Open Algorithms for Identity Federation

Thomas Hardjono
MIT Connection Science & Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
Email: hardjono@mit.edu

Sandy Pentland
MIT Media Lab &
MIT Connection Science & Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
Email: sandy@media.mit.edu

***Abstract*—In this paper we argue that the current attribute sharing approach adopted by the identity management industry for consumer identity provides only limited information about an individual, and therefore is of limited value to the participants of the ecosystem. Instead, we propose a new approach based on *algorithm sharing* as part of a broader *open algorithms* paradigm. Rather than sharing attributes (fixed assertions), the participants of the ecosystem should obtain richer insights for through the collective sharing of algorithms that have been vetted to be safe from a privacy perspective.**

I. Introduction

Today there is a resurgence in interest in the broad area of "identity" – driven amongst others by the emergence of blockchain technology which have been viewed by many as providing a future direction for decentralizing control, including control over one's identity. Often the problem of personal data privacy is equated with decentralization of control, even though it remains unclear how one is connected to the other. The problem of digital identities itself is not new, but rather a field that has been evolving since the mid-1990s with the advent of the DNS infrastructure, PKI and digital certificates.

It is not too far fetched to state that the "digital identity problem" of today (and for the past two decades) is more accurately a "problem of data" because much of the actual transaction value (based on an identifier) depends on the quality of data and information gleaned from the data. Thus the value of an identity system lies in the quality of information accessible regarding a subject or object managed by that identity system. Data itself gains more value when combined with other data of the same kind or data from different domains [1]. As such, we see the problem of "digital identity" are really being a problem of sharing information cross-domain (cross-verticals) while preserving privacy, and based on consent from the subject (user).

Today *Identity and Access Management* (IAM) represents a core component of the Internet infrastructure. IAM infrastructures are an enabler which allows organizations to achieve its goals. Enterprise-IAM (E-IAM) is a mature product category [2] and E-IAM infrastructures are already well integrated into other enterprise infrastructure services, such as directory services for managing employees and corporate assets. The primary goal of E-IAM systems is to authenticate and identify persons (e.g. employees) and devices as they enter the enterprise boundary, and to provide access control driven by corporate policies. In the case of Consumer-IAM (C-IAM), the primary goal is to reduce friction between transacting entities, such as between the consumer and an online service provider (e.g. merchant).

One of the key findings of the 2011 World Economic Forum (WEF) report on personal data [3], [4] is that the current ecosystems that access and use personal data is fragmented and inefficient. For many participants, the risks and liabilities exceed the economic returns. Furthermore, personal privacy concerns are inadequately addressed. Current technologies and laws fall short of providing the legal and technical infrastructure needed to support a well-functioning digital economy. Although today it might be evident that personal data "will emerge as a new asset class touching all aspects of society" the rapid rate of technological change and commercialization in using personal data is undermining end user confidence and trust.

A more recent development is the stated intent on the part of some traditional institutions possessing customer data in a given vertical to seek new sources of revenue from that consumer data, departing from the traditional roles of those institutions. In the financial industry the Bitcoin system [5] – albeit limited in transaction type and throughput – has provided a real working example of a peer-to-peer value-transfer system that does not depend on an intermediary entity (i.e. a bank). This in-turn has caused a number of large traditional financial institutions to question their business survival (and even the necessity of their existence) in the coming decades. A number of these financial institutions currently have begun to re-cast their role into "providers of identity" even though strictly speaking these institutions are not identity providers [6]–[8]. The same trend can also be seen in other verticals, such as within the telecommunications industry (i.e. consumer call data records).

However, as much as there is new interest in the digital identity problem, the existing challenges with data-sharing remains unsolved: (i) data remains siloed within the organizational boundaries, (ii) data maybe of limited type/domain only (e.g. financial data), and (iii) sharing of raw data with parties outside the organization remains unattainable, either due to regulatory constraints or due to business risk exposures and liabilities. The upcoming enactment of the new EU General Data Protection Regulations (GDPR) will place various other additional constraints on data (e.g. expanded definition of

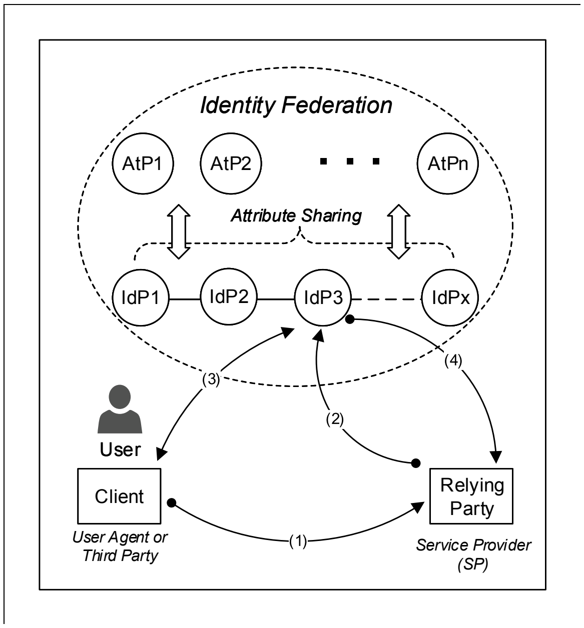

Fig. 1. Identity Federation and Attribute Sharing

personal data to include genetic and biometric data)

It is with this backdrop that we propose the *open algorithms* approach, using the identity-related use-case in the context of Consumer-IAM as a way to illustrate the usefulness of the paradigm in addressing the needs for information sharing and privacy. The open algorithms concepts and principles will be discussed further in Section III.

The remainder of this paper is organized as follows. In the next section we provide a brief history and overview of identity management and federation, providing some definitions of the entities and their functions. Readers familiar with the IAM industry and the current identity federation landscape are encouraged to skip this section.

In Section III we explain in further detail the concepts and principles underlying the open algorithms paradigm.

We use this paradigm in the context of identity federation in Section IV with the goal of identifying new architectures and interaction models for entities in a federation ecosystem that allows them to obtain better information about users (subjects) while providing better controls over privacy.

Section V briefly discusses the true difficulty today in developing an ecosystem for sharing of information, namely that of legal trust frameworks for the sharing of information. We argue that the open algorithms paradigm offers a way forward for ecosystem entities (e.g. traditional institutions) to develop new revenue sources based on their data, governed by relevant legal agreements and contracts that form the basis for a future legal trust framework.

II. Identity Federation and Attribute Sharing

Historically, notion of an *identity provider* entity emerged starting in the late 1990s in response to the growing need to aid users in accessing services online. Prior to the maturing of the identity provider services, a user would typically create an account and credentials at every new service provider (e.g. online merchant). This cumbersome approach – which is still in practice today – has led to a proliferation of accounts and duplication of the same user data across many service providers.

In 2001 an alliance of over 150 companies and organizations formed an industry consortium called the *Liberty Alliance Project* The main goal of this consortium was to "establish standards, guidelines and best practices for identity management" [9]. Several significant outcomes for the IAM industry resulted from this consortium, two among which were: (a) the standardization of the *Security Assertions Markup Language* (SAML2.0) [10], and (b) the creation of the widely used open source SAML2.0 server implementation called *Shibboleth* [11].

Today the industry category of "Identity Provider" has matured, and identity providers are touted as the solution to the "accounts proliferation" problem (also call the "NASCAR logos problem"). In this mediated-authentication model the user would first create an account at one of the Identity Providers (IdP) of his or her choice. When the user seeks to access a resource or service offered by a merchant (called the Relying Party or RP) the user would be redirected by the RP to the Identity Provider who will then challenge the user for her or her credentials (i.e. perform authentication). After successful authentication, the Identity Provider would redirect the user back to the calling RP. This flow pattern is known as the *Web Single Sign-On* (Web-SSO) pattern. The same pattern is followed in newer identity systems based on the OAuth2.0 framework [12], notably the now popular OpenID-Connect 1.0 protocol [13]. This flow pattern is summarized in Figure 1.

In Figure 1 Step (1) the user requests access to services or resources at the Relying Party (RP). The RP redirects the user in Step (2) to one of the Identity Providers (IdP) in the federation, who challenges the user in a front-channel authentication event (Step (3)). Upon successful user authentication, the IdP redirects the user (e.g. via call-back URIs) to the original RP in Step (4), accompanying it with one or more signed assertions (e.g. SAML2.0 assertions or signed JWT).

The emergence of federation was driven by the need for identity providers to achieve scaling of services in the face of an increasing number of Relying Parties (e.g. online merchants). The idea is straightforward: when a Relying Party seeks to ascertain the personal information regarding a user, the RP can inquiry into the federation, starting with the IdP with whom the RP has a service agreement. When that IdP is unable to furnish the necessary information about the user it can then inquire to other IdP members of the federation or from an attribute provider.

More formally, the primary goal of a *federation* among a group of *identity providers* (IdP) is to share "attribute" information (assertions) regarding a user [14]:

- An *Identity Provider* is the entity (within a given identity system) which is responsible for the identification of persons, legal entities, devices, and/or digital objects, the issuance of corresponding identity credentials, and the maintenance and management of such identity information for Subjects.

- An *attribute* is a specific category of identifying information about a subject or user. Examples for users include name, address, age, gender, title, salary, health, net worth, driver's license number, Social Security number, etc.

- An *Attribute Provider* (AtP) is a third party trusted as an authoritative source of information and responsible for the processes associated with establishing and maintaining identity attributes. An Attribute Provider asserts trusted, validated attribute claims in response to attribute requests from Identity Providers and Relying Parties. Examples of Attribute Providers include a government title registry, a national credit bureau, or a commercial marketing database.

- An *Identity Federation* is the set of technology, standards, policies, and processes that allow an organization to trust digital identities, identity attributes, and credentials created and issued by another organization. A federated identity system allows the sharing of identity credentials issued, and identity information asserted, by one or more Identity Providers with multiple Relying Parties (RP).

The *attribute provider* category of entities became recognized through the need for confirmed facts or assertions regarding a user as the subject. The separation between the identity provider and the attribute provider is useful one and necessary because not all attribute providers are (or sought to be) identity providers that performed mediated-authentication.

Although the federated identity model using the SSO flow pattern remains the predominant model today for the consumer space, there are a number of limitations with the model – both from the consumer privacy perspective and from the providers business model perspective:

- *Identity management as an adjunct service*: Most large scale consumer-facing identity services today are a side function to another more dominant service (e.g. free email service, social media platform). Very little potential revenue is available from the pure mediated-authentication model for the identity provider entity.

- *Limited data*: The attribute provider model currently used in federations delivers only a limited amount of information regarding a user. The scope of attributes permitted to be exchanged among federation members are usually defined in the legal service contracts which are referred to as *legal trust frameworks* (see Section V).

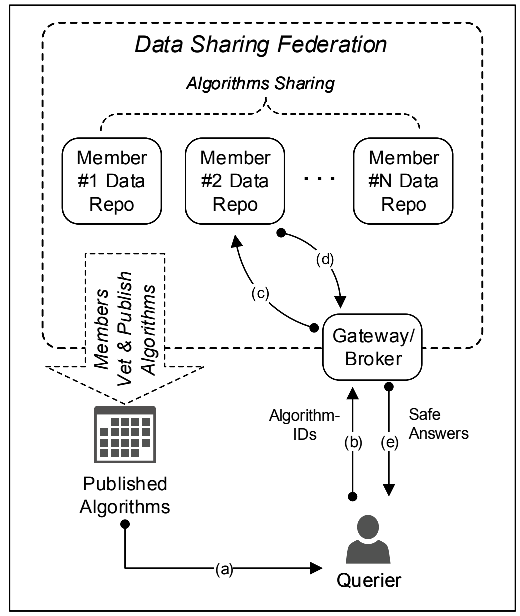

Fig. 2. Overview of OPAL

An example of the list of permitted attributes or claims can be found in the OpenID-Connect 1.0 core specifications (Section 5.1 on Standard Claims in [13]) for ecosystems deploying the OpenID-Connect flows.

- *Limited user control and consent*: Over the past years several efforts have sought to address the issue of user control and consent (e.g. [15], [16]). The idea here is that not only should the user explicitly consent to his or her attributes being shared, the underlying identity system should also ensure that only minimal disclosure is performed for a constrained use among justifiable parties [17], [18]. However, in reality there is little incentive on the part of the major incumbents today (e.g. free email providers; social network platforms) to pay attention to these fundamental aspects of user control and consent.

III. Open Algorithms: Key Concepts & Principles

As previously mentioned, data is crucial for the proper functioning of the digital society. Data is core to the day-to-day running of communities, governments and businesses. Furthermore, data increases in value when combined cross domains and verticals, and previously unconsidered insights can be obtained when data is combined. However, the current reality is that cross-organizational data sharing is prohibitive both from the regulatory perspective and from the business risk perspective.

The concept for *Open Algorithms* (OPAL) emerged from several MIT research projects over the past decade which focused on the future data-driven society [4]. It was increasingly becoming apparent that an individual's privacy could be affected through correlation of just small amounts of data [19], [20] and that models for user-centric personal data stores [21] needed to be developed further and championed in industry.

The principles underlying the MIT OPAL Project seeks to address the apparent conflict between (i) the need for the digital society to have access to personal data, and (ii) the need to address personal privacy concerns in the digital society in a satisfactory manner. The OPAL model – albeit simple – provides a useful framework within which industry can begin finding solutions for these information sharing constraints.

The following are the key concepts and principles underlying the open algorithms paradigm [22]:

- *Moving the algorithm to the data*: Instead of sending raw data to the location of data-processing (e.g. centralized analytics infrastructure), it is the algorithm that should be sent to the location of the data (i.e. data-repository) and be processed there.

- *Raw data must never leave its repository*: Raw data sets must never be exported, and must always be under the control of its owner or the owner of the data repository.

- *Vetted algorithms*: Algorithms must be studied and vetted by domain-experts to be "safe" from bias, discrimination, privacy violations and other unintended consequences.

  Vetted algorithms should be collectively agreed upon by members of the information sharing ecosystem or federation, governed by a legal trust framework for algorithm-sharing.

  Note that the vetting of an algorithm does not guarantee the quality of the output from executing the algorithm, (which is more a function of the quality of the input data).

- *Provide only safe answers*: When executing an algorithm on a data-set, the data-repository must always provide responses that are deemed "safe" from a privacy perspective. Responses must not release or leak personally identifying information (PII) without the consent of the user (subject). This may imply that a data repository return *only aggregate answers*.

  A data repository always has the option to decline responding if the resulting answers are deemed to violate privacy. This option to decline must be stated within the underlying legal trust framework.

- *Consent for algorithm execution*: Data repositories that hold user (subject) data must obtain explicit user consent when the user's data is to be included in a given algorithm execution. Vetted algorithms should be made available and understandable to users (e.g. "algorithm X is to be executed on data-set Y containing your private information, with the expected result being Z"), and consent should be unambiguous.

- *Trust Networks for Information Sharing*: The ecosystem must be governed by a legal trust framework for the sharing of algorithms (see Section V).

- *Decentralized Data*: New architectures based on Peer-to-Peer (P2P) networks should be employed as the basis for new decentralized data architectures [23]. Since data is a valuable asset, a proper design of a decentralized architecture should increase the resiliency of the overall data against attacks (e.g. theft, poisoning, manipulations, etc). New secure infrastructure models, such as MIT Enigma [24], should be further developed to address the resiliency issues.

  These decentralized data architectures should also incorporate the notion of personal data stores [21] as a legitimate data repository end-point in deploying the open algorithms approach. Protocol standards for managing access to personal repositories are in development [15], and further research is underway to use existing standards in a decentralized fashion (e.g. [25]).

In is important to note that the term "algorithm" has been intentionally left undefined. This is because we believe that each federation or network that implements the open algorithms principles must do so in their own data context, defining it within their legal trust framework and contracts. They must not only define the technical syntax but also the semantic intent of "algorithm". This provides flexibility and greater scope for each OPAL-based information sharing federation to define the nature or type of data within their federation and the class of algorithms to be used in that community.

From an engineering perspective, the OPAL model is fairly simple to understand (see Figure 2). A querier entity (e.g. person or organization) that wishes to obtain information from the information sharing federation selects one or more vetted algorithm (queries) from the set of published vetted-algorithms (each digitally signed by the federation to provide source authenticity). This is shown as step (a) in Figure 2.

After selecting one or more algorithms (or its algorithm-ID), the querier entity sends the algorithm (digitally signed) to the gateway or broker entity that mediates access to the information sharing federation. Optionally the querier entity may attach payment in order to remunerate the destination data repositories (step (b) in Figure 2)

The gateway entity forwards the algorithms to each relevant member data repository (step (c) in Figure 2), and obtains a safe response from the corresponding data repository (step (d) in Figure 2). Note that a member data repository may always decline (e.g. if it deems privacy to be affected; or if payment is insufficient). A data repository may place additional filters on the outgoing response. The getaway relays the results back to the querier entity (step (e) in Figure 2), possibly with additional filtering.

Currently, small test-bed deployments of the basic open algorithms concept are underway for specific and narrow data-domains [26], with other planned [27].

IV. OPAL for Identity Federation

In this section we offer a perspective on how the open algorithms paradigm can be used by an identity federation ecosystem to enable richer information sharing among participants in the ecosystem. The key idea here is for industry to move from an attributes-based unidirectional model to a richer algorithms-based interaction.

- *Algorithms instead of attributes*: Rather than attribute providers delivering static attributes (e.g. "Joe is over 18") to the relying party, allow instead the relying party to choose one or more vetted algorithms from a given data domain.

  The result from executing a chosen algorithm is then returned to the relying party in a signed response. The response can also include various metadata embellishments, such as the duration of the validity of the response (e.g. for dynamically changing data sets), identification of the data-sets used, consent-receipts, timestamps, and so on.

- *Persistent logging and audit*: All requests and responses must be logged, together with strong audit capabilities. Emerging technologies such as the blockchain systems could be expanded to allow for an immutable forward-hash chains to be created, effectively providing a digital notary service that can be used for a post-event auditing (e.g. by an external party or auditor).

  Logging and audit is not only crucial for operations and for regulatory compliance, it is also necessary in order to obtain buy-in and consent from individuals whose data is present within a larger data-set on which an algorithm is being executed. Users as stakeholder in the ecosystem must be able to see the audit trail of their data being used in computations.

- *Apply correct pricing model for algorithms and data*: For each algorithm and the data to which the algorithm applies, a correct pricing structure needs to be developed by the members of the federation. This is not only to remunerate the data repositories for managing the data-sets and for executing the algorithm (i.e CPU usage), but also to encourage data owners to develop new business models based on the OPAL paradigm.

  Pricing information could published as part of the vetted-algorithms declaration (e.g. as metadata), offering different tiers of pricing for different sizes of data sets. For example, the price for obtaining a credit-rating for a subject may perhaps be different from obtaining information based on a combined data set of the subject's GPS locations and credit-card spending behavior [4].

- *Remunerate the user and owner of data*: A correct alignment of incentives must be found for all stakeholders in the identity federation ecosystem. Users should see some return on the use of their data, even if it is tiny amounts (e.g. in the pennies or sub-pennies).

  The point here is that users will contribute more personal data if they are active participants in the ecosystem and understand legal protections offered by the trust frameworks that govern their personal data.

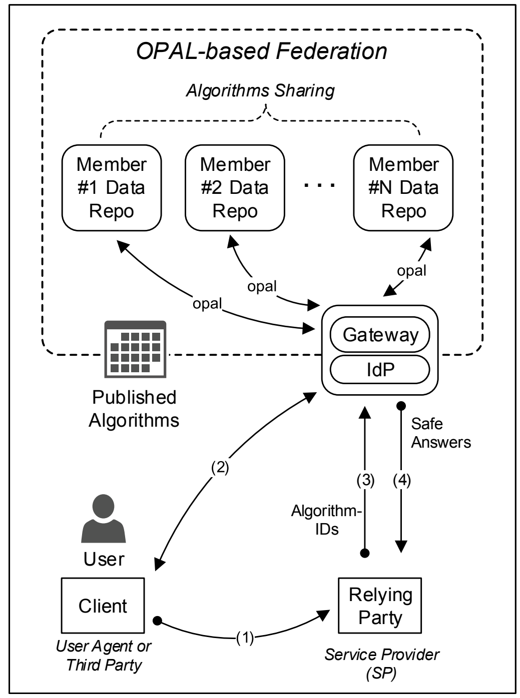

Fig. 3. OPAL-based Federation

The use of the OPAL paradigm for information sharing within an identity federation is summarized in Figure 3, using the the traditional SSO flows. Figure 3 shows an alternate flow pattern, which essentially replaces the attribute providers (AtP) in Figure 1 with the OPAL-based federation of Figure 2.

In Step (1) of Figure 3 when the user (subject) seeks access to resources or services at the relying party, the relying party has the option to request the execution of one or more of the vetted-algorithms (Step (3)) as part of the user-authentication event.

Note that Figure 3 collapses the Identity Provider entity into the Gateway entity, although the understanding is that different deployment architectures and topologies can be envisaged. The Gateway/Identity Provider entity now has the task of authenticating the user (as in the traditional SSO model) and also of delivering (possibly filtering) safe responses to the relying party (Step (4)).

The convergence of the Gateway entity with the Identity Provider entity in Figure 3 somewhat reflects the current sentiment on the part of a number of data-rich traditional institutions who are considering their role as an "identity provider" (e.g. see [8]). Here the Gateway/Identity Provider entity is under the control of the federation as a whole. A given federation may in fact deploy multiple of these Gateway/IdP entities (or even outsource the role to one or more federation members).

V. Trust Framework for OPAL Federation

Today trust frameworks for identity management and federation in the US is based on three types or "layers" of law. The foundational layer is the general commercial law that consists of legal rules that are applicable to identity management systems and transactions. This general commercial law was not created specifically for identity management, but instead are public law written and enacted by governments which applies to all identity systems, its participants – and thus enforceable in courts.

The second "layer" consist of general identity system law. Such law is written to govern all identity systems within its scope. The intent would be to address the various issues related to the operations of the identity systems. The recognition of the need for law at this layer is new, perhaps reflecting the slow pace of development in the legal arena as compared to the technology space. An example of this is the Virginia Electronic Identity Management Act [28], which was enacted recently.

The third layer is the set of applicable legal rules and system-specific rules (i.e. specific to the identity system in question). The term "trust framework" is often used to refer to these system rules that have been adopted by the community. A trust framework is needed for a group of entities to govern their collective behavior, regardless if the identity system is operated by the government or the private sector. In the case of a private sector identity system, the governing body consisting of the participants in the system typically drafts rules that take the form of *contracts-based rules*, based on private law.

Examples of trust framework for identity federation today are FICAM for federal employees [29], SAFE-BioPharma Association [30] for the biopharmaceutical industry, and the OpenID-Exchange (OIX) [31] for federation based on the OpenID-Connect 1.0 model.

In order for an OPAL-based information sharing federation to be developed, it must use and expand the current existing legal trust frameworks for identity systems. This is because the overall goal is for entities to obtain richer information regarding a user (subject), and as such it must be bound to the specific identity system rules. In other words, a new set of third layer legal rules and system-specific rules must be devised that must clearly articulate the required combination of technical standards and systems, business processes and procedures, and legal rules that, taken together, establish a trustworthy system [14] for information sharing based on the OPAL model. It is here that system-specific rules regarding the "amount of private information released" must be created by the federation community, involving all stakeholders including the users (subjects).

Taking the parallel of an identity system, an OPAL-based information sharing system must address the following:

- Verifying the correct matching between an identity (connected to a human, legal entity, device, or digital object) and the set of data in a repository pertaining to that identity;
- Providing the correct result from an OPAL-based computation to the party that requires it to complete a transaction;
- Maintaining and protecting the private data within repositories over its lifecycle.
- Defining the legal framework that defines the rights and responsibilities of the parties, allocates risk, and provides a basis for enforcement.

Similar to – and building upon – an identity system operating rules, new additional operating rules need to be created for an OPAL-based information sharing system. There are two (2) components to this. The first is the business and technical operational rules and specifications necessary to make the OPAL-based system functional and trustworthy. The second is the contract-based legal rules that (in addition to applicable laws and regulations) define the rights and legal obligations of the parties specific to the OPAL-based system and facilitate enforcement where necessary.

As the current work is intended to focus on the concepts and principles of open algorithms and their application to information sharing in the identity context, these two aspects will be subject for future work.

VI. Conclusions

In this paper we have described the concepts and principles underlying the open algorithms (OPAL) paradigm. We believe that this new paradigm offers a way forward for the data-driven society and industry more broadly to begin addressing the core issues around data and information-sharing. Some of these challenges includes siloed data, the limited type/domain of data, and prohibitive nature of cross-organization sharing of data.

We have used the identity-related use-case in the context of Consumer-IAM as a way to illustrate the usefulness of the paradigm in addressing the needs for information sharing and privacy.

In particular, the sharing of algorithms instead of fixed attributes within an identity federation offers the participants of the federation ecosystem with better access to broader information. This in turn allows them to develop new revenue sources based on their data, governed by relevant legal agreements and contracts that form the basis for a information sharing legal trust framework.

Finally, a new set of legal rules and system-specific rules must be devised that must clearly articulate the required combination of technical standards and systems, business processes and procedures, and legal rules that, taken together, establish a trustworthy system for information sharing in a federation based on the OPAL model.

References

[1] A. Pentland, T. Reid, and T. Heibeck, "Big Data and Health - Revolutionizing Medicine and Public Health: Report of the Big Data and Health Working Group 2013," World Innovation Summit for Health, Qatar Foundation., Tech. Rep., December 2013, http://www.wish-qatar.org/app/media/382.

[2] Gartner, "2017 Planning Guide for Identity and Access Management," Gartner Inc., Report, October 2016.

[3] World Economic Forum, "Personal Data: The Emergence of a New Asset Class," 2011, http://www.weforum.org/reports/ personal-data-emergence-new-asset-class.

[4] A. Pentland, *Social Physics: How Social Networks Can Make Us Smarter*. Penguin Books, 2015.

[5] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System." [Online]. Available: https://bitcoin.org/bitcoin.pdf

[6] UBS, "Digital Identity: a UBS Group Innovation White Paper," UBS, Tech. Rep., December 2016.

[7] W. Suberg, "Barclays will help you prove your identity to UK Govt," March 2016, [Online; posted 21-March-2016]. [Online]. Available: https://news.bitcoin.com/barclays-will-help-prove-identity-uk-govt/

[8] World Economic Forum, "A Blueprint for Digital Identity: the Role of Financial Institutions in Building Digital Identity," World Economic Forum, Tech. Rep., 2016, http://www3.weforum.org/docs/ WEF_A_Blueprint_for _Digital_Identity.pdf.

[9] "Liberty Alliance," https://en.wikipedia.org/wiki/Liberty_Alliance, accessed: 2017-05-29.

[10] OASIS, "Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0," March 205, available on http://docs.oasisopen.org/security/ saml/v2.0/ saml-core-2.0-os.pdf.

[11] R. L. Morgan, S. Cantor, S. Carmody, W. Hoehn, and K. Klingenstein, "Federated Security: The Shibboleth Approach," *EDUCAUSE Quarterly*, vol. 27, no. 4, p. 1217, 2004.

[12] D. Hardt, "The OAuth 2.0 Authorization Framework," RFC 6749 (Proposed Standard), Internet Engineering Task Force, Oct. 2012. [Online]. Available: http://www.ietf.org/rfc/rfc6749.txt

[13] N. Sakimura, J. Bradley, M. Jones, B. de Medeiros, and C. Mortimore, "OpenID Connect Core 1.0," OpenID Foundation, Technical Specification v1.0 – Errata Set 1, November 2014, http://openid.net/specs/openid-connect-core-1_0.html.

[14] American Bar Association, "An Overview of Identity Management: Submission for UNCITRAL Commission 45th Session," ABA Identity Management Legal Task Force, May 2012, available on http://meetings.abanet.org/ webupload/commupload/ CL320041/relatedresources/ ABA-Submission-to-UNCITRAL.pdf.

[15] T. Hardjono, E. Maler, M. Machulak, and D. Catalano, "User-Managed Access (UMA) Profile of OAuth2.0 – Specification Version 1.0," April 2015, https://docs.kantarainitiative.org/uma/rec-uma-core.html.

[16] M. Lizar and D. Turner, "Consent Receipt Specification Version 1.0," March 2017, https://kantarainitiative.org/confluence/display/infosharing/Home.

[17] K. Cameron, "The Laws of Identity," 2004. [Online]. Available: http://www.identityblog.com/stories/2004/12/09/thelaws.html

[18] A. Cavoukian, "7 Laws of Identity - the Case for Privacy-Embedded Laws of Identity in the Digital Age," Office of the Information and Privacy Commissioner of Ontario, Canada, Tech. Rep., October 2006. [Online]. Available: http://www.ipc.on.ca/index.asp?navid=46&fid1=470

[19] Y. A. de Montjoye, J. Quoidbach, F. Robic, and A. Pentland, "Predicting personality using novel mobile phone-based metrics," in *Social Computing, Behavioral-Cultural Modeling and Prediction (LCNS Vol. 7812)*. Springer, 2013, pp. 48–55.

[20] A. Pentland, "Saving Big Data from Itself," *Scientific American*, pp. 65–68, August 2014.

[21] Y. A. de Montjoye, E. Shmueli, S. Wang, and A. Pentland, "openPDS: Protecting the Privacy of Metadata through SafeAnswers," *PLoS ONE 9(7)*, pp. 13–18, July 2014, https://doi.org/10.1371/journal.pone.0098790.

[22] A. Pentland, D. Shrier, T. Hardjono, and I. Wladawsky-Berger, "Towards an Internet of Trusted Data: Input to the Whitehouse Commission on Enhancing National Cybersecurity," in *Trust::Data - A New Framework for Identity and Data Sharing*, T. Hardjono, A. Pentland, and D. Shrier, Eds. Visionary Future, 2016, pp. 21–49.

[23] P. De Filippi and S. McCarthy, "Cloud Computing: Centralization and Data Sovereignty," *European Journal of Law and Technology*, vol. 3, no. 2, 2012, available at SSRN: https://ssrn.com/abstract=2167372.

[24] G. Zyskind, O. Nathan, and A. Pentland, "Decentralizing privacy: Using blockchain to protect personal data," in *Proceedings of the 2015 IEEE Security and Privacy Workshops*. IEEE, 2015.

[25] T. Hardjono, "Decentralized Service Architecture for OAuth2.0," Internet Engineering Task Force, draft-hardjono-oauth-decentralized-00, February 2017, https://tools.ietf.org/html/draft-hardjono-oauth-decentralized-00.

[26] R. Frey, T. Hardjono, C. Smith, K. Erhardt, and A. Pentland, "Secure sharing of geospatial wildlife data," in *GeoRich '17: Proceedings of the Fourth International ACM Workshop on Managing and Mining Enriched Geo-Spatial Data*, May 2017.

[27] DataPop, "Data-Pop Alliance," 2017, http://datapopalliance.org.

[28] State of Virginia, "Virginia Electronic Identity Management Act," March 2015, VA Code 2.2-436 2.2-437; VA Code 59.1-550 59.1-555. [Online]. Available: https://lis.virginia.gov/cgi-bin/legp604.exe?151+ful+CHAP0483

[29] US General Services Administration, "U.S. Federal Identity, Credential and Access Management (FICAM) Program," 2013, http://info.idmanagement.gov.

[30] SAFE-BioPharma Association, "Trust Framework Provider Services," 2014, http://www.safe-biopharma.org/SAFE_Trust_Framework.htm.

[31] OIX, "Open Identity Exchange," 2013, http://openidentityexchange.org.