OMB No. 0925-0001 and 0925-0002 (Rev. 10/15 Approved Through 10/31/2018)

**BIOGRAPHICAL SKETCH**

Provide the following information for the Senior/key personnel and other significant contributors.

Follow this format for each person. DO NOT EXCEED FIVE PAGES.

|  |
| --- |
| NAME: Vilhuber, Lars |
| eRA COMMONS USER NAME (credential, e.g., agency login): lvilhuber |
| POSITION TITLE: Executive Director |

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

|  |  |  |  |
| --- | --- | --- | --- |
| INSTITUTION AND LOCATION | DEGREE (if applicable) | Completion Date  MM/YYYY | FIELD OF STUDY |
| Rheinische Friedrich Wilhelms Universität Bonn, Economics, Bonn, Nordrhein-Westfalen | OTH | 07/1992 | Economics |
| Université de Montréal, Economics, Montreal, QC | MS | 1993 | Economics |
| Université de Montréal, Economics, Montreal, QC | PHD | 1998 | Economics |
| York University, Toronto, ON | Postdoctoral Fellow | 06/1999 | Responsible for dissemination of admin/survey data |

### **A. Personal Statement**

My role in the proposed research is to provide general project oversight, specific use cases, knowledge of software development processes, and in cross-domain issues arising from disparate metadata schemas. The development of multi-disciplinary metadata, and the facilitation of discovery of available data to accelerate and enhance scientific discovery has been a core part of my scientific activities since my post-doc at York University, where I documented administrative datasets and survey extracts, made available to various researchers, in scalable ways long before more robust tools were made available. Our Cornell-based group's impetus to developing the NSF-funded Comprehensive Extensible Data Documentation and Access Repository (CED2AR) was driven by the observed lack of tools to disseminate even metadata on confidential datasets, greatly inhibiting the ability of a broader research community to develop novel research. The current proposal is squarely in line with this motivation and past activity. My experience in developing CED2AR, supervising scientific input as well as software development to create a robust, user-oriented software for editing schema-compliant metadata by non-specialists, will be particularly applicable in the current research proposal, where many of the same issues will arise. Our constant interaction with the user community, and rapid-response software development to respond to unforeseen needs, serve as a solid guideline for the expected development effort in this grant. I am a member of multiple scientific committees that advise networks of research data centers, and am involved with the Census Bureau's Center for Economic Research, which manages the Federal Statistical Research Data Centers. All of these entities face the problem on how to properly document an extraordinarily varied set of confidential datasets, making them accessible to a broad spectrum of scientific domains. Metadata alignment problems arise not just from the problems of cross-disciplinary incompatibility, but also from idiosyncratic backend systems that were designed with administration, not data dissemination, in mind, and often do not share the concepts that standard research-oriented metadata schemas expect - a situation not dissimilar to the one faced by this project. Thus, my past experience has prepared me well for both the organizational and scientific challenges of this project.

1. Lagoze Carl, Vilhuber Lars, Williams J, Perry B, Block W. CED2AR: The Comprehensive Extensible Data Documentation and Access Repository. IEEE/ACM Joint Conference on Digital Libraries. 2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL); ; London, United Kingdom. IEEE; c2014.
2. Lagoze Carl, Block W, Williams J, Abowd John, Vilhuber Lars. Data Management of Confidential Data. International Journal of Digital Curation. 2013 June 20; 8(1):265-278.
3. Lagoze Carl, Willliams J, Vilhuber Lars. Communications in Computer and Information Science. Cham: Springer International Publishing; 2013. Chapter Chapter 13, Encoding Provenance Metadata for Social Science Datasets. 123-134p.
4. Abowd J, Vilhuber Lars, Block W. Lecture Notes in Computer Science. Hutchison D, Kanade T, Kittler J, Kleinberg J, Mattern F, Mitchell J, Naor Moni, Nierstrasz O, Pandu Rangan C., Steffen B, Sudan M, Terzopoulos D, Tygar Doug, Vardi M, Weikum G, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. Chapter Chapter 17, A Proposed Solution to the Archiving and Curation of Confidential Scientific Inputs. 216-225p.

### **B. Positions and Honors**

**Positions and Employment**

|  |  |
| --- | --- |
| 1999 - 2001 | Economist, LEHD Program, U.S. Census Bureau, Washington, DC |
| 1999 - 2002 | Adjunct Professor, York University, Toronto |
| 2001 - | Economist (various contracts), U.S. Census Bureau, Washington, DC |
| 2006 - | Senior Research Associate, Cornell University, Economics, Ithaca, NY |
| 2007 - | Vice-President and Senior Researcher, ACES-Research, LLC |
| 2011 - | Executive Director, Labor Dynamics Institute, Cornell University, Ithaca, NY |

**Other Experience and Professional Memberships**

|  |  |
| --- | --- |
| 1996 - | Member, American Economic Association |
| 2010 - | Member, American Statistical Association |
| 2012 - | PI of the national Coordinating Office, NSF-Census Research Network |
| 2013 - | Institutional Member, Association of Public Data Users |
| 2014 - | Member, Scientific Advisory Committee, Quebec Inter-University Centre for Social Statistics |
| 2015 - | Chair of Scientific Committee, Centre d'accès sécurisé aux données (France) |
| 2015 - | New York Federal Statistical Research Data Center, Senior Advisor |
| 2016 - | Member of Privacy and Confidentiality Committee, American Statistical Association |
| 2016 - | Member of Advisory Board of International Data Service Center, IZA - Forschungsinstitut zur Zukunft der Arbeit |

**Honors**

### **C. Contribution to Science**

1. Development of metadata repository, metadata standards, and metadata viewer with features to support confidential data and restricted-access environments
   1. Lagoze Carl, Vilhuber Lars, Williams J, Perry B, Block W. CED2AR: The Comprehensive Extensible Data Documentation and Access Repository. IEEE/ACM Joint Conference on Digital Libraries. 2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL); ; London, United Kingdom. IEEE; c2014.
   2. Lagoze Carl, Block W, Williams J, Abowd John, Vilhuber Lars. Data Management of Confidential Data. International Journal of Digital Curation. 2013 June 20; 8(1):265-278.
   3. Lagoze Carl, Willliams J, Vilhuber Lars. Communications in Computer and Information Science. Cham: Springer International Publishing; 2013. Chapter Chapter 13, Encoding Provenance Metadata for Social Science Datasets. 123-134p.
   4. Abowd J, Vilhuber Lars, Block W. Lecture Notes in Computer Science. Hutchison D, Kanade T, Kittler J, Kleinberg J, Mattern F, Mitchell J, Naor Moni, Nierstrasz O, Pandu Rangan C., Steffen B, Sudan M, Terzopoulos D, Tygar Doug, Vardi M, Weikum G, editors. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. Chapter Chapter 17, A Proposed Solution to the Archiving and Curation of Confidential Scientific Inputs. 216-225p.
2. PI on NSF-funded Coordinating Office of the NSF-Census Research Network, combining researchers from geography, statistics, computer science, economics, and federal agencies, interacting and communicating with the federal and state statistical systems, and the academic community. Maintenance of a central repository, cataloging network-generated output (publications, education, software, events) at <http://www.ncrn.info>.
3. Studies in labor dynamics, labor mobility, and local labor markets
   1. Abowd J, Vilhuber Lars. Did the Housing Price Bubble Clobber Local Labor Market Job and Worker Flows When It Burst?. American Economic Review. 2012 May; 102(3):589-593.
   2. Abowd J, Vilhuber Lars. National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail. Journal of Econometrics. 2011; 161(1):82-99.
   3. Dunne T, Jensen J, Roberts M, Abowd J, McKinney K, Vilhuber Lars. Producer Dynamics. University of Chicago Press; 2009. The Link between Human Capital, Mass Layoffs, and Firm Deaths. 447-472p.
   4. Abowd J, Vilhuber Lars. The Sensitivity of Economic Statistics to Coding Errors in Personal Identifiers. Journal of Business & Economic Statistics. 2005 April; 23(2):133-152.
4. Development of protection measures for confidential data, application of such measures to releasable data.
   1. Miranda J, Vilhuber Lars. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2014. Chapter Chapter 18, Using Partially Synthetic Data to Replace Suppression in the Business Dynamics Statistics: Early Results. 232-242p.
   2. Drechsler Jörg, Vilhuber Lars. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2014. Chapter Chapter 19, Synthetic Longitudinal Business Databases for International Comparisons. 243-252p.
   3. Abowd J, Gittings R, McKinney K, Stephens B, Vilhuber Lars, Woodcock S. Dynamically Consistent Noise Infusion and Partially Synthetic Data as Confidentiality Protection Measures for Related Time Series. SSRN Electronic Journal. 2012; :-.
   4. Machanavajjhala A, Kifer D, Abowd John, Gehrke J, Vilhuber Lars. Privacy: Theory meets Practice on the Map. 2008 IEEE 24th International Conference on Data Engineering. 2008 IEEE 24th International Conference on Data Engineering (ICDE 2008); ; Cancun, Mexico. IEEE; c2008.

Complete List of Published Work in My Bibliography:  
<http://bit.ly/2fjTCDK>

### **D. Additional Information: Research Support and/or Scholastic Performance**

**Ongoing Research Support**

1507241, National Science Foundation - Directorate for Social, Behavioral & Economic Sciences

Lars Vilhuber (PI)

10/01/14-06/30/17

RCN: Coordination of the NSF-Census Research Network

Role: PI

1012593, National Science Foundation - Directorate for Computer & Information Science & Engineering

Lars Vilhuber (PI)

07/15/10-06/30/17

TC: Large: Collaborative Research: Practical Privacy: Metrics and Methods for Protecting Record-level and Relational Data

Role: PI

G-2015-13903, Alfred P. Sloan Foundation

Vilhuber, Lars (PI)

04/01/15-03/01/18

To study the economics of socially efficient protocols for managing research databases containing private information

Role: CPI

1131848, National Science Foundation

Vilhuber, Lars (PI)

10/01/11-09/01/17

NCRN-MN: Cornell Census-NSF Research Node: Integrated Research Support, Training and Data Documentation

Role: PI

**Completed Research Support**

1042181, National Science Foundation

Lars Vilhuber (PI)

09/15/10-08/31/15

Synthetic Data User Testing and Dissemination

Role: PI

0820349, National Science Foundation

Lars Vilhuber (PI)

10/01/08-09/30/13

The economics of mass layoffs: displaced workers, displacing firms,and causes and consequences

Role: PI

0922005, National Science Foundation - Directorate for Social, Behavioral & Economic Sciences

Lars Vilhuber (PI)

07/01/09-06/30/13

Social Science Gateway to TeraGrid

Role: PI

0922494, National Science Foundation

Lars Vilhuber (PI)

07/01/09-06/30/11

Joint NSF-Census-IRS Workshop on synthetic data and confidentiality protection, July 2009 Washington, DC

Role: PI

R41 AG029756-01

VILHUBER, LARS  (PI)

09/15/07-08/31/10

Synthesis of HRS-SSA linked data

Role: PI