructure as well as incentivizing participation. The system must be designed to mitigate risks for special groups. Advancing this work would additionally require education and support from a wide variety of stakeholders, including associations representing organizations that benefit from data sharing, influential groups (e.g., labor), groups outside of the traditional healthcare system who may engage with data sharing, and groups representing communities and priority populations.

## Next Steps

Policy recommendations derived from this workshop will be further developed with the COVID-19 Steering Committee and released later this fall. These forthcoming recommendations will be peer-reviewed and informed by engagement with relevant stakeholders.

## Appendix A: Institutional Affiliations of Workshop Participants

| |
|---|
| American Civil Liberties Union |
| Boston Consulting Group |
| CCST Science and Technology Policy Fellow |
| California Department of Public Health |
| California Department of Public Health |
| California Department of Public Health |
| California Department of Public Health |
| California Department of Public Health |
| California Department of Public Health |
| California Polytechnic State University |
| California State Assembly |
| California State Assembly |
| Lawrence Berkeley National Laboratory |
| Manifest Medex |
| National Institutes of Health |
| National Science Foundation |
| Public Health Alliance of Southern California |
| Radiologist (Affiliation not provided) |
| University of California Berkeley |
| University of California Davis |
| University of California Davis |
| University of California Los Angeles |
| University of California Riverside |
| University of California San Diego |
| University of California San Francisco |
| University of California San Francisco |

## Appendix B: Data Sharing Use Cases

Prior to the workshop and with input from relevant stakeholders, CCST developed four use cases to discuss during the workshop (see Appendix B). The diverse scenarios provided concrete examples for which the groups would consider current and potential challenges, barriers, opportunities, and risks of data sharing. Each breakout group sequentially reviewed and discussed two use cases.

### Use Case A: Missed Opportunities for Treatment

A woman is diagnosed with syphilis who later becomes pregnant. The public health department was notified of the original diagnosis, but they lack accurate contact information to follow up with the patient after she becomes pregnant to recommend a shot of penicillin to ensure her baby does not contract congenital syphilis. The patient later goes to the Emergency Department for injuries after a fall. The patient does not notify the ER provider that she has syphilis. She therefore still does not receive the necessary dose of penicillin.

### Use Case B: Questions of Disease Severity

The public health department realizes that there is an increase in hospitalizations that may be due to COVID-19, but they do not have access to the clinical data that could help them understand whether these patients are in the hospital because of their COVID-19 infection or if they are in the hospital with COVID-19 but have been admitted because of something besides their COVID infection. They cannot easily see what vaccines the patients have had to understand if these are vaccine failures. Lastly, the public health department cannot easily see what treatments the patients in the ICU have had to understand if these are treatment failures.

### Use Case C: Database Notifies Providers of Rare Communicable Disease

A local public health officer receives a notification from a local emergency department about a case of acute flaccid myelitis (an uncommon but serious neurological condition that can be caused by viral infection). These data are then deidentified and uploaded into a database of all infectious disease reports accessible by researchers and other qualified parties for analysis and reporting. Concurrently the same system initiates an SMS alert to all providers within a self-selected geographic range who have opted in for alerts of the diagnosis and presenting symptoms

### Use Case D: Community Notifications about Possible Disease Spread

A university researcher is sampling wastewater streams from student dormitories. When positive results of a communicable disease are found, both dorm residents and the local public health department are notified. The public health department then notifies the general population about possible infection in the area with recommendation for testing.

## Appendix C: Barriers to Data Sharing

Participants identified numerous barriers to effective, efficient, and secure health data sharing (below). Because participants identified these barriers in two separate discussion groups, there are some redundancies. Participants in each group were asked to vote for the three barriers they believed were most amenable to policy solutions; cumulative votes from group participants are indicated in parentheses.

- Lack of infrastructure (capacity/staffing, data systems, funding at local levels, etc.) to collect and distribute data (11 votes)
- Lack of standardization in electronic systems, practice at point of care (7 votes)
- Inability to protect consumers when there are a number of parties outside healthcare that are not subject to regulation related to de-identification and re-identification (6 votes)
- Variability of interpretation in privacy laws and different laws (6 votes)
- Lack of data quality, data management, and analysis capacity (5 votes)
- Gaps between guidelines and execution; lack of consistency in implementation of policies and regulations (5 votes)
- Lack of funding and policies to be able to update our existing data/tech systems and have the proper staff to maintain these systems, as well as react as public health emergencies occur (4 votes)
- Gaps in identifying which data are shared and for what purpose (4 votes)
- Lack of sustained and significant investment in a) construction of public health data systems; and b) leadership and HR to manage those systems (4 votes)
- Lack of clarity on which actors and intermediaries can access data; the need to build firewalls – clarity on what accountability looks like (3 votes)
- Lack of interoperability between systems (3 votes)
- Concerns around reidentification or data breaches, including difficulty in providing guarantees against reidentification (3 votes)
- Fear and distrust about what will happen with data after you share and lack of control after sharing (2 votes)
- Potential weaponization of diseases and medical conditions that have been criminalized (2 votes)
- Lack of technical interoperability between healthcare organizations, local health jurisdictions, and state (1 vote)
- Concerns in minority communities about the use of data by law enforcement/immigration and health systems
- Concerns around discrimination and stigmatization
- Lack of a culture that the value of data is not distinctly for individual, or population uses and encourages innovation across health, community, and public health sectors to drive equity (1 vote)
- The state legislation, regulation, governance, and funding doesn't leverage California's ability to use technology at scale to reach the local level

- Lack of adequate, user-centered distribution of information
- Lack of training related to requirements and policies
- Lack of consistency and clarity in knowledge and implementation of what data can and cannot be shared
- Lack of clear articulation of what data is being collected, the benefit, and how data is being protected
- Misunderstanding of who data should be shared with and what is allowed
- Players outside of healthcare are more technologically advanced, creating risk of hacking
- Gap in codifying data sharing standards and privacy policies in statute
- Challenges to keep up with changes in technology, which creates a moving target to protect privacy
- Lack of IT expertise and training, particularly on the public health side
- Lack of ability or knowledge to mount response based on the data and insights
- Desire to avoid bad publicity / fear of public or political backlash in sharing data or preliminary insights with the public
- Concerns about legal liability may lead to caution or reluctance around data sharing
- Financial incentives