Opening Statement

Rep. Daniel Lipinski (D-IL)
Ranking Member
Subcommittee on Research
House Committee on Science, Space, and Technology

Committee Hearing:
Scientific Integrity and Transparency

March 5, 2013

Thank you Chairman Bucshon and thanks to all of the witnesses for being here.

The U.S. has for decades represented the world's gold standard for scientific integrity. But no one should mistake this observation as an argument for complacency. In the COMPETES Act of 2007, which we worked on in this subcommittee, then- Subcommittee Chairman Brian Baird included a provision on Responsible Conduct of Research that required every institution receiving NSF grant funding to provide training on the ethical conduct of science to all students and postdocs covered under those grants. Today, all U.S. research universities have implemented research ethics training for their STEM students and trainees.

The bigger challenge to the progress of science is not misconduct, but rather poor methodology and bad statistical analysis that take a long time to uncover. Or for that matter, discoveries in one field that have broad multidisciplinary relevance but take time to be known in other fields. To that end, the open sharing of scientific data is good for science and it's good for society. We must, of course, respect issues of privacy and intellectual property. But the more data are open, the faster we will validate new theories and overturn old ones, and the more efficiently we will transform new discoveries into innovations that will create jobs and make us healthier and more prosperous. The movement toward open data is not primarily about scientific integrity, it's mostly about speeding up the process of scientific discovery and innovation.

However, there are some big challenges to the widespread implementation of open data. Someone must define what exactly data sharing is going to mean and how it is going to be done, beginning with a standard. The February 22nd OSTP memo on increasing access to the results

of federally funded scientific research, which by the way was also a direct response to a requirement in COMPETES, takes on many of these issues in detail.

Specifically, we must consider such questions as:

- What does it entail and how much does it cost for researchers to develop a data management plan and to prepare their own data for sharing? Do they have adequate assistance from professional information managers?

- Are funding agencies sufficiently aware of the costs and skills required for good data management plans, and how should they evaluate and budget for data management proposals?

- What are the IT infrastructure needs for data-sharing, including technical standards, and what, if any, scientific or technical barriers exist to developing that infrastructure?

- What are the most important factors to consider in the economics of digital data access and preservation?

- What should be the respective roles of science agencies, universities, and the private sector in supporting and preserving public databases? How can these groups work together to minimize costs and maximize benefit to the scientific community?

- And finally, are there any policy or legal barriers for sustainable digital access and preservation?

In light of the Majority's suggestion of a possible legislative outcome for this hearing, I hope that today's dialogue will include a thoughtful discussion of some of these practical issues of implementation. I know that all four expert witnesses before us have a lot to contribute to this discussion and I look forward to learning from them.

With that I yield back.