



**Interview with
Dr. Carrie Wolinetz
US National Institutes of Health**

**For podcast release
Monday, February 1, 2021**

KENNEALLY: A year after the novel coronavirus, SARS-CoV-2, appeared, vaccination programs are already underway to protect millions of people around the world. Drug development at this remarkable speed depends on access for the global community of researchers to the very latest science. A new policy for data management and sharing of the US National Institutes of Health, the world's largest funder of biomedical research, will ensure such data is open and discoverable.

Welcome to Copyright Clearance Center's podcast series. I'm Christopher Kenneally for Velocity of Content. Dr. Carrie Wolinetz is acting chief of staff as well as associate director for science policy and director of the office of science policy at the National Institutes of Health. Dr. Wolinetz advises the NIH director on science policy matters of significance to the agency, the research community, and the public, including data management and sharing policies. She joins me from Maryland in the United States. Welcome to the program, Dr. Carrie Wolinetz.

WOLINETZ: Thank you, Chris. I'm delighted to be here.

KENNEALLY: We look forward to exploring with you the recently released final NIH policy for data management and sharing. This came out at the end of October of 2020. It's all part of the agency's efforts to make the results of publicly funded biomedical research broadly available to the public – no small matter, as the NIH invests almost \$42 billion a year in medical research. So it will mean this policy will have wide impact on science and public health in the United States and around the world. What's behind the policy, Dr. Wolinetz? Why does NIH want the global science community to share scientific data openly?

WOLINETZ: Well, we, too, are very excited about this policy. It's been a long time coming. We see responsible data management and sharing to be part of the important stewardship of the biomedical research enterprise. It underlies the advancement of science. It ensures that we are able to advance rigorous and reproducible research, that we can validate results. And we think it promotes public trust in research by fostering transparency and accountability. It maximizes the contributions of human volunteers who participate in biomedical research by ensuring that their data is responsibly accessible and maximizes



their time and contributions to the enterprise. So we really think it's an important foundation for the advancement of science.

KENNEALLY: Will this policy change the way that research is conducted? If so, how?

WOLINETZ: We really hope that this actually incites a cultural revolution in science in many ways. We are hoping that data sharing and responsible management becomes very much part of the fabric of the conduct of science, so that when people are thinking about their scientific questions and designing their experiments, right from the start, they're also thinking about how are they going to make sure that that data is organized and managed and accessible and that the metadata associated with it is accessible very much as part of the fabric of conducting the research itself. We also hope that this broader accessibility of data will help ensure more rigorous research, will increase the overall quality of science, and increase the innovation from the use or reuse of data that is more broadly accessible.

KENNEALLY: The policy we're speaking about, the process for developing it has been underway for nearly 20 years. Who contributed to the process – which particular stakeholders? And how did NIH conduct that activity?

WOLINETZ: It might be easier to ask who hasn't contributed to this process. (laughter) It has been a long time in coming. NIH has been at the forefront of data sharing in many ways in the life sciences for a very long time. We really set a lot of standards with our genomic data sharing policies and expectations. Then, we've moved into this broader policy, which really applies to all of the different types of science and areas of research that we fund.

We've heard from stakeholders internal to NIH, external to NIH. We've heard from research participants and patient advocacy organizations, institutions, scientists, folks who are experts in data science – really a broad range of stakeholders over the course of many years as we've put out elements of this policy for public comment and input and feedback. And we've really benefited a great deal from that feedback.

I think at the end of the day, when the policy was released, we heard folks who said we haven't gone far enough, and we also heard from folks who said maybe we've gone too far, which leads me to believe we probably got it right by striking that balance in the middle.

KENNEALLY: One of the balances that you're after is to give an even opportunity to both have data transparency as well as respecting concerns for privacy and protection. How do you work to get that right?



WOLINETZ: That's exactly right. It's not just about data sharing. It's about responsible data sharing. That means making sure that when the data involves human participants, particularly if it's identifiable data, that you have the appropriate protections around that. Making data accessible doesn't mean that you're posting it on the internet for the world to see. There are well-established models for accessibility – controlled access, accessibility to data when there are issues of privacy or identifiability. There are established procedures for deidentifying datasets as well. So we think that it is entirely possible to make sure that the data is accessible and usable for science without violating any of the privacy or confidentiality obligations that we have to participants in research.

KENNEALLY: This new policy requires researchers to prepare and file with NIH a data management and sharing plan. What are the essential elements of such a plan?

WOLINETZ: That's exactly right. We recognized very early on that it would be very difficult given the diversity of data types that NIH supports to have a one-size-fits-all, thou shalt share thy data exactly this way type of policy. So instead, what the policy requires is for investigators to submit a plan telling us the how, the where, the when, and the details of their data management and sharing policy that makes sense for the type of experiments and research that they're conducting.

Within that data plan, we expect to see things like the data type they expect to be generating, related tools, software, or code associated with that data, standards associated with the data, any timelines related to data preservation or access, any additional access, distribution, or reuse considerations – for example, whether there are issues related to informed consent or privacy and confidentiality, vulnerable populations, any technological considerations that might need to be considered. Basically, everything that helps to identify, again, the how, the where, the when of sharing data and allows us as an agency to be able to assess whether that plan makes sense.

I'm just going to add finally that it's important to note that investigators will be able to update these plans, because we recognize that science is dynamic, and it changes throughout the course of the research timeline. So investigators, if their science shifts, will also be able to shift appropriately their data sharing and management plans.

KENNEALLY: Nevertheless, it does not sound like any small task to put together these kinds of plans and to keep them updated. So I have to ask whether there are certain communities, certain researchers, that are better prepared than others for this policy and to put it in practice.

WOLINETZ: Absolutely. I think there are some scientific fields that are way out on the leading edge of data sharing. I've already mentioned genomics, which has long had a culture of



data sharing, and there are a lot of well-established repositories and cultural practices. Other fields are going to be less prepared, which is why we have given a very long implementation window here.

This policy, although it's out now and it's final, it's not actually effective until the end of January 2023. So there's a lot of time here, and we are trying to use this time thoughtfully to make sure that particularly for those segments of the community where this might not be part of their current culture and practice, that they have time to get their ducks in a row, get the tools that they need in place, the guidance that they need, in order to be able to implement this smoothly.

KENNEALLY: What will NIH do to assist those researchers and communities to comply with the policy?

WOLINETZ: We've put out a lot of supplemental information already – things about costs associated with data sharing, for example, guidance to help investigators think through what constitutes an appropriate repository or criteria for selecting a repository. And we expect as we do this outreach and engagement with the community and hear back from them about what they need and what would be most useful to continue to release information like that.

We're also working with groups like the National Academy of Sciences on workshops to deal with things like the cost of data sharing and establishing reasonable costs. There are a lot of groups out there working on data standards, for example, that we're paying a lot of attention to. So we are continuing to try to make sure that we've got the tools and information and guidance available to help investigators and institutions and the entire community hit the ground running.

We're not expecting this to be absolutely perfect on day one. We expect that there's going to be a learning process here. We're aiming for culture shift, and that doesn't happen overnight, and we recognize that. But we want to make sure that the community is as prepared as they can possibly be.

KENNEALLY: That culture shift you speak of is part of the NIH's priority around open science. Can you tell us what some of those top concerns are when it comes to open science for NIH?

WOLINETZ: Well, we want to make sure that as a publicly funded research agency, we are being as transparent and adherent to both our mission and our obligation to the public as possible by being good stewards. Part of that is making sure that we have the mechanisms and the expectations in place for appropriate and responsible sharing of data. Part of that is



really making sure that in addition to the policy, that we're also thinking through the infrastructure piece here, recognizing that we don't want to create the analogous situation of 1,000 channels on satellite TV and nothing to watch.

We're not looking for sort of a cacophony of open science here. We want to make sure that we're putting in place the infrastructure, whether that's in the form of repositories or data standards and expectations, that allow data sharing that will ultimately move science forward.

KENNEALLY: Dr. Carrie Wolinetz, acting chief of staff as well as associate director for science policy and director of the office of science policy at the National Institutes of Health, thank you for speaking with us today about the new policy for data management and sharing.

WOLINETZ: It's been a pleasure. Thank you.

KENNEALLY: Our co-producer and recording engineer is Jeremy Brieske of Burst Marketing. You can subscribe to the program wherever you go for podcasts and follow us on Twitter and Facebook. I'm Christopher Kenneally. Thanks for listening and join us again soon for another Velocity of Content podcast from CCC.

END OF FILE