In a world of “alternative facts,” scientists and their very methods feel under attack. With the last March for Science and the People’s Climate March both taking place in the past two weeks, we see a perhaps unprecedented political mobilization of scientists and science supporters. We also see many stories of coordinated, grassroots efforts to duplicate or rescue data from government web sites to prevent potential politically-motivated data removal. Few commentators have drawn lines between these initiatives, but they are at heart intimately connected. To support science is to support data, both the collection of new data, and the protection of data already held by government-funded organizations. The public and policy makers need to understand what it really means to support data - ensuring that we continue to have data that are available and useful both now and in the future.

The March for Science included one policy-explicit call on its web site, namely for “political leaders and policy-makers to enact evidence based policies in the public interest.” The key phrase here is “evidence based.” In the context of science, if somebody is talking about evidence, they are talking about data. We have overwhelming evidence of climatic change at both global and regional scales in the form of observations, computational simulation results, and combinations of the two. In another example, Thomas Piketty’s book Capital in the Twenty-First Century presented evidence of changes in wealth distributions based on data related to a variety of economic factors. As both of these cases show, having data does not mean that people can’t debate what those data mean, or what to do with them. But having openly available data that can be used as evidence is a prerequisite for any “evidence based” policy making.

In another case, the People’s Climate March web site lists a lengthy and ambitious agenda encompassing greenhouse gas reduction, energy industry transition, and minimum wage increases, among other things. How do you determine whether greenhouse gases are increasing or decreasing? How do you determine the economic impacts of new forms of energy production or increased minimum wages? By gathering, documenting, and analysing data, and then putting the data in a reliable place where anybody can look at them.

The United States government has been one of the global leaders in the production, collection, and distribution of scientific data for much of the past century. Numerous data centers based in federal science agencies – including NASA, NOAA, USGS, DOE and the EPA, among others – provide access largely without restrictions to thousands of data sets on a wide range of topics, from air quality and weather patterns to earthquakes and building energy usage. Many of these government data centers have existed for decades, and are part of national or international collaborations that exist to collect, compile, and distribute data worldwide in a consistent fashion.

Scientific data centers within the US government emphasize coordination, sharing (of data and expertise), and openness. Some undergo formal certification as trustworthy data centers, and

most follow well-established best practices for keeping track of data, like creating geographically distributed copies of their data, and creating robust documentation that describes how data were collected and what (if any) problems might be known to exist within the data. These best practices are specifically designed to prevent loss of data for any reason, whether technology or human errors, natural disaster, or political manipulation. Government open data policies reinforce the need for high quality data centers, and knowledgeable data center staff.

New efforts to rescue government data may not account for the established data center approaches to mitigating risk of data loss. Many data sets that have been “rescued” via recent grass-roots efforts are not in any risk of being lost by their government data centers, and are already duplicated in multiple locations around the country and the world. This is not to say that all government-produced scientific data are managed in the same way. Many data sets are not provided the ideal amount of care, and efforts to rescue or duplicate them are very important. Like everything else, data archiving takes money, and agencies allocate funds to data centers based on priorities from Congress and input from their relevant scientific expert communities.

The grass-roots “data rescue” efforts of the past few months emerged out of the fear that the new US presidential administration would destroy scientific data collections, either by directly ordering data deletion or by defunding data archives. These are not idle fears. The recent Canadian government of Prime Minister Stephen Harper defunded scientific data centers and libraries, ostensibly as part of cost-cutting measures. According to a September 2015 report by the Canadian news magazine Maclean’s, data about pollution in the Great Lakes, Canadian household incomes, and fishing stocks, among other data records, were lost or made inaccessible as a result of government cuts. This should not be repeated in the United States. Cuts to government data collections would represent the destruction of decades of economic investment, and would be a critical blow to numerous industries. The weather forecasts we use everyday are based on NOAA and NASA data. The agriculture, energy, and transportation industries likewise use government data extensively. Local governments’ ability to prepare for and react to natural disasters like earthquakes and floods depend on government-produced scientific data. Innumerable other examples exist of the critical economic and social value of scientific data provided by the US government.

Science is not something that only exists in the heads of scientists. Science rests on a base of data, the best evidence we have for how our world and society work. The scientific process, which critically involves the sharing of data and results, leads to medical and technological achievements that improve the standard of living for everybody. There cannot be *evidence-based* policies or concerted action in addressing our many societal challenges without data. March organizers, industry leaders, science advocates, and citizens contacting their congressional representatives should make this explicit.