



11-2018

Researcher Credentialing for Administrative Data: Easing the Burden, Increasing the Efficiency

Allison R.B. Tyler

University of Michigan-Ann Arbor

Johanna Davidson Bleckman

University of Michigan-Ann Arbor

Margaret C. Levenstein

University of Michigan-Ann Arbor

Follow this and additional works at: [https://repository.upenn.edu/
admindata_conferences_presentations_2018](https://repository.upenn.edu/admindata_conferences_presentations_2018)

Tyler, Allison R.B.; Davidson Bleckman, Johanna; and Levenstein, Margaret C., "Researcher Credentialing for Administrative Data: Easing the Burden, Increasing the Efficiency" (2018). *2018 ADRF Network Research Conference Presentations*. 36.
https://repository.upenn.edu/admindata_conferences_presentations_2018/36

DOI <https://doi.org/10.23889/ijpds.v3i5.1050>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/admindata_conferences_presentations_2018/36
For more information, please contact repository@pobox.upenn.edu.

Researcher Credentialing for Administrative Data: Easing the Burden, Increasing the Efficiency

Abstract

While not collected primarily for research purposes, administrative data gathered as part of normal agency or program operations present unique and vital opportunities for researchers. However, as with other types of restricted access data, these data are often not made available or require special authorization to access. The risks of re-identification, social stigma, and privacy violations for individuals represented in the data, especially for special populations, require that data be securely held and access authorization be moderated. Obtaining this authorization, especially for multiple datasets from one institution or spread across multiple data providers, imposes a significant administrative burden on researchers and data repository staff who must repeatedly and redundantly provide and validate user identities.

This presentation will offer the Inter-university Consortium for Political and Social Research's solution to ease the burden on both researchers and data repository staff -- the Researcher Passport. The researcher passport will incorporate standardized, community-normed identity verification criteria, data security level interpretations, and restricted data access training requirements. The researcher passport, using these standards, will be issued by a central authority who carries out the identity verification process and issues a tiered-access passport to users who are authorized for a streamlined data access request process for certain levels of secure data. Visas issued by data custodians control "entry" by passport holders to particular data sets. We describe these standards, how they fit into the repository workflow, and how they make the data access process for efficient and effective for administrative data users and providers.

Comments

DOI <https://doi.org/10.23889/ijpds.v3i5.1050>

Researcher Credentialing for Administrative Data: Easing the Burden, Increasing the Efficiency

Allison Tyler
University of Michigan
November 13, 2018

Introduction

- 2017-2019 Alfred P. Sloan Foundation funded research study
- Research team:
 - PI: Margaret Levenstein, Director, ICPSR
 - Allison R. B. Tyler, UMSI
 - Johanna Davidson Bleckman, ICPSR



Background

Introduction

- An increasing push to make research data available for **reuse** and **secondary analysis** (Holdren, 2013)
- Restricted data security presents a challenge to this goal because of time-consuming and redundant **identity verification processes**

Introduction

- Interested in the repository process of **digital access credentials** assignment for restricted data users
- Recommend a novel **transferable** researcher credentialing system to make the data access process for restricted data more efficient



Research

Methodology

- Environmental scan
 - Interviews
 - Document analysis

Repositories	
United States	16
International	7
Data Types	
Repository Documentation	355
Interviews	13

Identity Verification Criteria

Identity		Training		Project	
Name	18	Data privacy & confidentiality	4	Statement of purpose	15
Institutional affiliation	18	Responsible data use	2	Signed data use agreement	13
Contact information	13	Information security	2	Data security plan	8
Country of residence	10	Disclosure control	2	IRB/Ethics review	6
Supervisor information	7	CIPSEA	1	Project timeline	5
CV	3	Title 13/26	1	Justification for data	5

Terminology

Security Level	Most Common Term at this Level	Other Terms
Low	Public-Use	Open access, Green, Data available, Open access for registered users, Unrestricted, Category (Cat) 0
Moderate	Restricted Controlled Access	Yellow, Cat A, Limited data set, Secure dissemination, General
High	Restricted	Confidential, Government regulated, Red, Cat B, Cat C, Restricted release, Research identifiable, Virtual enclave, Physical enclave, Closed, Special, Other access, Safeguarded

Results

- Which leads to:
 - Unwillingness of repositories and data custodians to accept identity verification by other institutions

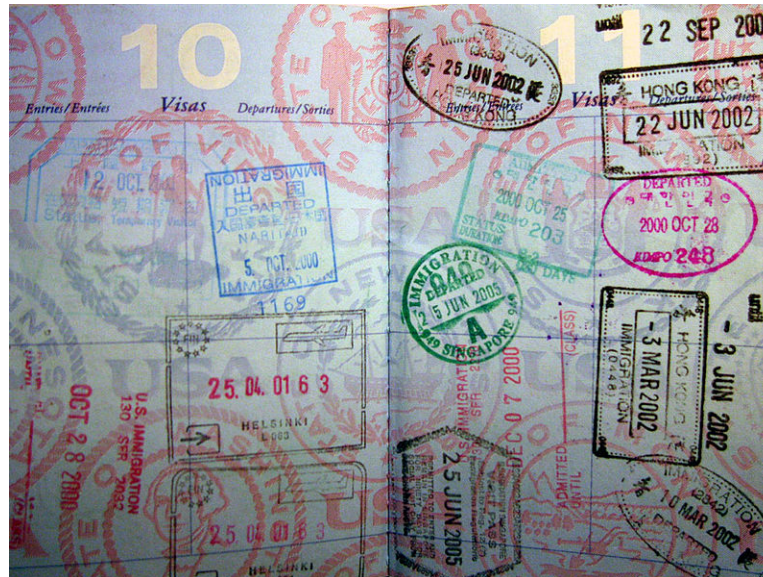
ICPSR Implementation

- Researcher passport
 - Issued to verified users indicating the highest data security level they have a streamlined review process for
 - Requests for higher security data require higher scrutiny

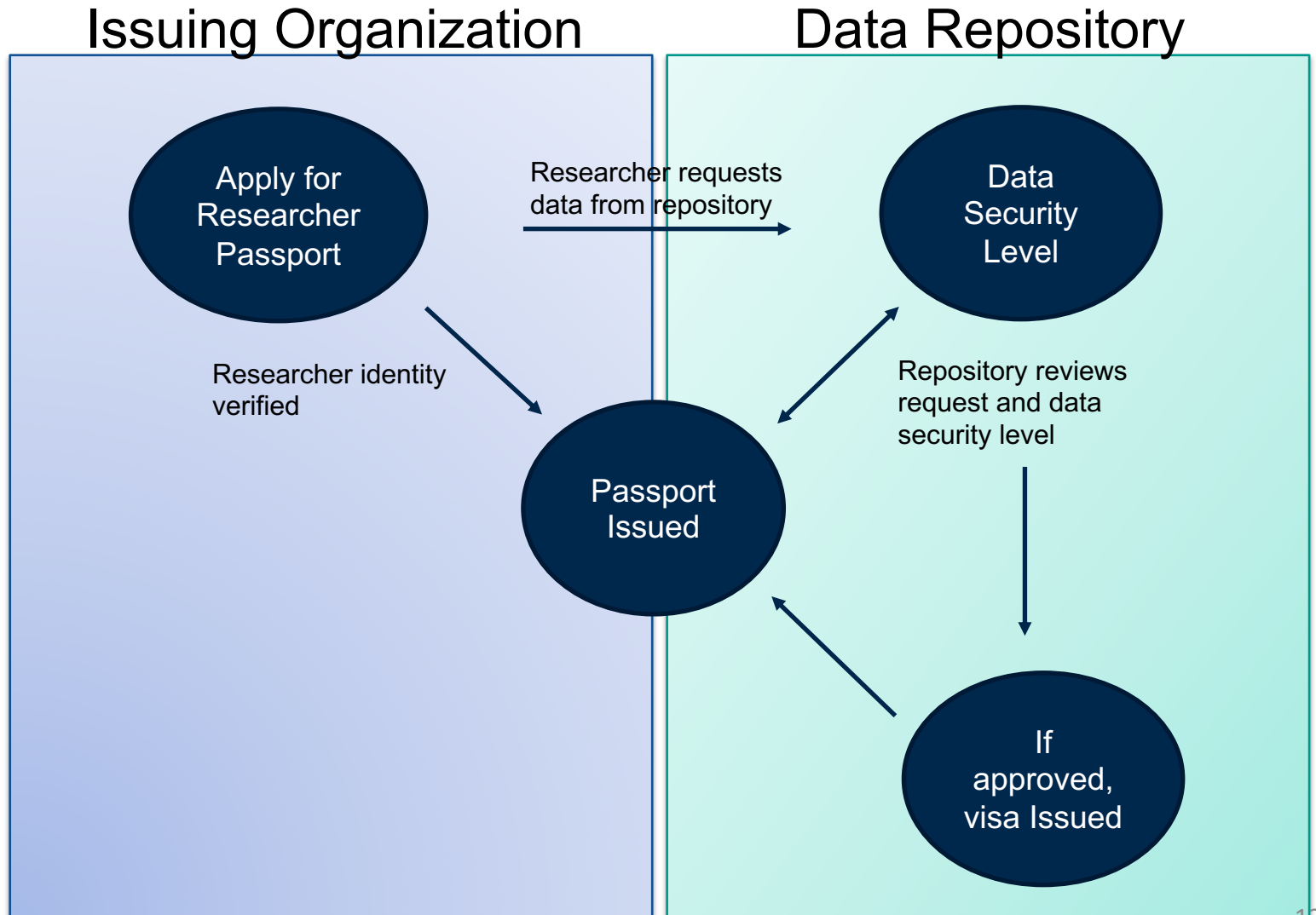


ICPSR Implementation

- Visas
 - Connecting credential holders to the data sets and access methods they are authorized for
 - Record of prior use



Credential & Access Process



Data Security Levels

- Low
 - *No identifiable information (generally public-access)*
- Moderate
 - *Not fully anonymized, non-trivial risk*
- High
 - *Not anonymized though without direct identifiers*
- Highest
 - *Includes direct identifiers, require special handling*

PI SCORE	DATA SECURITY LEVEL				
	LOW	MODERATE	HIGH	HIGHEST	
8+	unrestricted	secure download	secure download	VDE / physical enclave	Platinum
7	unrestricted	secure download	virtual enclave	physical enclave	Gold
6	unrestricted	secure download	virtual enclave	no access	Silver
5	unrestricted	virtual enclave	no access	no access	Bronze
4	unrestricted	virtual enclave	no access	no access	Copper
0-3	unrestricted	no access	no access	no access	Tin



Where We Are Now

Improving Data Access and Confidentiality Protection



What is a Researcher Passport?

It's a digital identity, or profile, that captures and verifies the information that data repositories need to know in order to share their data with you. It can then be provided to participating repositories to expedite your access to their data.

- ✔ Complete your profile
- ✔ Submit your application
- ✔ Share your passport as you apply for data access

Watch our one-minute



Open Badges

- Innovate Passport component:
 - Identity information
 - Academic qualifications
 - Professional experience
 - Training completion



Administrative Data

Administrative Data

- Administrative data is often restricted and subject to **legal** and **administrative requirements** and **confidentiality protection**
- Limits access for research and evidence building

Administrative Data

- Benefits:
 - Open badges to identify experience and training with restricted and administrative data
 - Record of prior data use
 - Build consensus about what it means for data to be restricted
 - training completion

Administrative Data

- The Researcher Passport will:
 - Streamline access
 - Implicitly and explicitly provide guidance to data custodians by establishing/reinforcing shared norms
 - Increase trust by creating common language and maintaining history of researcher “travels” in restricted data

Next Steps

- Test credentialing system and interface
- Collaborate with restricted data community partners
- Provide guidance for evaluating disclosure risk consideration for administrative data

Conclusion

- To meet the goal of increasing the utility of restricted research and administrative data, we are developing a trusted ***Researcher Passport***
- Recommend standardizing the process of **identity verification** and **data security level assignment**

Further Information

Levenstein, M.C., Tyler, A.R.B., Davidson Bleckman, J. 2018. The Researcher Passport: Improving Data Access and Confidentiality Protection: ICPSR's Strategy for a Community-normed System of Digital Identities of Access. ICPSR White Paper Series No. 1. Ann Arbor, MI: University of Michigan Inter-university Consortium for Political and Social Research. <https://deepblue.lib.umich.edu/handle/2027.42/143808>



Additional Slides

Research Questions

- How do restricted data repositories authorize users to access restricted data?
- What challenges exist that impede the transferability of restricted data access?

User Identity Criteria

USER ATTRIBUTES	POINTS ATTRIBUTED
Highest degree held	
Doctoral/terminal degree	3
Graduate degree (non-terminal)	2
Undergraduate	1
No degree	0
Professional Position (choose one of the following two options)*	
Option 1 Academic faculty/staff: Highest institutional appointment/affiliation	
Full/Associate professor	3
Assistant professor	2
Student	1
Research staff	1
Option 2 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

Data Security Levels

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
<hr/>	
Data security score <i>(after totalling above)</i>	Range
low	0-3
moderate	4-5
high	6-9
highest	10+

Researcher Passport - Professor

For data set with Moderate security level

USER ATTRIBUTES	RESULTS	POINTS
Highest degree held	Ph.D. →	3
Professional Position	Associate Prof. →	3
Other		
Recognized Federal clearances	none →	0
Current (2pts) or recent (1pt) Federal grant	yes - recent →	1
Research publications (1 or more publications)	yes →	2
Restricted Data Use Experience (1 or more projects)	none →	0
Potential dataset- or repository-specific user requirements		total = 9 pts
Country- or region-specific citizenship or residency status	US citizen	
Affiliation with Carnegie Classified Academic Institution	yes	
Badges earned and verified		
Trainings*		
Data security — Levels I-III	Level I	
Research conduct — Levels I-III	Level II	
Specific expertise	n/a	
Contributions — data stewardship		
History of data sharing	DOI	
History of metadata enhancement	no	
History of code/syntax sharing	DOI	
Confirmed research misconduct	no	

Researcher Passport - Professor

PI SCORE	DATA SECURITY LEVEL				
	LOW	MODERATE	HIGH	HIGHEST	
8+	unrestricted	download 	secure download	VDE / physical enclave	Platinum
7	unrestricted	secure download	virtual enclave	physical enclave	Gold
6	unrestricted	secure download	virtual enclave	no access	Silver
5	unrestricted	virtual enclave	no access	no access	Bronze
4	unrestricted	virtual enclave	no access	no access	Copper
0-3	unrestricted	no access	no access	no access	Tin

VISA RECOMMENDATION

User	Total Points	9
	User's Passport Level	Platinum
Data	Security Level of NISVS Data	Moderate
Visa	Visa recommended for this User + these data?	Yes
	Mode of Access recommended for this user/data combination?	Secure Download

Researcher Passport – NYT Reporter

For data set with Moderate security level

USER ATTRIBUTES	RESULTS	POINTS
Highest degree held	B.A. —→	1
Professional Position: Media years experience	3 —→	2
Other		
Recognized Federal clearances	none —→	0
Current (2pts) or recent (1pt) Federal grant	none —→	0
Research publications (1 or more publications)	none —→	0
Restricted Data Use Experience (1 or more projects)	yes —→	2
Potential dataset- or repository-specific user requirements		total = 5 pts
Country- or region-specific citizenship or residency status	Spanish citizen, currently living in US (3 yrs)	
Affiliation with Carnegie Classified Academic Institution	no	
Badges earned and verified		
Trainings		
Data security — Levels I-III	Level I	
Research conduct — Levels I-III	Level I	
Specific expertise	n/a	
Contributions — data stewardship		
History of data sharing	no	
History of metadata enhancement	no	
History of code/syntax sharing	no	
Confirmed research misconduct	no	

Researcher Passport - NYT Reporter

PI SCORE	DATA SECURITY LEVEL				
	LOW	MODERATE	HIGH	HIGHEST	
8+	unrestricted	secure download	secure download	VDE / physical enclave	Platinum
7	unrestricted	secure download	virtual enclave	physical enclave	Gold
6	unrestricted	secure download	virtual enclave	no access	Silver
5	unrestricted	secure download	no access	no access	Bronze
4	unrestricted	virtual enclave	no access	no access	Copper
0-3	unrestricted	no access	no access	no access	Tin

VISA RECOMMENDATION

User	Total Points	5
	User's Passport Level	Bronze
Data	Security Level of NISVS Data	Moderate
Visa	Visa recommended for this User + these data?	Yes
	Mode of Access recommended for this user/data combination?	VDE