Commentary on Sustainability and Infrastructure to Support Research

Richard B. Rood
September 20, 2008

Summary

Scientific investigation of climate change is a multi-investigator, multi-institutional, trans-disciplinary enterprise. Community-wide assessments of knowledge are a routine and necessary activity. Infrastructure to support scientific communities is not simply enabling; it is an essential element of scientific investigation. Infrastructure improves the ability of controlled experimentation and validation. Infrastructure enables investigator groups to leave a footprint of their research and deliberations. This allows transparency of process and validation, which improves the ability of others to evaluate and apply the knowledge that is generated by the science community. Infrastructure supports the communication of information from the confines of the science community to society as a whole. If well implemented, infrastructure reduces startup costs of investigations, enables the re-use of tools, promotes the sharing of intellectual capital, and facilitates collaboration across individuals, institutions and communities.

With these attributes, infrastructure is an element of sustainability. If the next student that comes along can assess the quality of tools developed in the research group, can trust the reliability of the data quality control, can rely on the information that describes the attributes of experiments, and can rely on the documentation that describes processes and applications, then the path for that student to produce new knowledge is eased. The same is true for scientists, resource managers, policy makers, indeed, all who have a vested interest in reliable, quality-assured knowledge.

Commentary

In 2000 I was an author on a report delivered to the Office of Science and Technology Policy (OSTP) High-End Climate Science: Development of Modeling and Related Computing Capabilities. This report identified the lack of software infrastructure as the largest missing element of the U.S. efforts in high-end climate modeling. This report followed years of *ad hoc* and community efforts by scientists to develop infrastructure, which was often characterized as a quest to develop plug compatible models. These activities were reported in Kalnay et al. (1989) and Dickinson et al. (2002).

In my case, the 2000 report for OSTP followed several years of trying to develop more rigorous processes in the development of large software systems for global modeling and assimilation at the National Aeronautics and Space Administration. These efforts were highlighted with some successes and failures and many lessons learned. Following the 2000 report there have been several federally funded programs to support the development of elements of community infrastructure. This commentary follows from direct experience in Earth science modeling, and consultations with scientists in other fields. First, the motivations to build infrastructure are introduced. Then the controversial aspects of infrastructure development are introduced. Finally, an extension of these motivations to include infrastructure as part of robust scientific process and sustainability is introduced.

The Call for Development of Infrastructure

The call for development of infrastructure in the atmospheric science community has come from both the grass roots and agency management. Within the atmospheric science community there are those who not only called for the development of software and systems infrastructure, but invested time and resources in self-organizing activities.

Grass-roots call for infrastructure came from those who were interested in and responsible for developing atmospheric models. The motivation for infrastructure is a natural path for those interested in organization, especially if there are tasks that seem to be duplicated or, like calendar functions, are required by all of those interested in the field. The desire for infrastructure grows as collaborations increase. In weather forecasting the mutual desires of the research and the operations communities to migrate research to operations is an important factor (see: From Research to Operations in Weather Satellites and Numerical Weather Prediction: Crossing the Valley of Death). A natural and central focus to grass-roots infrastructure efforts are code sharing and re-use; hence, the articulation of plug compatibility as in Kalnay et al. (1989).

When faced with time and budget constrained development and management of large software systems to support weather forecasting, climate predictions, and data analysis an infrastructure that is far more extensive than required to support collegial code sharing and re-use is needed. Attention is brought to systems design, and design is defined by both application and the computational environment in which the system will be deployed. Development of multiple modules, *de facto* subsystems, by a team of people requires the definition of interfaces, both technical and social. Verification and validation plans are needed to both assure that the software does what it is expected to do and that the software works in highly customized computational environments. Infrastructure grows to include not only the tools that support code development and testing, but also to include the formal definition of process. The tools and process of software development have been formalized as software engineering. The advocacy of infrastructure at this scale follows from those responsible for the management of large systems.

From the level of those who sponsor research and development applications as well as manage large national programs comes another call for the development of infrastructure. There is at this level the apparent duplication of effort across many groups. In addition there is the constant need for significant resources to keep software systems viable on evolving, exotic hardware systems. At this program level there is the expectation that the software systems that they sponsor will deliver, for instance, validated simulations to support the assessment of climate change.

The motivation to develop infrastructure, therefore, comes from top to bottom. This sits in relation to the resistance of the development of infrastructure that comes from organizational, resource, technical, sociological and emotional sources. The development of effective infrastructure is the development of an evolutionary, large-scale system. This system must have buy-in from users and the process to tie together users and developers. It is not easy to develop an effective infrastructure; it requires sustained investment and the integration of many varieties of expertise, some scientific, some not.

The Controversies of Infrastructure

The call to develop infrastructure in the atmospheric science community came from a subset of scientists who perceived value in the development of infrastructure. This subset of scientists did not represent the voice of the entire community. Efforts to build infrastructure reveal a deep resistance to building infrastructure, even amongst institutions that are advocates of infrastructure building.

Some pieces of infrastructure to support scientific investigation manifest themselves as facilities, and these often have broad community support. These projects are funded and built on the potential of the unique characteristics of these facilities to support discovery of fundamental knowledge. Examples include telescopes to support astronomy, accelerators to support high energy particle physics, satellites to support Earth observations, and supercomputers to support a wide range of computational-based investigation. These expenditures are not without controversy, because their cost often represents a major fraction of a field's funding "allotment." Members of the field can and do make cogent arguments that concentration of funds in these centralized facilities is risky business. More abstract arguments are anchored around the unpredictable nature of discovery and the human contribution to synthesis, innovation, and breakthroughs. An additional controversy arises, especially in the case of high-performance computers, about the value of curiosity-driven research versus the routine or operational generation of products.

Expenditures of infrastructure that are not tangible, community facilities are subject to the same skeptical eye and controversy. This type of infrastructure is most often anchored in information technology. It is easy to cite examples of computational and data systems that have been ineffectively designed, implemented or executed. These examples provide fuel for controversy.

The basis for some skepticism is principled and philosophical – anchored in belief as described above. Another class of skepticism arises from more concrete considerations. Many of these are based on the fact that, at least initially, the participation of a research group in a community activity is costly. It might require using tools that are not familiar to a group, or that a mature group will already have customized. If there is the requirement that the people in the community must be totally responsible for the development of the infrastructure, then this is a direct cost of effort. There are few groups with resources (fiscal, skill base and intellectual) on the margin to carry out such a development. Thinking more broadly, a community that spans across disciplines will require individuals to participate in activities that are not of apparent benefit to their careers or interests. Then there are issues of evolving, maintaining and servicing infrastructure tools.

More subtle are arguments that IT infrastructure cannot be built to support scientific research. Often cited in these arguments is the unpredictable nature of scientific research, or broadly, research in general. It is stated that infrastructure development imperils the integrity of "the science," leading to the conclusion that scientific infrastructure must be built exclusively by scientists to protect this integrity. This point of view places the development of infrastructure to a position subsidiary to science. It excludes the expertise and the intellect of the non-scientist; they, too, are subsidiary.

These arguments that activities to develop software and systems infrastructure to support scientific investigation somehow imperil the integrity of scientific investigation point to an even deeper source of resistance to the development of infrastructure. This final source is anchored in the sociological and psychological nature of scientific research and researchers. In any individual researcher or existing group there is an implicit or explicit infrastructure. This represents how things are done, and often there is a history of arriving at how things are done. Challenging these established processes and procedures challenge the culture of organizations and the sensibilities of individuals; it is exquisitely personal. In fact, the development of community infrastructure might require competitors to come to agreement in a way that is perceived as an impossible compromise. There would be winners and losers, and the losers, implicitly and instinctually, giving up something to the winners. The development of community infrastructure is as much a problem in sociology as technology.

Infrastructure and Scientific Method

As described above the development of infrastructure to support scientific investigation is construed as tools developed to support efficiency in the execution of research. This subsection asserts that in the case of collaborative science, infrastructure supports that robustness of the scientific method.

For many years there have been assessments of state of the knowledge in, for example, our understanding of ozone depletion and climate change. While often viewed in the community of scientists as a programmatic or a societal burden on scientific research, it can be argued that these community activities represent the unifying characteristics of scientific investigation that coexist with the reductionist investigation. One of the challenges faced in assessment is precise knowledge of the heritage of data sets, whether observations or simulations. Hence, infrastructure that supports management of information that provides accurate and detailed descriptions of data, metadata, is a natural part of the inventory or taxonomy of scientific information.

The role of infrastructure to manage information is intuitive. How might infrastructure be more deeply represented in scientific investigation?   In the development and execution of climate models elemental components are evaluated from geographically and institutionally diverse communities. Exercises to perform cause and effect experiments to evaluate these components are carried out. In the model-development centers there are multiple configurations of models and component models. These component models are constantly evolving; they are subject to documented changes and more subtle ancillary changes that are not documented. The cause and effect experiments take on the flavor of a “bake off” of model configurations. These configurations are evaluated against metrics or development priorities that are often defined through *ad hoc* processes. Infrastructure that helps to manage and document the changes in the development environment contributes to the performance of controlled experimentation. This helps to determine cause and effect, or minimally, to clarify arguments based differences in models and simulations that are presumed to conceptually equivalent.

As stated above, the definition of infrastructure grows out of being a collection of tools that ease the work of individuals and groups. Infrastructure also includes consideration of process and behavior – minimally, there are requirements of verification, validation, certification and documentation. The practices that assure the robustness of infrastructure are easily extended to the scientific process. At the core of the scientific process is validation of results. This validation has two essential aspects. The first is the validation against independent sources of information or data. The second is the independent evaluation of the results by other scientists. The infrastructure that supports controlled experimentation also supports transparency of the validation process – an essential element of the scientific process.

For a field like climate change where the knowledge generated by the field has, potentially, profound impact on the behavior of all society this transparency is a fundamental obligation. Infrastructure, therefore, evolves to be an element of the scientific method, contributing to validation.

Infrastructure and Sustainability

Infrastructure to support research is part of a sustainable enterprise. In a generalized sense infrastructure includes tools and behavior, best practices of tool use. As an element of sustainability image the following case. A student researcher needs to acquire a combination of observational and model-simulated climate data. The observations are from a ground station and contain observations of varying quality. The model data is from a suite of simulations with varying carbon dioxide concentrations, and there is a matrix of model parameters that characterize the details of the simulation. The student spends several weeks going through the observations to identify missing and obviously incorrect data. Several phone calls are required to a national modeling center to identify the parameters used in the simulation. The information about the data quality assessment performed by the student can be documented and archived as metadata with observational information. This information then serves to accelerate the research of the next student, and provides a transparent record that can be verified by independent researchers.

If the researcher leaves of footprint of their activity, then this helps to sustain the community. It eliminates a waste of energy; it improves the robustness of information. Traditionally, the product of research is a written report. The work of the researcher is, however, completely represented by a set of activities that include generation or collection of data, development of tools to extract information from the data, analysis to extract knowledge from the information, validation that the knowledge is robust, and reporting of results. Infrastructure that records the information and process of this set of activities, the workflow, accelerates the extraction of knowledge and is a foundation of sustainability.

In a field such as climate research, knowledge generated by scientific investigation has consequences for society as a whole. While the scientific report is the end product, and the research is subject to verification by other scientists, it is often the details of the workflow, and especially the details of verification and validation, that demand scrutiny as the science knowledge diffuses from the scientific community. Therefore, it is the responsibility of the scientist to provide accurate and transparent information of process, of their practice. The recording and documenting of this information as metadata stand to accelerate the use of scientific information. It helps to sustain the field and contributes to the culture of a sustainable society.

Summary

The scientific investigation of climate change is, today, a multi-investigator, multi-institutional, trans-disciplinary enterprise. Community-wide, assessments of knowledge are an essential element of the enterprise. Infrastructure is not simply enabling; it is an essential element of scientific investigation. Infrastructure improves the ability of controlled experimentation and validation. Infrastructure enables investigator groups to leave a footprint of their research and deliberations. This allows transparency of process and validation, which improves the ability of others to evaluate and apply the knowledge that is generated by the science community. Infrastructure supports the communication of information from the confines of the science community to society as a whole. If well implemented infrastructure reduces startup costs of investigations, the re-use of tools, the sharing of intellectual capital, and collaboration across individuals, institutions and communities.

With these attributes, infrastructure is an element of sustainability. If the next student that comes along can assess the quality of tools developed in the research group, can trust the reliability of the data quality control, can rely on the information that describes the attributes of experiments, and can rely on the documentation that describes processes and applications, then the path for that student to produce new knowledge is eased. The same is true for scientists, resource managers, policy makers, indeed, all who have a vested interest in reliable, quality-assured knowledge.

Points for further elaboration:

- Devaluation of infrastructure
- Haves and have nots
- Operational versus research communities
- Infrastructure and access to infrastructure as an equalizing fabric of society