
---

title: "The Researcher Passport: A Digital Credential to Improve Restricted Data Access"
author:
- name: Margaret C. Levenstein
  affiliation: Inter-University Consortium for Political and Social Research
  email: julia.lane@nyu.edu

---
date:
2019
cross-
refYaml:
"./pandoc-
crossref-
settings.yaml"
...
Problem
State-
ment
==========

In the last two decades, research data repositories joined libraries and archives in promoting access to and use of data by making it available directly on the internet. While this led to a radical expansion in the avail-

The access challenge is magnified when researchers require data sets from multiple data repositories, each of which has its own access requirements. There is often inconsistency between data custodians in the in-for-

To better understand this challenge, the Interuniversity Consortium for Political and Social Research (ICPSR) conducted a yearlong study of 23 restricted data repositories around the world. The environmental scan and interviews

---

- User verification - Training requirements
- Terminology

For each of these areas, no two repositories used the same processes or interpretations of legal and regulatory requirements. Data users who need data from multiple sources must negotiate the data access processes, each

Repositories require verification of different information about users as they review data access requests. Some repositories require in-depth background checks, while others simply validate the application information

To solve these challenges and promote a research community with a shared understanding of the needs of researchers and repositories, ICPSR is developing the Researcher Passport. This digital identifier captures verified information

Researcher
pass-
port
==========

The foundation of the Researcher Passport is the level of trust that data repositories place on their users: the more the repository trusts the user to use data in accordance with the data use agreements, based on data use experi-

The first time that a user requires restricted data from a participating institution, the user submits an application to ICPSR for a Researcher Passport. ICPSR conducts the in-depth identity verification process and is-

This Passport does not replace the authority of repositories to implement their own additional requirements for information about users or the project proposal, nor does it mandate that any user who is issued a Pass-

Within the Researcher Passport system, the Visa is what authorizes access to the data in the repository. Over time, as users access data, the Passport's visas also serve as the record of prior data use that some repositories already

ICPSR is developing this system within the ICPSR infrastructure. In November 2018 ICPSR will launch phase 1 of the Researcher Passport as part of its ICPSR MyData user accounts.[a] The system will be piloted at three restricted data repos-

Open Badges

An innovative component of ICPSR's Researcher Passport is the use of Open Badges[1] as the mechanism for embedding user characteristics—identity information, academic qualifications, professional experience, training completion—within the Passport. The Passport is the digital container for these badges. As part of the initial Passport issuing process, select badges will be identified as "verified." Verified badges will include the components of data access requests that were identified as most important for developing trust in potential users (Levenstein, Tyler, and Davidson Bleckman 2018). Badges can be used by repositories in defining their access requirements. If an applicant does not meet the badge requirements for specific data sets or access methods, then further review is required to determine which access method would be appropriate, or if access should be granted at all. Beginning in 2019, ICPSR and the University of Michigan School of Information will build and integrate ICPSR badges into the Passport.

Community Norms

In addition to the benefits to repositories and data users in terms of more efficient data access request evaluation and user authorization, the Researcher Passport project seeks to establish shared norms and expectations around the access and use of restricted data. The Researcher Passport is most useful if it is widely accepted by the repository community; the more it is accepted and implemented, the more standardized the expectations will be. Our analysis of existing repository practices identified three specific areas where this standardization is needed and for which the Researcher Passport provides that standardization:

- User evaluation criteria
- Data evaluation criteria
- Data management training

User evaluation criteria

First, we propose a point-based user evaluation process to determine the Passport level eligibility. Passport applications are reviewed and points assigned based on the highest academic degree earned, the professional position (with separate attributes for non-academic researchers, e.g., media, non-profit, for-profit, and government employees), possession of a government-issued clearance, history of federal grants, publication history, and restricted data use experience (Appendix B). Badges will be indicate these attributes, as well as additional attributes relevant for access to specific data sets but not necessary for the Passport level determination (e.g., nationality).

Data evaluation criteria

Similarly, we propose standardizing the data security level assignment process. Currently, as discussed above and in Table 1, there are a wide variety of terms and interpretations of data security levels, based on repository naming conventions and the specific needs of the data set in question. We recommend

[1] Open Badges: https://openbadges.org/

a spectrum of Low, Moderate, High, and Highest, each of which has a point evaluation range comparable to the user evaluation criteria. Data sets will be reviewed for different characteristics including sensitivity, disclosure risk, and legal or statutory limitations; total points map to a data security level. An example of this metric can be viewed in Appendix C. This metric provides flexibility for repositories to add additional evaluation criteria (e.g., detailed geography, additional legal requirements); the data security score range is adjusted accordingly. This will build a common understanding of the meaning of data security levels, even as repositories maintain use of additional requirements for access to particular datasets.

Training

ICPSR evaluated the training requirements for access to restricted data at the 23 repositories in our study. Most repository training requirements refer to Human Subjects Research and Responsible Conduct of Research, in accordance with IRB requirements. We also evaluated the content of trainings required by repositories; we found only two that included content specifically focused on restricted data. In both cases, the training modules on restricted data were developed for use only at their specific institutions. Even in these training modules, data protection and data handling topics are discussed only briefly. No training program exists that covers all topics to which repositories said they wanted their restricted data users exposed. As part of the development of the Researcher Passport, ICPSR will develop training modules to meet the requirements and expectations of repositories. We will also continue to try to identify other training programs that meet these requirements. Completion of training will be identified through appropriate Badges on the Passport.

CONCLUSION

We live in a data-intensive world. We create data as we sleep and walk and eat, with every purchase we make, every email we send, every camera we stroll by. These data are valuable for research and evidence-building. Analyses of such data are used to inform more science and more policy making than ever. But it is also easier than ever to identify individuals, or use data inappropriately or inconsistently with the expectations of those being measured.

Given the unprecedented availability of digital data and the continuing need to interrogate "old" data, secure and efficient mediation of data access is a priority for the research community. ICPSR's development of the Researcher Passport, a digital researcher credential based on shared norms about users and about data, represents our contribution to the challenge of balancing access and privacy. The process begins with developing a shared understanding of how data are classified and how users are evaluated and authorized to access that data, and then turns to the design and implementation of a system that operationalizes the trust imbued in those users in a digital identifier.

For further details about the researcher credentialing research project, please consult our May 2018 report to the Alfred P. Sloan Foundation, available at https://deepblue.lib.umich.edu/handle/2027.42/143808.

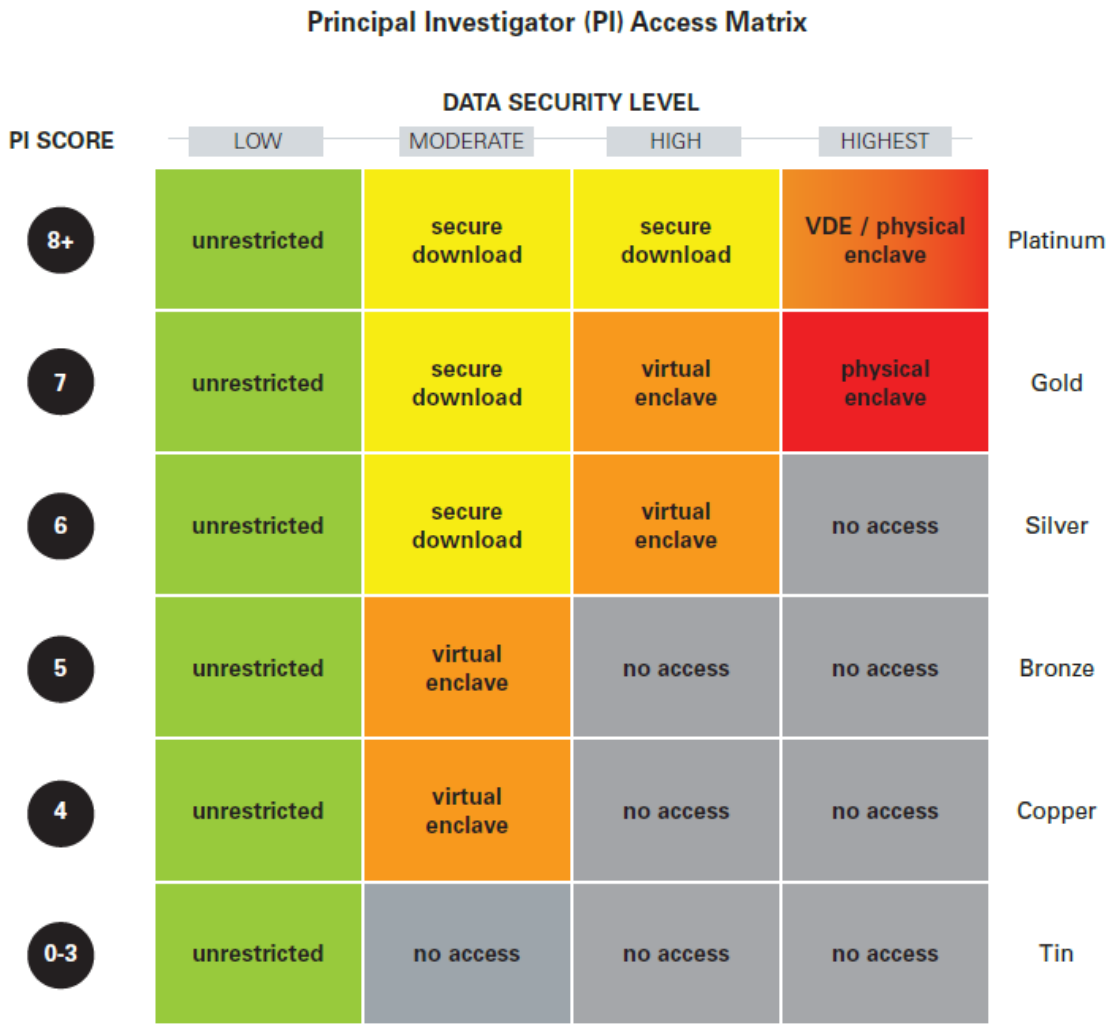

Figure 1: Researcher Passport Access Matrix[2]

APPENDIX A

APPENDIX B

APPENDIX C

References

Levenstein, Margaret, Allison R. B. Tyler, and Johanna Davidson Bleckman. 2018. *The Researcher Passport: Improving Data Access and Confidentiality Protection.* ICPSR White Paper Series No. 1. Ann

[2]Levenstein, Tyler, & Davidson Bleckman, 2018, p. 23.
[3]Levenstein, Tyler, & Davidson Bleckman, 2018, p. 21
[4]Levenstein, Tyler, & Davidson Bleckman, 2018, p. 22.

USER ATTRIBUTES	POINTS ATTRIBUTED
Highest degree earned	
Doctoral/terminal degree	3
Graduate degree (non-terminal)	2
Undergraduate	1
No degree	0
Professional Position (choose one of the following two options)*	
Academic faculty/staff: Highest institutional appointment/affiliation	
Full/Associate professor	3
Assistant professor	2
Student	1
Research staff	1
Non-profit, for-profit, government, or media staff: Years of relevant experience	
5+	3
3-4	2
0-2	1
Other	
Recognized Federal clearances	4
Current (2 pts) or recent (1 pt) Federal grant	2/1
Research publications (1 or more publications)	2
Restricted data use experience (1 or more projects)	2
Potential dataset- or repository-specific user requirements	
Country- or region-specific citizenship or residency status	specify
Affiliation with Carnegie-classified academic institution	yes/no
Badges earned and verified	
Trainings	
Data security — Levels I-III	specify
Research conduct — Levels I-III	specify
Other	specify
Specific expertise	
Restricted qualitative data use	specify
Other	specify
Contributions — data stewardship	
History of data sharing	citation/DOI
History of metadata enhancement	citation/DOI
History of code/syntax sharing	citation/DOI
Confirmed research misconduct (unintentional procedural violations and/or intentional data disclosure or misuse)	yes/no

Figure 2: Proposed Researcher Passport User Attributes Evaluation Metric[3]

DATA CHARACTERISTICS	POINTS ATTRIBUTED
Sensitivity level	**If yes, then add...**
protected population	+ 3
proprietary data	+ 4 to 6
potentially harmful personal information	+ 4
Disclosure risk level	
sample size	+ 1 to 4
geographic region size	+ 1 to 4
rare sample attributes	+ 1 to 4
link to public data	+ 3
Legal or statutory limitations	
HIPAA	+ 6
FERPA	+ 6
other legislated restrictions	+ 3 to 6
Data security score *(after totalling above)*	**Range**
Low	0-3
Moderate	4-5
High	6-9
Highest	10+

Figure 3: Proposed Data Characteristics Evaluation Metric[4]

Arbor, MI: University of Michigan Inter-University Consortium for Political and Social Research. Accessed May 21, 2019. https://deepblue.lib.umich.edu/handle/2027.42/143808.