



Considerations for Private-by-Design Dynamic Consent to Digital Health Research

Yuriy (Kat) Pavlov^{1*}, Jonathan Dursi¹, Ming Hin Fok², Guillaume Bourque³ and Michael Brudno¹

1: University Health Network, 2: Independent Researcher, 3: McGill University

* ORCID iD: Yuriy Pavlov <https://orcid.org/0000-0002-4847-7284>

for GenoPri 2021

Sept 22 2021

Hello everyone! I'm going to be speaking with you today about our experiences designing a private dynamic consent system for use on a federated national health data-sharing project.



CanDIG: about us



The **Canadian Distributed Infrastructure for Genomics**

- Multi-jurisdictional and multi-institutional
- Sharing of genomic data and metadata while keeping the data under local access control and security protocols
- Driver project for GA4GH; especially interested in federated solutions

(Dursi et al. 2021)

My name is Yuriy Pavlov. I'm a technical analyst at the DATA lab of the University Health Network, working on the Canadian GA4GH driver project CanDIG. The multi-jurisdictional nature of Canadian healthcare means that national health research projects that rely upon data centralization are difficult to negotiate and sustain. Thus, the CanDIG team is developing a distributed, or federated, infrastructure for health research computing.



Outline

- What is **dynamic consent**, and why should it be an access control priority?
- Consent theory: key **functional specs** of dynamic consent systems
 - **Disclosure** of research protocol
- **Private** and **secure** dynamic consent

In the course of designing and implementing a dynamic consent system for CanDIG, we have so far encountered various challenges pertaining to security and the preservation of research participant privacy. We would like to share our experiences with the attendees of GenoPri 2021, and suggest considerations that any health research platform interested in dynamic consent would benefit from including in their Research & Design.

I'll begin this talk by defining and justifying dynamic consent as a complement to privacy preservation. I hope to convince all of you to think of consent as a priority on your roadmap for Authentication and Authorization Infrastructure, rather than an add-on. From the point of view of consent theory, I'll then discuss key functional specifications for dynamic consent systems, such as disclosure mechanisms. I'll then mention considerations for designing private and secure dynamic consent.



Defining dynamic consent

What is:

- Autonomy?
- Consent?
- Dynamic Consent?

Consent is primarily a mechanism for the preservation of the right to autonomy, also known as self-determination. Dynamic consent is a consent framework specific to research. I will define dynamic consent by way of these concepts.

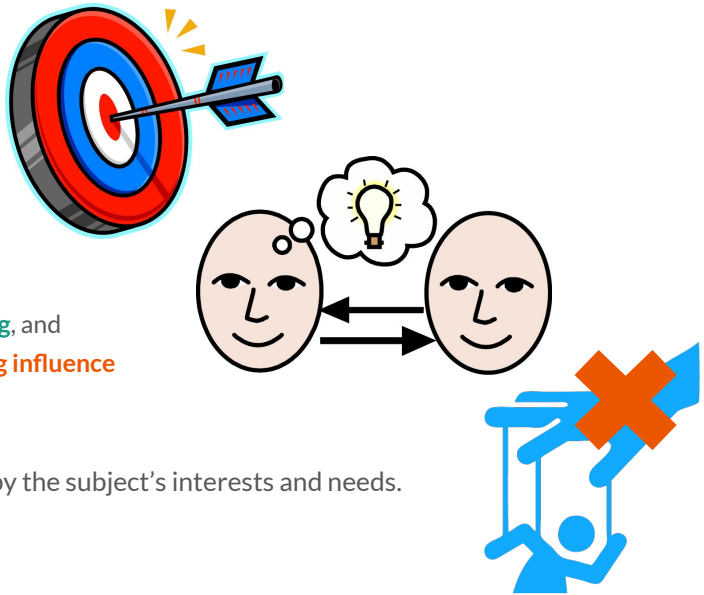
Autonomy

P acts autonomously only if P acts:

1. [substantially] **intentionally**,
2. with [substantial] **understanding**, and
3. without [substantial] **controlling influence**

(Faden and Beauchamp 1986)

Autonomous actions are informed by the subject's interests and needs.



[slowly]

In their book on the History and Theory of Informed Consent, Faden and Beauchamp define autonomy as follows: “P acts autonomously only if P acts substantially (1) intentionally, (2) with understanding, and (3) without controlling influences.”

A person that performs an action *autonomously* is substantially empowered in that action. The practical consequences of this empowerment is that they choose to perform this action in the context of their interests and needs. This has positive emotional consequences for the person. More importantly, this personal empowerment has material consequences, especially in the context of structural inequity where institutions cannot reasonably be trusted to have the interests and needs of oppressed people at heart.



Informed consent and refusal


Informed consent has a sense of being an autonomous authorization by a subject.

Similarly, **informed refusal** is an autonomous nonauthorization.

(Faden and Beauchamp 1986)

Autonomy is a moral duty and a right, and consent is a mechanism for the preservation of this right. Faden and Beauchamp describe the moral definition of informed consent as being an autonomous authorization by a subject, or an autonomous nonauthorization in the case of informed refusal.

This notion of consent as autonomous authorization, as well as the criteria for autonomy laid out in the previous slide, will help us to manifest the ideal concept of consent into a technical solution:



Key components of a consent system

1. An **authorization infrastructure** that uses participant-writeable data access decisions
2. An **information-disclosure system** that empowers participants to make these decisions autonomously

First, an authorization infrastructure that uses participant-writeable data access decisions, and

Second, an information-disclosure system that empowers participants to make these decisions substantially autonomously.



Why dynamic consent, specifically?

Theoretically, dynamic consent ~= informed consent applied to **networked digital infrastructures**

Historically, dynamic consent is a response to the limitations of the **broad consent** framework, ex.:

- Reliance on “consent” to *future* research protocols, which cannot by definition be informed consent
- Reliance on a constant level of risk, which does not scale with the accelerating commodification and criminalization of data
- Biomedical research subjects are often solicited under duress


Practically, dynamic consent synthesizes the moral imperatives of informed consent with the benefits of **patient portals**, facilitating the building of trust and participation by the data subject.

Having defined a high-level set of technical specifications for a consent system (using only principles of informed consent elaborated many decades back,) why are we specifically interested in dynamic consent?

Dynamic consent can roughly be defined as the attempt to apply informed consent to the networked digital context. It leverages the relatively widespread availability of home internet to extend the consent process from the clinic to the home.

Dynamic consent is also historically foregrounded by the framework of broad consent, which was a popular model for managing participant consent in the context of data-sharing, but has limitations that may be overcome by the technologies available to us today.

In practice, dynamic consent has a lot of overlap with principles of patient portal design. It benefits research efforts by facilitating the building of trust and participation by the data subject, and potentially by the broader culture that the subject is situated within.



Key components of a **dynamic** consent system

1. An **authorization infrastructure** that uses participant-writeable data access decisions
2. An **informational system** that empowers participants to make these decisions autonomously
3. An accessible **user interface** for #1 and #2

With that said, the main elaboration contributed by dynamic consent to our definition of a feature-complete consent system is this:

- An accessible user interface for participants to review pertinent information and edit their authorization-related data



Why discuss consent at a privacy conference?

The solicitation of consent is the basis of a research **participant's trust** in a project. (Woolley 2019)

Consent is often treated as an add-on to AAI systems, despite the facts that:

- Consent solicitation is a **moral duty**
- Consent systems are **technically feasible**


Privacy and security technologies are often leveraged to **justify the absence of consent**, despite the facts that:

- Privacy does not atone for the emotional nor material consequences of a breach of consent
- Security does not atone for the emotional nor material consequences of a breach of consent

Patrick J. Woolley (2019) asserts that the solicitation of consent is the basis of a research participant's trust in a project. Yet despite the moral duty to solicit consent, especially in the research sphere, and the technical feasibility of elaborate authorization systems at this point in time, consent is often deprioritized.

One reason to assert this point at the International Workshop for Genome Privacy and Security is that privacy and security are very often leveraged to justify the absence of consent. This happens so frequently that many of us are accustomed to accepting privacy and security as good-enough substitutes for consent. While privacy and security are important, and are often useful for **enabling** a prospective participant's consent, fundamentally they solve a different problem from consent and cannot replace it.

If you don't want an action done to you or your data, secrecy will not resolve the violation of your consent. While it can reduce embarrassment, it will not eliminate the sense of disempowerment, nor the material consequences of the action that was done to you against your interests.



Functional specs of dynamic consent

- Digitizing consent
- Informing participants
- Soliciting “meaningful” consent

Hopefully I have convinced some of you of the “whys” of dynamic consent. Next I would like to turn to the “hows”.

How can consent be digitized?

What must one do to adequately inform a participant in the digital context?

How can one solicit sustainably *meaningful* consent?

Since concrete specifications will vary according to regulations and resources, I will focus on asking questions rather than offering solutions, but I’ll provide some suggestions where appropriate.



Informed consent and refusal

Informed consent has a sense of being an **autonomous** authorization by a subject.

Similarly, **informed refusal** is an autonomous nonauthorization.

(Faden and Beauchamp 1986)

Let us briefly return to Faden and Beauchamp's definition of informed consent as an autonomous authorization by a subject.

What is authorization? (1/3)

The authorizing party assumes responsibility for an action, while simultaneously transferring the authority to enact that action to the authorized party.

The transfer involves both **permission and responsibility** to do something.

(Faden and Beauchamp 1986)



Faden and Beauchamp further clarify their definition of informed consent by defining authorization as being:

The transfer of action-enacting authority from one party to another, with the authorizing party retaining responsibility for the action. This transfer involves both permission and responsibility to act.



What is authorization? (2/3)

Data access authorization

= **permission and responsibility for the AAI system** to grant researcher R access to a datum, (or to at least communicate to the authorizer if access is not granted.)

Informed consent for data sharing

= autonomous data access authorization

= **autonomously given permission and responsibility for the AAI system** to grant researcher R access to a datum, (or to at least communicate to the authorizer if access is not granted.)

Thus we can define informed consent for data sharing as being an autonomous data access authorization given by the data subject. We can similarly define informed refusal as being an autonomous data access nonauthorization given by the data subject.

Consents can be digitized as authorization metadata, used by access control for resource-level data access decision-making.

[slowly]

Therefore: consents can be digitized as authorization metadata, used by access control for resource-level data access decision-making.



What is authorization? (3/3)

Authorization may be **broad** (follow general rules) or **narrow**.

Assent (a submission to a plan) points towards authorization but is **explicitly not an authorization**.

(Faden and Beauchamp 1986)

Assent example: incomprehensible End-User License Agreements

The specific formats that this authorization metadata takes can be diverse; indeed, Kaye et al. (2015) encourage you to support broad consent within your dynamic consent system.

However, it's important to mindfully distinguish between what consent fundamentally is and is not. For example, the acceptance of an End-User License Agreement that is long, inaccessible, and offers no alternatives is *at best* an act of assent, not consent. In this example, there is a failure to solicit informed consent stemming from a failure to meaningfully inform the user of the protocol that they are agreeing to. Also, if the user must absolutely gain access to the software, say for example due to monopoly, their agreement to the license is partially coerced. Thus they cannot do better than assent to the terms.

Information / disclosure

One of the requirements for an action to be “autonomous” is that it is substantially informed. The disclosure phase of informed consent is critical.



Autonomously-given authorization requires information

What human intermediaries will be available for acting as **informants** in production?

What features are needed for these intermediaries to be able to work efficiently?

Plan to **manage the remaining burden of knowledge** on the participant.

In a digital context, the burden of knowledge upon the participant can get to be very high, especially when there is no easily accessible clinical intermediary involved. Depending on the resources available, it may be impossible to get some subjects to the level of understanding that they could reach when undergoing a clinical consent process with a human intermediary.

It's important to make an honest assessment of what human resources will be available once the dynamic consent system goes into production. Once this has been assessed, make a realistic plan for minimizing the remaining burden of knowledge upon the participant.



Disclosure for substantial understanding has two components

Core disclosure (~= “reasonable person standard”)

Initiate communication process with the disclosure of core information that the professional perceives as being **substantially relevant and generally material**.

Goal: subject and requestor are negotiating the **same task**.

Subjective disclosure

Continue the communication process by soliciting and integrating **feedback**/active participation.

Goal: subject understands the information that is **subjectively material to their decision**.

(Faden and Beauchamp 1986)

The goal of disclosure is to convey substantial understanding of a protocol, its risks, and its alternatives to the participant. There are two components to disclosure. Core disclosure is the goal when disclosure is initiated. Subjective disclosure may be accomplished in the conversations that follow.

The purpose of core disclosure is to ensure that the subject and requestor are negotiating the same task. This is accomplished by selecting and communicating a core set of generally material information about the protocol.

Subjective disclosure then elaborates upon this baseline understanding by delivering information that is subjectively material to the participant's decision. The goal is to ensure that any information that would make a participant change their mind if known *is known*. Feedback loops are necessary for achieving this; it is impossible to determine what is material to a specific participant without feedback.



Core disclosure: minimizing the burden of knowledge

- **Software documentation**
 - Written for the needs and interests of the data subject (consult a patient group or clinician?)
 - Written in accessible language
- **Data Access Committee (DAC) portal**
 - Dataset-level authorization decisions to filter out outright inappropriate data access requests
 - Ask if DAC members could help inform data subjects through this portal
 - ex. the Resource Entitlement Management System (REMS) (Linden et al. 2013)

If access to clinical intermediaries is limited, consider leveraging some of the following resources to reduce the burden of knowledge upon the participant during core disclosure:

Document the key decisions and tradeoffs made in designing your software, and translate the main ones into accessible terminology targeted at research participants.

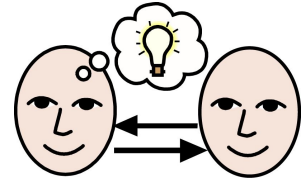
Also, you can leverage the experience and screening protocols of a Data Access Committee (or DAC) by incorporating a DAC portal, such as REMS: the Resource Entitlement Management System. DACs can make dataset-wide authorization decisions through this portal, and could potentially furnish participants with information about the research project and its protocol.

Linear notification systems cannot accomplish subjective disclosure.

Notification systems that are only capable of delivering information from the requestor to the subject can accomplish core disclosure.

Unfortunately, however, they lack the embedded feedback loops necessary to accomplish subjective disclosure. We could collect feedback from subjects and intermediaries to iteratively improve the materiality of the core disclosure, but some subjects will still fall through the cracks and not receive the level of material understanding that they would be able to receive in-clinic, where they can ask questions or clearly demonstrate failure to understand a protocol.

Digitizing subjective disclosure



- Dedicate as many **human intermediaries** (informants) as possible
- Build the **high-quality interfaces** necessary to enable human intermediaries to do their work efficiently
- In theory, could leverage **ML** informants that train continuously on user-specific data
 - Adds a feedback loop to the disclosure process
 - Fitting a model this closely to user data is a privacy risk
 - Federated learning a promising solution?

In my opinion this is a very major challenge for dynamic consent at the large scale, and one of the core reasons that dynamic and informed consent cannot be equated. The gold-standard solution to this problem is to leverage the skills and minds of as many trusted human intermediaries as possible.

It's also worth noting that in theory, federated learning algorithms could also build in the feedback processes needed to accomplish subjective disclosure.



Keeping digitized consent meaningful

- **Study the clinical consent processes** that your users undergo, and replicate them where possible
- Involve consent system **end-users** (ie. data subjects) in the design of the consent system as **early** as possible
 - Improves user adoption
 - Reduces future re-development work
- Make the consent and notification process as **user-configurable** as possible
 - Reduces information overload consent fatigue

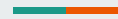
Although subjective disclosure presents a substantial challenge to digitization, other aspects of meaningful consent are more easily implemented.

The data subject will more easily follow a digitized consent process if it resembles the consent process they experienced in-clinic.

Involving patient groups in the design of the consent system will both improve user adoption out of the gate and reduce future development hours.

Making consent style and notification user-configurable can reduce information overload and consent fatigue.

These strategies can help keep data subjects engaged and participative in the research.



Secure dynamic consent

Separating data domains

- Physically
- Temporally

Federation

The final segment of this talk will cover the relationship between consent and privacy, and posit considerations for secure dynamic consent systems.

I will especially rely upon domain-driven design to define distinct boundaries between data types and enforce physical separation between them. I will also briefly discuss dynamic consent in the federated context, for situations where privacy and security requirements are heterogenous and access control must be decentralized.



Consent: an antagonist to anonymity

Participants who wish to be **notified** of new requests on their data **give up anonymity**, since there will be some association between their contact information and health data.

(Anonymity may persist for participants that choose not to enroll in the dynamic consent system, and possibly even those who simply choose to forfeit notifications.)

Research participants who wish to receive notifications for dynamic consent must register some contact information that re-identifies them. These participants give up anonymity in order to fully participate in the dynamic consent system.

To manage the risks that this linkage introduces, we must be careful to design consent systems to be as private and secure as possible.



Separating and securing domains

The scope of a breach can be minimized by the separation of domains, such that:

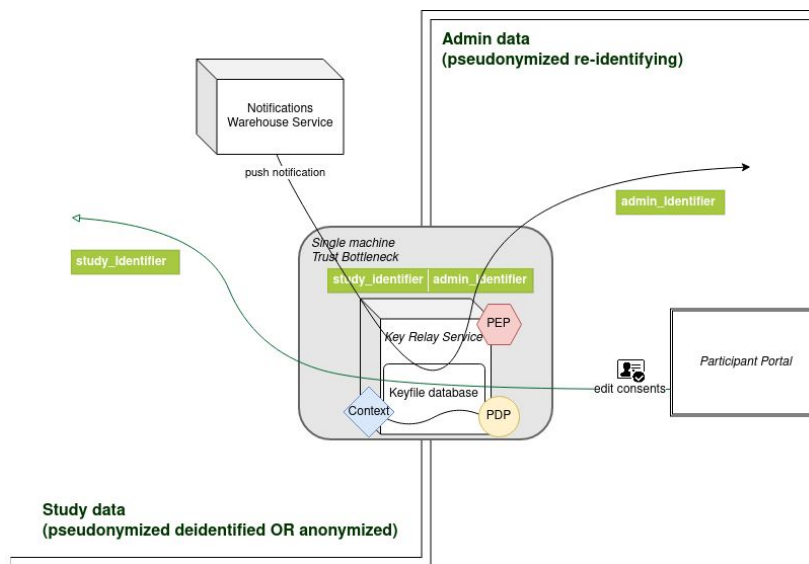
- Each domain is protected by its **own access control policies**
 - Each user role is only able to access the **domains that pertain to them**
- Domains may be **physically separated**

Goal: **minimize data** and **deter linkage attacks across domains**.

The scope of a breach can be minimized by separating domains such that each domain is protected by its own access control policies, including restrictions by user role. If possible, domains should be physically separated as well; they should be served from different machines.

**There is no need to keep
re-identifying contact data on
the same machines as medical
study data.**

Consider that there may be no need to keep re-identifying contact data on the same machines as medical study data.



There is no need to keep the **administrative <-> study data associations** on the same machines as either data service.

Moreover, there may be no need to keep the associations between contact and study records on the same machines as either of their servers. If the three data types live on different machines, and keep each other at arms' length, a breach to any of these services would not be sufficient to re-identify the medical data.

The diagram pictured here visualizes the movement of consent modifications and notifications between the study and administrative domains, through a trust bottleneck that translates the calls between these domains.



Designing speed bottlenecks

Where appropriate, design **slow processes** into the system.

So long as **time-related expectations** are communicated to research participants, it's okay to prioritize **security** over speed.

Remember: clinic-based research consent/revocation often takes days to propagate through the system.

Another way to reduce the risk of linkage across domains is to build in intentionally slow processes, where appropriate. As long as time-related expectations are communicated to participants, it's okay to prioritize security over speed. After all, clinically-mediated informed consent for research can be very slow, and still constitute a satisfactory and ethical consent framework.



Bottleneck-by-design example 1:

Immediate notifications for new data access requests are

- minimally beneficial
- potentially identifying

Thus, participant-bound notifications can and likely should take time to propagate through the system. It can even be a good idea to randomize the amount of time taken for a notification to propagate.

The first example of a good bottleneck candidate is the notification system. Sending notifications immediately following new data access requests is not only negligibly beneficial, but is actually risky from a privacy standpoint, in that the delivery of a notification soon after the submission of a new data access request may strongly indicate membership in the requested dataset.

Thus, participant-bound notifications can take time to propagate through the system. It may even be a good idea to randomize the amount of time taken for a notification to propagate, or to hold back notifications and send many of them in scheduled waves.



Bottleneck-by-design example 2:

Immediate updates to consensual authorization metadata are:

- minimally beneficial
- potentially identifying

Thus, participant updates to authorization metadata do not need to be immediately enacted (although the delay should likely not exceed several days.)

Similarly, another good candidate for a bottleneck is the propagation of consent updates through the system. Again, it is tolerable for updates to take a day or two to be enacted in authorization decisions, so long as this time expectation is communicated to the participant.

Propagating updates to consents on a scheduled basis, rather than immediately after the update is written, can even help to protect the privacy of the participant.



Bottleneck-by-design example 3:

On the other hand, authorization filtering for the real-time retrieval of study data should be relatively fast.

Thus, authorization metadata always should be “within reach” of the study data service’s access control.

On the other hand, filtering study data by its associated authorization metadata should happen in relatively real-time. This means that authorization metadata, derived from consent metadata, should always be within reach of the study data’s access control.



Bottleneck-by-design examples 2+3:

Participant-writeable authorization metadata can both:

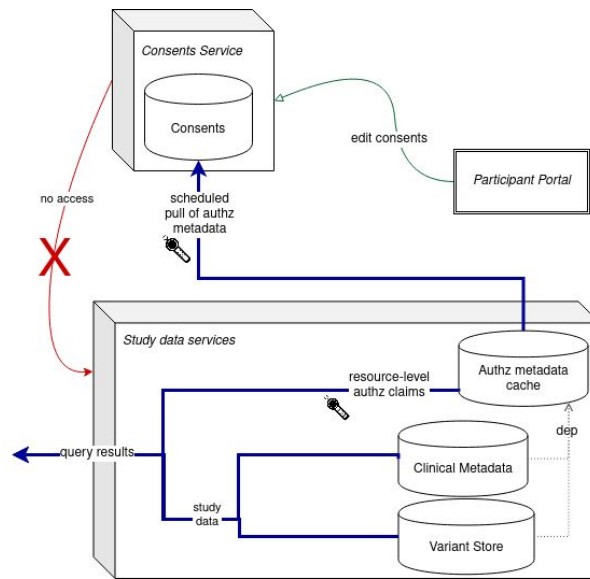
- have its source of truth in a Consents service
 - Unlike a study data server, a dedicated consents server **can accept read/writes from participants**
- be cached in the data services' databases for use in rapid authorization.
 - **Tagging study data with its authorizations** improves the retention of contextual integrity, in Nissenbaum's sense of contextual integrity as the basis for privacy (2004)

Updates from the Consents service can be pulled on a nightly basis.

Combining the speed-related specs from the previous two examples, we could, for example, approach the storage and use of consent metadata as follows:

Consents can have their source of truth in a dedicated consents service, and also be cached in the data services' database for use in rapid authorization.

This design allows us to tag the study data with its authorization context, without needing to expose the study data service to reads or writes from participants. The cache can be updated on a nightly basis.



Separating **consents-derived authorization metadata** from **consents** to optimize security, speed, and contextual integrity.

This diagram visualizes this separation, as well as the cache-update process.

[pause]

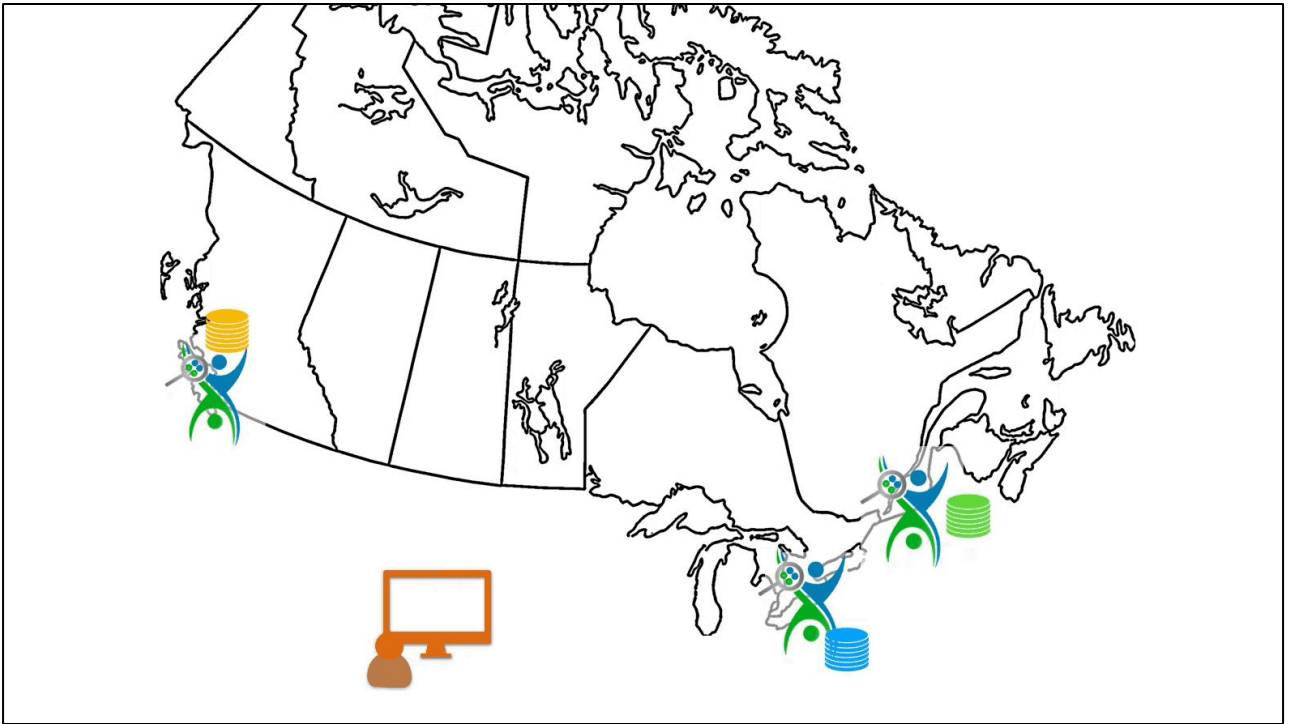


Federation

Federated systems

- A solution for inter-jurisdictional data sharing
 - ex. Interprovincial healthcare data-sharing networks in Canada
 - ex. International data-sharing networks
- Features local **access control** and **security**
 - Satisfaction of heterogeneous privacy and security specifications

Data federation is a good strategy for enabling the sharing of data between jurisdictions, such as provinces or nations. Federated systems can satisfy a heterogeneous set of privacy and security specifications by enabling each peer to “bring their own” access control and security solutions.



Here, an incoming query is federated to two other CanDIG sites in faraway provinces, who respond to the query according to their own policies, before the data is aggregated together and returned.



Federated dynamic consent

Federated dynamic consent

- Local **storage of authorization metadata** for use in local access control
- Local **schemas** for consents and the authorization metadata derived from them
- Custom/local **frontends** for participants, intermediaries, etc.

There's no need for consents/authorization metadata to move out of their home site.

Dynamic consent plays well with federated systems.

Most of the functional specs of dynamic consent systems pertain to access control and frontend design. Since these functions can be customized at their home site without interfering with the data-sharing functionality of the federated system, dynamic consent and federation are very mutually compatible.

Thus, federated infrastructures are promising premises for dynamic consent in international data-sharing contexts.



Conclusion

- What is **dynamic consent**, and why should it be an access control priority?
- Consent theory: key **functional specs** of dynamic consent systems
 - **Disclosure** of research protocol
- **Private** and **secure** dynamic consent

And with that, our time together today is over.

I hope that I have succeeded in justifying dynamic consent as an access control priority, and provided a cursory set of functional specifications for dynamic consent systems. I spent some time discussing the process of disclosure in the digital context. Finally, I posited some considerations for designing dynamic consent systems with privacy and security built-in.

In so doing, I have provided you with an overview of the topics and challenges pertinent to private and secure dynamic consent in health research. Thank you so much for having me, and my colleagues from the CanDIG project, here today.

Thank you for your time!

**Please reach out to me if you have
any questions I'm not able to answer
here today:
katpavlova.ccm@gmail.com**



Works Cited

- Dursi, Lewis Jonathan, et al. "CanDIG: Secure Federated Genomic Queries and Analyses Across Jurisdictions." bioRxiv (2021).
- Emam, Khaled El, et al. "A review of evidence on consent bias in research." *The American Journal of Bioethics* 13.4 (2013): 42-44.
- Faden, Ruth R., and Tom L. Beauchamp. *A history and theory of informed consent*. Oxford University Press, 1986.
- Kaye, Jane, Edgar A. Whitley, David Lund, Michael Morrison, Harriet Teare, and Karen Melham. "Dynamic consent: a patient interface for twenty-first century research networks." *European Journal of Human Genetics* 23, no. 2 (2015): 141.
- Linden, Mikael, Tommi Nyrönen, and Ilkka Lappalainen. "Resource Entitlement Management System." Maastricht, Netherlands: TERENA, 2013.
- Nissenbaum, Helen. "Privacy as contextual integrity." *Wash. L. Rev.* 79 (2004): 119.
- Ploug, Thomas. "In Defence of informed consent for health record research-why arguments from 'easy rescue', 'no harm' and 'consent bias' fail." *BMC Medical Ethics* 21.1 (2020): 1-13.
- Pricor, Megan, et al. "Dynamic consent: an evaluation and reporting framework." *Journal of Empirical Research on Human Research Ethics* 15.3 (2020): 175-186.
- Rothstein, Mark A., and Abigail B. Shoben. "Does consent bias research?" *The American Journal of Bioethics* 13.4 (2013): 27-37.
- Woolley, J. Patrick. "Trust and Justice in Big Data Analytics: Bringing the Philosophical Literature on Trust to Bear on the Ethics of Consent." *Philosophy & Technology* 32, no. 1 (2019): 111-134.



Works Cited (cont.)

Bullseye clipart: <http://clipart-library.com/clipart/5TRxEgkc.htm>

Understanding clipart: <http://clipart-library.com/clipart/395053.htm>

Marionette clipart: http://clipart-library.com/clipart/coercion-cliparts_11.htm

Frozen flower photo: <https://unsplash.com/photos/hUp58GsPKAw>

Fingerprint photo: <https://unsplash.com/photos/SRFG7iwktDk>

Extra slides



Common consent styles

Opt-in: **by default, I do not consent** to be a part of the dataset. I may consent to becoming part of the dataset once I am informed about the research protocol.

Opt-out: **by default, I consent** to be a part of the dataset. I may revoke consent once I am informed about the research protocol.

Consent may be digitized and managed in a variety of formats. Ideally, users should be able to pick the consent styles that work for them. An opt-in consent style means refusal-by-default, whereas an opt-out consent style means consent-by-default. Both should be supported in most dynamic consent solutions.

Consent: a complement to privacy

Consent can be a protocol for **transparency**: communication about risks that can't yet be eliminated.

- Anonymity is difficult to future-proof
- Genetic data is inherently identifying
- Many strong privacy-preserving methods are only applicable to aggregated data



Consent can complement privacy, especially where privacy technologies near their limits, like in case-level biological data. Where anonymization, deidentification, and privacy-preservation fall short, dynamic consent can empower a subject to make risk-informed decisions about their data.

Effects of dynamic consent upon data analysis

Now I will briefly touch on the effects of dynamic consent systems upon data analysis.



Consent bias

- **Real**, but very **unlikely to ever be one of the primary sources of bias**
(Rothstein and Shoben 2013), (Emam et al. 2013)

Consent bias, the usual argument against informed consent in research, is real, yet it is negligible compared to the moral imperative to solicit informed consent. Additional information is provided in the additional slides at the end of this deck, but I will skip this discussion to focus on bias specific to dynamic consent.



Consent bias: additional information

- The solutions proposed by consent bias discourse (data privacy and security) are not the safeguards to autonomy that they are made out to be.
- **Are the benefits of eliminating consent bias worth the violation of respect for autonomy?**
 - Consider the impacts of mistrust in research
- See Ploug's analysis of the ethical importance of informed consent to secondary research (2020)



Effects of drop-out upon a dataset

Worst case: all members of a dataset prefer to opt-out, and do so at staggered rates

Probable case: anecdotally, it seems that data subjects seldom use dynamic consent systems to revoke consent after giving it (discussed in the Feb 27 2021 *Workshop on Dynamic Consent* hosted by DNV and Australian Genomics).

May be able to offset these effects by freezing datasets.

The effects of participant drop-in and drop-out upon a dataset, which can cause the size and contents of a dataset to fluctuate over time, constitute sources of bias specific to dynamic consent that can substantially affect experimental conditions.

The worst-case effects of drop-out, for example, would involve the slow and unpredictable depopulation of the dataset over time. Anecdotally, dynamic consent systems are seldom used to revoke consent to research after consent has been given.

A viable strategy for managing both drop-in and drop-out effects is dataset freezing.



Effects of drop-in upon a dataset

Worst case: all members of a dataset prefer to *opt-in*, and do so at greatly delayed and staggered rates

Probable case: unknown

May be able to offset these effects by:

- Implementing differential privacy algorithms for aggregated data analysis
- Educating data subjects about differential privacy
- Encouraging data subjects to set their consent styles for **research on differentially private aggregated data** to *opt-out*

The worst-case effects of drop-in upon a dataset would involve the slow and unpredictable population of the dataset over time. Implementing differentially private algorithms for use in research on aggregated data, and encouraging participants to set *opt-out* preferences for their data in this case, could help to reduce the severity of drop-in effects.

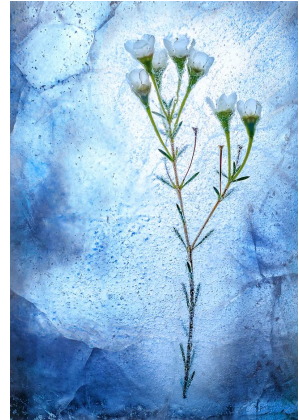
Dataset freezing

Dataset freezing is a **compromise** between:

- support for dynamic consent
- complete static datasets

A deadline for opt-in/opt-out decisions is communicated to the data subjects, and the dataset is frozen thereafter.

- Freezing can be indefinite or fixed-term



Dataset freezing is a compromise between support for dynamic consent and dataset completeness. After the deadline for consent decisions is passed, the dataset (or a copy of it) is frozen for analysis, either for a fixed period or indefinitely. This yields a good-enough solution for negligibly consent-biased secondary data analysis.



Analysing your dynamic consent solution

See Prictor et al.'s (2019) **evaluation and reporting framework for dynamic consent** for suggestions on quantifying and analysing the outcomes of dynamic consent upon your datasets and participant groups.

Prictor et al.'s framework for the evaluation of and reporting on dynamic consent is a useful resource for quantifying the effects of dynamic consent upon your own datasets and participant groups.