



# FEDERAL DEMONSTRATION PARTNERSHIP

Redefining the Government & University Research Partnership

## Data Stewardship Subcommittee: Survey Results

Melissa Korf, Harvard University

Rick Ikeda, NIH

FDP Meeting – January 2017



# Sample Description

- 65 institutions responded (though every respondent did not answer every question).
- Responding institutions represented all categories of amount of federal funding per year:
  - 24.58% (16) receive less than \$50 million per year
  - 18.48% (12) receive \$50-99 million per year
  - 27.67% (18) receive \$100-299 million per year
  - 23.06% (15) receive \$300-699 million per year
  - 6.19% (4) receive more than \$700 million



## Sample Description (cont.)

- The primary source of Federal funding is NIH for most respondents (59.34%) followed by NSF (24.96%) and DoD (9.39%). Only 4 institutions provide an “other” response.
- In terms of status under HIPAA:
  - 47.65% (31) are a covered entity
  - 26.17% (17) are a hybrid entity
  - 26.17% (17) are not a covered entity



# Sources of Third Party Data

Overall Rank	Item	Rank Distribution	Score	Total Respondents
1	a. Federal repositories		218	59
2	c. Individual Collaborators at other institutions		201	56
3	b. Non-Federal repositories		179	55
4	d. Companies/industry		165	58
5	e. Other: Enter comments below:		50	26



# Data Use and Sharing Burden Assessments

- All items had an average burden assessment over 3, but the items with the highest ratings were:
  - Compliance with overly prescriptive data security requirements (3.86 average)
  - Adapting to constantly changing data security requirements (3.85 average)
  - Inconsistent data management plan and data sharing requirements across Federal agencies (3.8 average)
  - Lack of guidance and training on data management issues by funders for researchers (3.7 average)
  - Ensuring compliance with data sharing obligations (3.66 average)
  - Lack of guidance and training on data management issues by funders for institutions (3.66 average)
- The number of data use agreements processed and complex and difficult negotiations of data use agreements received average ratings of 3.28 and 3.2, respectively



# Other Data Use and Sharing Burdens

- “Other” factors mentioned as burdens:
  - Building up brand new work-flows, procedures, equipment, and personnel for interdisciplinary use and sharing of data will also be a burden. An underappreciated component is data integrity and re-use issues. Instructors and researchers may need to use data beyond the purpose of the initial research design.
  - Export Control issues related to data sharing present a relatively high burden.
  - Coordination across non-research units is a high burden.
  - Differing standards for the EU and US can make for difficult negotiations.
  - Dealing with conflicting data laws/requirements, often for organizations that misunderstand what is required (e.g., requiring us to sign a BAA for research data for research use that was gathered pursuant to informed consent-when compliance with data security is the core concern).



# How can FDP best help?

- Of the topics included in the survey, the FDP can be the most help in reducing the burden of:
  - Inconsistent data management plan and data sharing requirements across Federal agencies (26 votes)
  - Lack of guidance and training on data management issues (11 votes)
  - Adapting to constantly changing data security requirements (8 votes)
  - Compliance with overly prescriptive data security requirements (7 votes)



# Data Transfer and Use Agreements( DTUAs)

- **Estimated Volume**
  - 10.9% (7) process 0-10
  - 32.8% (21) process 11-50
  - 23.4% (15) process 51-100
  - 17.2% (11) process 101-200
  - 15.6% (10) process more than 200
- **Average Turn-around time**
  - 37.5% (24) have a turnaround time of 14 days or less
  - 43.8% (28) have a turnaround time of 15-30 days
  - 17.2% (11) have a turnaround time of 31-90 days
  - 1.6% (1) have a turnaround time of over 90 days



# DTUA contacts

- Many respondents selected more than one office, and many respondents chose “other” in order to write in that the nature of the data and the specific circumstances changed which office would handle the agreement.
  - The sponsored programs office is often involved in DTUAs (73.4% of respondents) followed by the Technology Transfer Office (45.3% of respondents).
  - The general counsel’s Office and IRB received far fewer responses (25% and 17.2%, respectively).
  - The research compliance and clinical trials offices were also mentioned in a few write-in responses.
- There appears to be moderate confidence that faculty at responding institutions know who to contact for assistance with a data transfer and use agreement. 53.8% agreed with this statement, and 24.6% provided the neutral response. Only 4.6% strongly agreed while 16.9% disagreed.



# Knowledge about DTUAs

- In terms of how well the needs for, use of, and compliance with DTUAs are understood:
  - By Faculty and Research Staff: 49.2% gave this the neutral score and 30.8% gave this a 2 (little understood).
  - By Research Administration: 41.9% scored this a 4 (well understood) with somewhat even numbers on either side (22.6% gave a 3 while 24.2% gave a 5).
  - By Others: 44% gave this the neutral score with ratings tapering off to either side.
- Comments:
  - Information can be silo'd or it can be hard to gauge understanding either because DTUAs are uncommon or because faculty don't raise them as a concern.
  - Further education and dissemination of information is needed.
  - Developments in this space are so new for some faculty that there hasn't been time for them to learn the information.



# DTUA tracking system?

- Responding institutions are relatively equally split on whether they use an electronic system to track data sharing and use agreements. 47.7% do use a system, and 52.3% do not.
- In the comments, a wide variety of off-the-shelf systems and home grown/built tracking databases were represented.



# Data Management Plan resources

- Faculty at my institution have sufficient resources available to them for assistance in developing a Data Management Plan.
  - 4.6% (3) Strongly Agree
  - 36.9% (24) Agree
  - 33.8% (22) Neutral
  - 21.5% (14) Disagree
  - 3.1% (2) Strongly Disagree
- Several responding institutions are working on developing resources or have new resources, so they are unsure whether faculty are aware of what is available.



# Training

- 42.2% offer training on data management, security, and/or data use agreements, and 20.3% require this training.
- Of the institutions who offer training, about half also require it (13 out of 27 respondents).



# Data Sharing and Public Access

- 16. My institution will have the resources to comply with the data sharing component of Public Access requirements.
  - 7.8% (5) Strongly Agree
  - 29.7% (19) Agree
  - 17.2% (11) are Neutral
  - 15.6% (10) Disagree
  - 6.3% (4) Strongly Disagree
  - 23.4% (15) say that it depends on the specific requirements



# Data Sharing and Public Access (cont.)

- “Costs associated with compliance with these requirements would be flowed down to the sponsor to the extent possible.”
- 4 comments included a variation on “Varies by discipline and data type”
- “Hardware and software needs related to storage, long-term preservation and providing perpetual access will require a robust infrastructure with a variety of components. Costs associated with this a concern, as equipment and building the infrastructure is expensive.”
- “We have an institutional repository that will accommodate data sets up to 50GB. We leverage many external repositories for our data. However, some data sets are too large, sensitive, or do not have a clear home that may not fall within the purview of these existing services. We are exploring solutions for providing public access as well as resources to fully solve this issue.”
- “Much of the burden will fall on the agencies themselves (e.g., to ensure appropriate evaluation on submitted DMPs), but if some activities are tasked to the University (e.g., ensuring compliance), it gets harder. Also, for example, if an agency required deposit of data in university repositories, we would not be ready (don’t have the capacity for huge increase).”
- “Stricter and more specific the requirements (and the extent to which they are actually enforced) will determine the level of help researchers need and therefore the amount of resources needed to help them.”
- “Because we are mostly NIH-funded, and faculty are aware of the requirements, we should be able to comply. If we had a lot of federal sponsors, it would likely be much more challenging.”



# Data Standardization related to Public Access

- All of the aspects included in the survey received moderate to high rankings of anticipated burden. The lowest average ranking was above 3.5.
- “Funding required to accommodate data standardization” received the highest ranking of anticipated burden with an average score of 4.02.
- “Staff effort/hours” came in second with an average score of 3.94. Notably, not a single respondent provided the lowest ranking (1) for this item.
- “Data Security” came in third with an average score of 3.79
- “Data storage requirements” received an average score of 3.56.



# FISMA compliance

- 16.4% (10 respondents) have a plan that enables compliance with FISMA low requirements
- 21.3% (13 respondents) have a plan that enables compliance with up to FISMA moderate requirements
- 9.8% (6 respondents) have a plan that enables compliance with FISMA requirements at all levels
- 37.7% (23 respondents) are currently working on developing a plan for compliance
- 14.8% (9 respondents) do not have a plan for compliance and are not currently working on developing one.
- With regards who feels the burden of their institution's efforts towards compliance:
  - The highest level of burden is attributed to Information Technology. Their average score, over 60 responses, was 4.13.
  - Research Administration and Faculty and Research Staff were just about tied for second with average scores of 3.63 and 3.58 respectively.
  - Only 37 respondents provided a score for "Others" and most attributed a moderate level of burden to this category. The overall average score was 2.95.



# Institutional Burdens of FISMA Compliance

- With regards FISMA requirements, respondents assessed the burden that was/is/will be imposed on their institution's resources to implement compliant plans:
  - Overall, all of the items were assessed on the higher end of the burden scale. Notably, no respondent gave any of these items the lowest burden score (1), and the majority of respondents assessed each item at either a 4 or 5.
  - "Funding required to build compliant infrastructure" received the highest average score (4.30).
  - "Staff effort/hours required for implementation" and "Funding required to maintain compliant infrastructure" were almost equally ranked with average scores of 4.19 and 4.13, respectively.
  - "Staff effort/hours required for maintenance" received a slightly lower average rating of 3.88. Although this rating is lower than the others, it is still well above a moderate burden rating.



# DFAR 252.204-7012 and NIST 800-171

- With regards a plan for compliance with DFAR 252.204-7012 Safeguarding Covered Defense Information and Cyber Incident Reporting including NIST 800-171 Protecting Controlled Unclassified Information in Nonfederal Information Systems and Organizations
  - 7.9% (5 respondents) have a plan and have implemented the compliance requirements
  - 22.2% (14 respondents) have a plan but are still working to implement the plan
  - 42.9% (27 respondents) are currently working on developing a plan for compliance
  - 27.0% (17 respondents) do not have a plan for compliance and are not currently working on one
- Who feels the burden? Responses in relation to these requirements were very similar to responses related to FISMA. The greatest burden is on IT (4.27 average) with Research Administration and Faculty and Research Staff next with averages of 3.8 and 3.62, respectively.



# Institutional Burdens of CUI Compliance

- With regards the DFAR 252.204-7012 and NIST 800-171 requirements, respondents assessed the burden that was/is/will be imposed on their institution's resources to implement compliant plans:
  - Responses were very similar to those regarding FISMA and the majority of respondents rated each item either a 4 or a 5. Interestingly, the burden to implement and build a compliant infrastructure is anticipated to be slightly higher than the burden of maintenance (but not by much).
  - "Staff effort/hours required for implementation" and "Funding required to build compliant infrastructure" were both rated the highest, with an average score of 4.19.
  - "Funding required to maintain compliant infrastructure" had a slightly lower average rating of 4.
  - "Staff effort/hours required for maintenance" received the lowest average rating (3.72), but this is still higher than moderate burden.



# General Comments

- “Collating all the standards has been the most difficult part - any external guidance in grouping/classifying these would be extremely helpful to institutions. We can then focus on the actual governance and implementation.”
- “We have challenges with the inconsistencies across federal agencies with data use and storage requirements.”
- “In theory, data management plans will tie to proposal budgets in a reasonable way, so adequate resources should be available. Realistically, though, PI's might be underestimating the time and expense of data management in some cases. And because these costs often come late in the project cycle, funds may already have been spent on other activities. Anticipate some PI's coming to their department, college, office of research, or the library for help.”



## General Comments (cont.)

- “Academic institutions must prepare for a shift unlike many things that Universities have undertaken in the recent past, towards a new infrastructure (technological, political, economic, and cultural) to support solutions to these new and changing requirements. National Information Standards Organization (NISO) has been working to create recommended standards (e.g., data citation, metadata, and networked data). Such work is slow.”
- “These requirements appear to have come on rather sudden and did not give universities adequate time to reorganize their IT systems and research administration procedures to prepare in advance. Given the amount of funding that is required to be invested to remain compliant, these requirement pose a significant burden.”



# Next Steps?

- In terms of where FDP can be of most help, “Inconsistent data management plan and data sharing requirements across Federal agencies” was the clear winner.
- Where to start?
  - Matrix of requirements of member agencies?
  - Other suggestions?