Data Sharing: Strategies for an Uncertain World

Presented by MIT CDOIQ in collaboration with Privacy Analytics, held on 19 May 2021

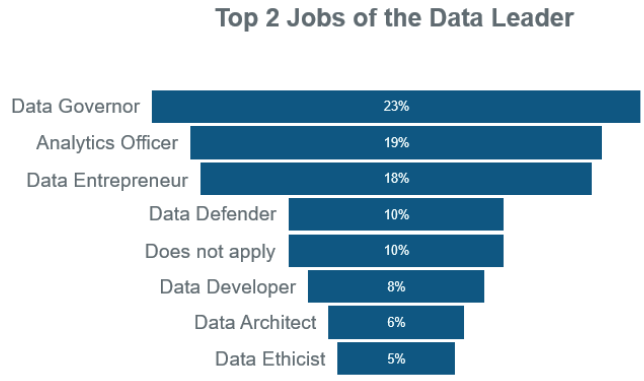
Based on background research conducted and in consultation with MIT CDOIQ, Privacy Analytics identified "data sharing" as an issue of significant importance to Chief Data Officers (CDOs) (and similar senior-level data and analytics leaders). In particular, our research found that CDOs care most about how to share data in an uncertain business environment. In this invite-only deep dive discussion, executive data leaders explored what's worked and what hasn't from across multiple industries, in both the private and public sectors.

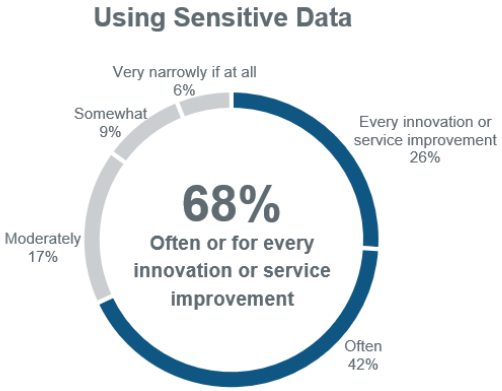
Three key takeaways emerged from this event:

- Effective internal data sharing is largely about addressing cultural issues and breaking down silos.
- External data sharing is no longer a nice-to-have but a necessity.
- Advances in technology allow for finely calibrated control over data access, making it easier to share data while addressing risk.

Why data sharing, and why now?

Of the 60 or so attendees, over 50 chose to participate in audience polls over the course of the two sessions (over an 80% response rate). The introductory poll results below show participant focus areas and the perception they have of their organization's risk appetite regarding the uses of sensitive data. We described sensitive data to be any data derived from people, be it identified with their name or transformed in any way, including aggregated. The results suggest that the top CDO jobs are governance, analytics, and entrepreneurship, and that sensitive data is often or always used to innovate and improve services.



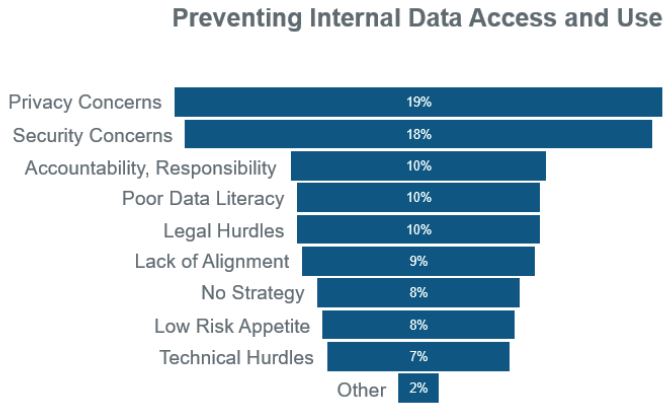


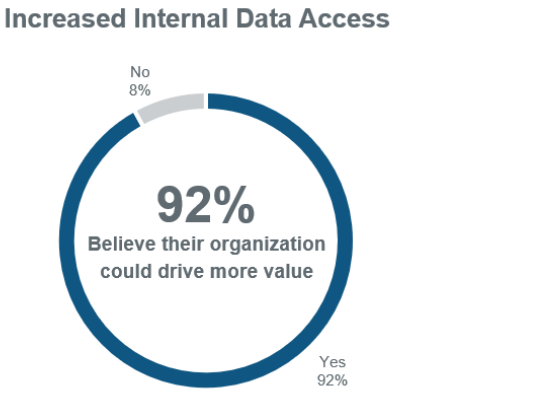
Enable data access and use to drive innovation

The panels explored how to achieve alignment across functions, break down traditional silos, and get data to internal teams to improve decision making and drive innovation. Advances in internal data sharing, often called data access, should set the stage for organizational success and a move to external data sharing.

Admittedly, one organization's "internal sharing" could be another organization's "external sharing", such as a state harmonizing data across departments. Legal restrictions may dictate some of these distinctions. For our purposes, we defined internal sharing as sharing and access from within the same legal entity, and external sharing as third-party access.

The poll results below show audience perspectives on opportunity and possible organizational pain points. They suggest that privacy and security concerns are main factors preventing internal data access and use, whereas increasing internal data access would drive more value for the organization.





The following ideas emerged from the roundtable discussions.

How to facilitate internal data sharing

- Move to a data-driven approach: Experiment and fail fast, foster a culture of data, promote data literacy.
- Create mechanisms to provide access to data internally: Make data findable, accessible, interoperable, and reusable (FAIR), unify and harmonize data, make it accessible for researchers, protect privacy, improve efficiency (example given of scalable anonymization).
- Know what data you have and its value by developing a good repository—knowing what's available is half the battle.
- Provide the necessary technology on the backend to support data management and analytics, and modernize enterprise architecture.
- Ensure buy-in from senior management.
- Set up a governance structure where responsibility is shared between business and IT.

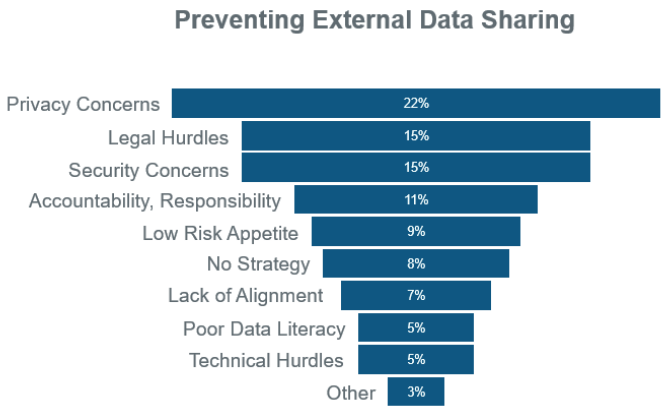
How to overcome silos and cultural barriers

- A barrier to internal data sharing can be different systems within a company that originated through acquisitions of other companies.
- Another barrier can be decentralized and fragmented systems: e.g., decentralized commercial healthcare in the US.
- Address cultural issues and how we think about data: e.g., data sharing is universally praised but seldom practiced.
- Avoid talk of “data owners”, which reinforces silos (language matters).
- Focus not so much on protection but on creation of value: e.g., data enablement through trustees.
- Recognize that one can adjust risk to various levels as appropriate (data collections, external reference structure, full transactional hubs).

Accelerate external data sharing to increase ROI

By external data sharing we mean making data accessible to people outside one's organization, such as allowing third-party access to data to collaborate on R&D initiatives, releasing data to partner organizations to leverage expertise such as AIML, or even publicly as open data. The challenges with external data sharing will be different, but some of the solutions for moving past traditional silos may be the same. Ultimately, we wanted to discuss data sharing on a large scale to ensure safe, accessible, and useful data.

The poll results below show audience perspectives on opportunity and possible organizational pain points. They suggest that privacy and security concerns, as well as legal hurdles, are main factors preventing external data sharing, whereas increasing external data access would drive more value for the organization.



The following ideas emerged from the roundtable discussions.

Why share data externally?

- To create benefits for the public while protecting privacy. We have a responsibility to do so as good data stewards.
- To improve business processes. The more real-time it can be, the better.
- To reduce costs, such as through consortiums.
- To monetize data (sharing and exchange). Demand exists because of a need for external data to drive business processes across the organization.

How to win stakeholder trust

- Offer something that stakeholders need: e.g., barter with marketing departments.
- Collaborate on strategic projects.
- Act like a data librarian: be the connector that they can't live without.

What is required for sharing access to data externally?

- Privacy by design is important.
- Take into account public emotions associated with privacy.
- Consider what the benefit of sharing will be and what is expected of the organization.
- Develop solid data policies for swapping of data.
- Take a marketing approach when communicating these policies. Alignment will come from how you package information, how you organize the discussion.

Delivering Results

The role of the CDO

- The CDO can play the role of breaking barriers. One barrier is that different groups often don't speak the same language. Another is that they don't trust one another.
- The CDO can be a translator, a facilitator of the conversation who can organize the discussion and build alignment.
- When the CDO reports through the business, this can help balance between business and IT. When the CDO comes in, there is often a sense of relief.
- External data sharing has to be a new strategic capability that the CDO needs to drive. And they need a repeatable, predictable set of technologies to do this.

The discussion was rich with insights and this summary only captures some of the main points by theme.

Acknowledgments

We wish to thank the executive data leaders that participated in the discussion by sharing strategic insights based on their combined knowledge and experience.

- Besa Bauta, Chief Data Officer, MercyFirst Hospital
- Cortnie Abercrombie, CEO and Founder of AI Truth
- Derek Strauss, CEO, Gavroshe
- Dr. John Talburt, Acxiom Chair of Information Quality, University of Arkansas at Little Rock
- Maria Villar, Head of Enterprise Data Strategy and Transformation, SAP Americas
- Michael Conlin, Chief Technology Officer, Definitive Logic
- Robert Elliott, Director of Bioinformatics – Anonymization Lead, Novartis Pharmaceuticals Corporation
- Robert McGough, Deputy Chief Data Officer, State of Arkansas
- Sorana Ionescu, Director, Smart Metering, Independent Electricity System Operator (IESO)
- Stephen Bamford, Head of Clinical Data Standards & Transparency, Janssen Research & Development

Finally, we wish to thank the host, Richard Y. Wang, Founder & Director of the MIT CDOIQ program.

About Privacy Analytics

Privacy Analytics provides data anonymization services and software with auditable proof of privacy protection. Our expertise has enabled business leaders at more than 200 organizations across consumer and healthcare industries to safely deploy transformative data strategies. With 100+ in-house experts and scalable commercial software, we've anonymized thousands of datasets over our 14 years in business, giving clients an edge in the hyper-competitive race for innovation and new revenue streams. Our proven approach to data anonymization makes sure our clients are CCPA, HIPAA, and GDPR-ready, turning their most sensitive data into their most powerful asset. www.privacy-analytics.com

About MIT CDOIQ

The International MIT Chief Data Officer and Information Quality Symposium (MIT CDOIQ), now in its 15th year, is one of the key events for sharing and exchanging of cutting-edge ideas, content and discussions. Our purpose is to advance the knowledge and accelerate the adoption of the role of Chief Data Officer (CDO) in all industries and geographical countries. As Data is a critical aspect of every organization, the symposium is focusing on the management and leadership of this critical element for the 21st century that will benefit every organization.

Against the backdrop of Data Analytics, Machine Learning, Data Quality and Data Management, the MIT CDOIQ Symposium will host its event virtually and will explore delivering mature data & analytics capabilities for ROI, including local organizational issues to global challenges, through case studies from industry, academic, financial, government and healthcare leaders. mitcdoiq.org