NYU | WAGNER

NASPAA 50 YEARS
The Global Standard in Public Service Education

Data Science for the Public Sector

Laurel McFarland
Julia Lane
Sherry Glied

NYU WAGNER

NASPAA 50 YEARS
The Global Standard in Public Service Education

Purpose

- Document the need for a new approach to data science that could be linked to NASPAA Schools
- Illustrate an example of how the pieces fit together
- Provide a suggested overview of data science competencies
- Discuss the potential role of NASPAA and/or a set of Deans/Schools in developing the new approach

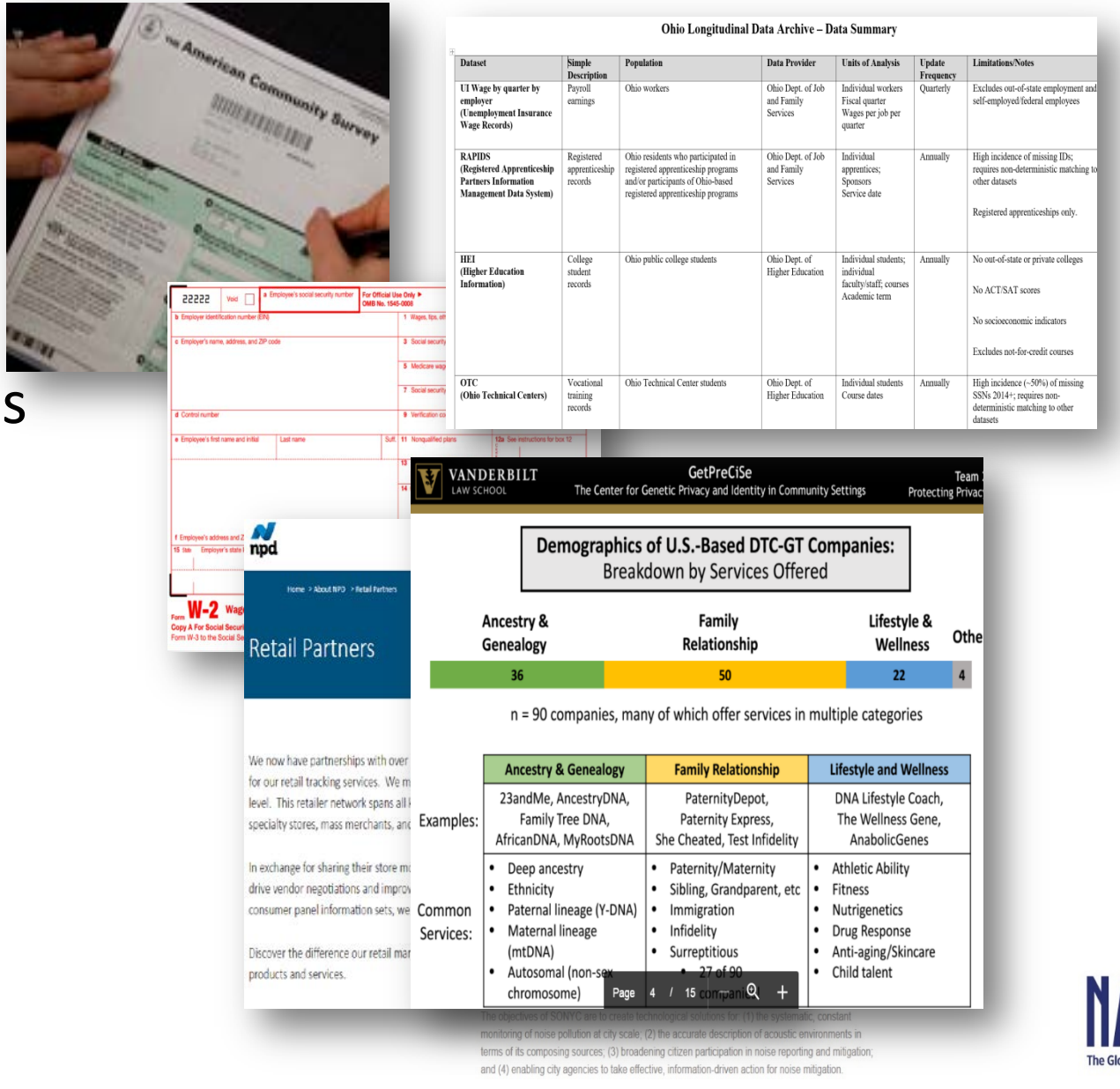
Why a new approach is needed

Key points

- Public policy data has changed; schools need to change what we teach
- New opportunity with federal legislation and associated response
- Successful pilot approach could be tested or expanded
- Potential role of NASPAA

Change in data

- Surveys
- Administrative records
- Transaction data



NYU WAGNER

NASPAA 50 YEARS
The Global Standard in Public Service Education

Change in legislation

H.R. 1831: Evidence-Based P
2016

Introduced: Apr 16, 2015
114th Congress, 2015–2017

Status: Enacted — Signed by the Presid
This bill was enacted after being s

Law: Pub.L. 114-140

Sponsor: Paul Ryan
Representative for Wiscons
Republican

Text: Read Text »
Last Updated: Mar 18, 201
Length: 5 pages

- 2020 Census ($
design, howeve
on the new desi
- CEDCaP ($78M
- American Com
continuing our e
- Geographic Su
smarter geograp
- Administrative
and federally sp
linkage techniqu
protection, and
- Economic & G
are relevant to t

BIPARTISAN POLICY CENTER

Foundations for Evidence-Based Policymaking Act of 2018

The bipartisan Foundations for Evidence-Based Policymaking Act of 2018 builds off the work of the U.S. Commission on Evidence-Based Policymaking to strengthen data privacy protections, improve secure access to data, and enhance the federal government's capacity for producing and using evidence.

Strengthens Privacy Protections

Maintains Strong Confidentiality Protections for Sensitive Data. Reauthorizes the Confidential Information Protection and Statistical Efficiency Act (CIPSEA), an existing law that gives the American public strong privacy safeguards and legal protections for appropriate uses of confidential data.

Institutes Processes to Assess Data Risks. Strengthens efforts to protect confidentiality while making data accessible for evidence building and transparent to the public by requiring comprehensive risk assessments for certain publicly released data.

Enhances Public Trust in Data. Improves public trust in statistical activities by codifying language directing certain agencies to establish procedures to protect trust in data activities by appropriately maintaining objectivity, independence, and confidentiality.

Establishes Consistent Leadership on Key Data Issues. Ensures a senior leader in each agency is responsible for protecting privacy and ensuring confidentiality protections are appropriately applied by creating chief data officers.

Improves Secure Data Access

Encourages Agencies to Make Data Public and Open When Possible. Takes steps to improve the public information about what data government currently holds and make data publicly available when possible and in the public interest.

Requires Development of Data Inventories. Enables

Makes Administrative Records Available for Evidence Building. Under a strong set of confidentiality protections, encourages that government data can and should be used to generate evidence about policies and programs, unless otherwise restricted by law.

Creates a Common Portal for Researcher Applications to Access Restricted Data. Reduces burden on researchers for applying to access government data by establishing a common application system for qualified individuals to access restricted, confidential data for approved projects.

Facilitates Continuous Feedback about Data Coordination. Promotes the use of data for evidence building by establishing a government advisory committee to review existing coordination and availability of data.

Enhances Government's Evidence Capacity

Directs Agencies to Develop Evidence Plans. Enables agencies to better prioritize evidence building by requiring that agencies document their key research questions, data needs, and planned activities.

Prioritizes Evaluation Activities in Agencies. Improves agency capacity to engage in and use program evaluation by establishing evaluation officers in government agencies and requiring agencies to develop written evaluation policies.

Develops Baseline Information about the Resources Available for Evidence Building. Directs government agencies to periodically assess and report on their capabilities to engage

arch Facility: The
h Facility is a pilot
access to analytical
overy services, and
s for users, includ-
l government ana-
chers. The Census
ers developed the
aborative Training
nalytics sponsored
, New York Univer-
Maryland.[1] It is cur-
ith users accessing
ining program. The
l-based computing
security approvals,
ected confidential
ent of Housing and
Census Bureau, as
y agencies, and an

NYU | WAGNER

NASPAA 50 YEARS
The Global Standard in Public Service Education

What Role for NASPAA Schools?

- Prepare current student for careers using administrative data for evidence building
  - NASPAA data competency
  - Both in the public sectors and in the non-profit sector

- But – data are available now and most current government employees do not have this capacity
  - Develop programs for current government employees (including alumni)

Possibility…

- Can NASPAA Schools work together to develop an educational program that could function:
  - As a data science track within existing MPA/MPP programs?
  - As a stand-alone executive training program for current government employees?

Necessary Components

Technical

- Create secure environment where data providers can share their data across agency and jurisdictional lines and students (current MPAs or agency employees) can access the data under appropriate safeguards

Create Curriculum

- Develop pedagogical tools to enable students to
  - Link and analyze disparate data
  - Learn to ask questions that available data can answer

Value

- Document value associated with data
  - Consistent with NASPAA mission
  - Useful to participating external agencies

Case Study – Coleridge Initiative

A Model for Training Government Staff – Coleridge Initiative

Technical

- Create secure environment where data providers can share their data across agency and jurisdictional lines
- Census Authorization to Operate; HHS and USDA in process

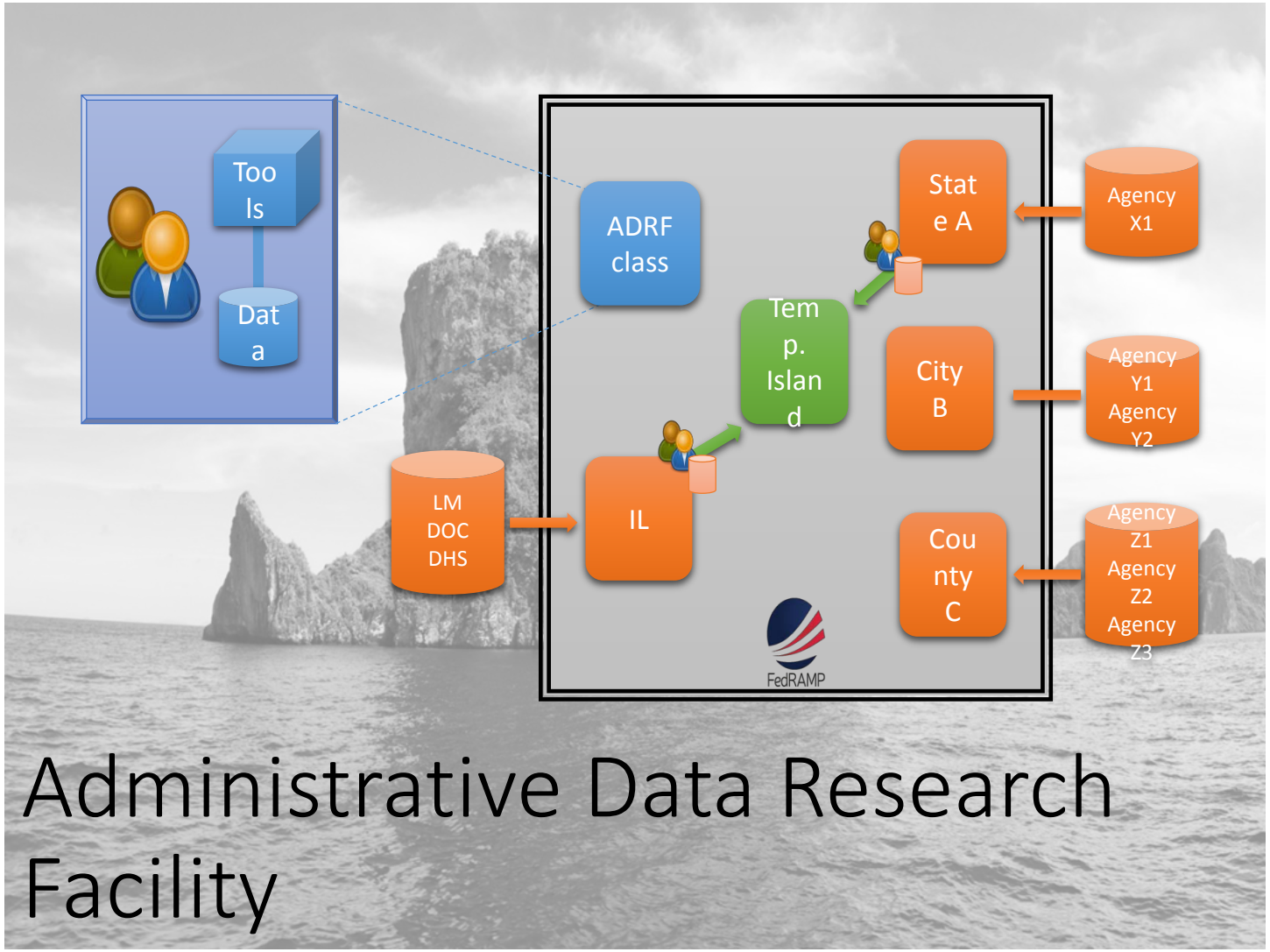
Create Curriculum

- Link disparate data
- Analyze data
- Develop new analytical capacity

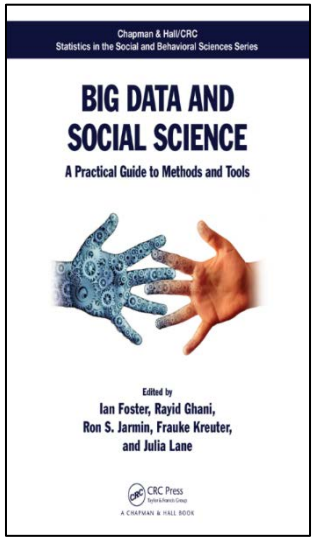
Value

- Document value associated with data
  - Consistent with the agency mission
  - Useful enough to engage decision-makers

Details

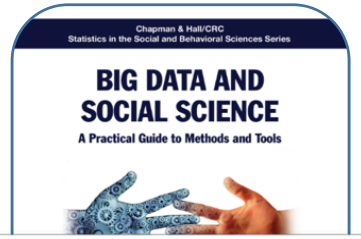

Too
ls
Dat
a
ADRF
class
Stat
e A
Agency
X1
Tem
p.
Islan
d
City
B
Agency
Y1
Agency
Y2
LM
DOC
DHS
IL
Cou
nty
C
Agency
Z1
Agency
Z2
Agency
Z3
FedRAMP
Administrative Data Research Facility

Curriculum -- Textbook



"If you work in social science and would like to explore the power of big data, this book is clearly for you…This book is complete and comprehensive. It covers all necessary steps to finish a big data project; collecting raw data, cleaning and preprocessing data, applying various modeling tools to analyze the data, evaluating results, protecting privacy, and addressing ethical problems…All the important topics concerning big data are covered, making this book a good reference that you should always keep on your desk." (2017) Book Reviews, ***Journal of the American Statistical Association***, 112:518, 878-882, DOI: 10.1080/01621459.2017.1325629

THE UNIVERSITY OF CHICAGO



UNIVERSITY OF MARYLAND 18 56

OHIO STATE

Book Contents

Chapter 1: Introduction

I. Capture and Curation

Chapter 2: Working with Web Data and APIs

Chapter 3: Record Linkage

Chapter 4: Databases

Chapter 5: Programming with Big Data

II. Modeling and Analysis

Chapter 6: Machine Learning

Chapter 7: Text Analysis

Chapter 8: Networks: The Basics

III. Inference and Ethics

Chapter 9: Information Visualization

Chapter 10: Errors and Inference

Chapter 11: Privacy and Confidentiality

Chapter 12: Workbooks

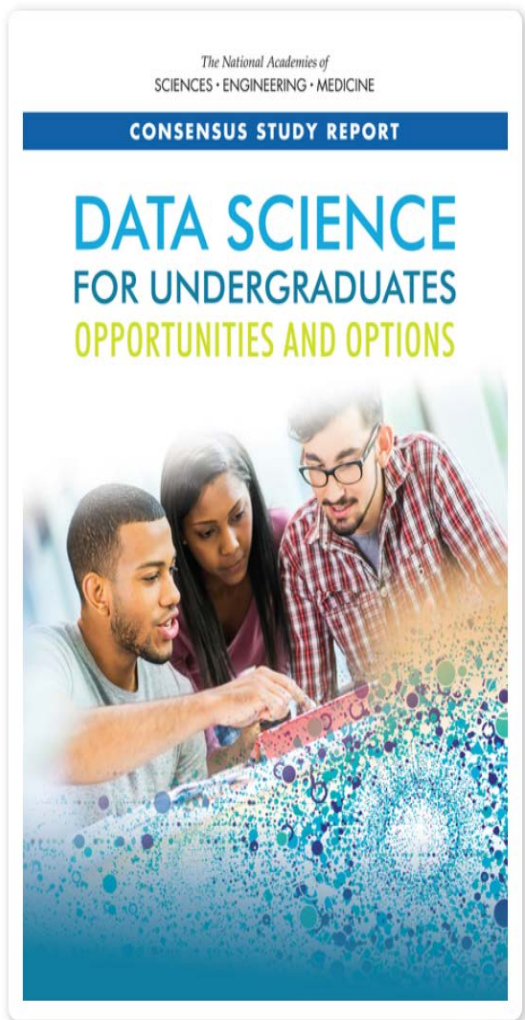
N

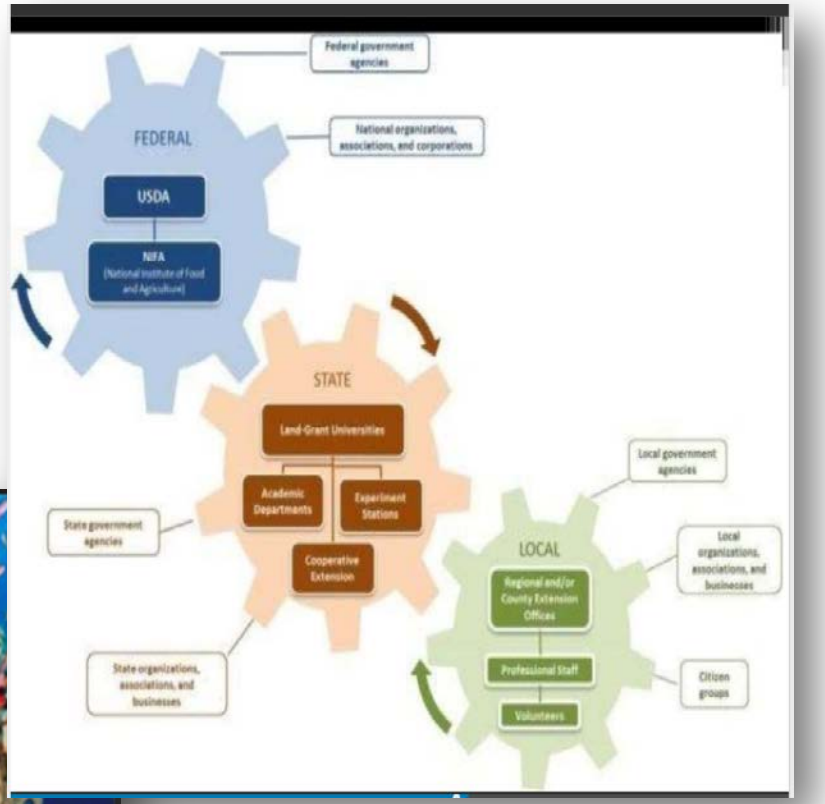
University of Missouri

INDIANA UNIVERSITY

SPAA 50 YEARS
Standard in Public Service Education

Broader Vision





NYU | WAGNER

NASPAA 50 YEARS
The Global Standard in Public Service Education

Key points

- Public policy data has changed; schools need to change what we teach
- New opportunity with federal legislation and associated response
- Successful pilot approach could be tested or expanded
- Potential role of NASPAA

Moving Forward- Potential Role of NASPAA

Opportunities for NASPAA Schools

- Develop a framework for establishing and maintaining a secure facility to access confidential data
  - Identify and address challenges
  - Offer case studies of success
- Collaborate on curriculum/pedagogical models
  - Develop shared pedagogical learning environment, case studies etc.
  - Leave explicit space for local experimentation and modification
- Establish evidence of value
  - Develop evidence-based paradigm for assessing value of these efforts both in MPA/MPP programs and in executive programs

Questions for the Group

- Is there interest in pursuing this agenda as a collective enterprise?
- Next steps would be to develop background white papers on technical aspects, pedagogic aspects, and evaluation aspects and discussing at NASPAA meeting in October
  - Are these the correct topics?
  - Would you be interested in participating in the development of the White Papers (as a reviewer/editor)?
  - Would you be interested in attending a meeting to discuss the papers at NASPAA?

Thank you

NYU | WAGNER

NASPAA 50 YEARS
The Global Standard in Public Service Education

Background information

- Website: https://coleridgeinitiative.org/
- GitHub organization: https://github.com/Coleridge-Initiative