**EarthCube GEO Domain Workshops**

Gathering Input to Address Geoscientist Cyberinfrastructure Needs

**Introduction:** For EarthCube to achieve its potential as a new data and knowledge management system for the 21st Century, the collective needs and desires of geoscientists across the disciplines must be made known so similarities and difference between user groups and disciplines can be identified and addressed. These workshops are being held to provide information so EarthCube can be designed to help geoscientist more easily do the science they want and need to do. Its goal is to provide them with access and tools so they can take leadership roles and be more competitive in the increasingly diverse and interdisciplinary scientific world in which we live. To this end, EarthCube is funding workshops so geoscience communities can articulate and document their cyberinfrastructure needs and what it is they would like to do in the future, both within their disciplines as well as in terms of accessing data and information from other fields, geoscience or not. This process is termed “requirements gathering”. The science-drivers, data utilities, user-interfaces, modeling software, tools, and other needs identified in this process will help direct and shape the function and form of what EarthCube will be.

**Provisional Workshop Goals:** EarthCube GEO domain workshops are designed to listen to the needs of the end-user groups that make up the geosciences and to understand better how data-enabled science can help them achieve their scientific goals. It is hoped that these workshops have a significant contingent of early career scientists since they will be the ones that EarthCube and the new world of data-enabled science is most likely to impact in terms of building their careers and creating educational materials.

The format is simple and workshop participants have the freedom to tell the EarthCube team what they feel should be communicated in terms of their community’s needs. Please be sure that each workshop understands that it should have some sort of virtual component so additional input from people who might not be able to come but are interested can have input. Please make sure that the outcome of the workshop includes the following, in addition to the other goals your community might have:

1. A list of important science drivers and scientific challenges that the community would like to tackle (or tackle more easily) both now and in the future, within a 5 to 15 year time frame).
2. within the discipline
3. across disciplines ***within the geosciences*** (also include, if possible, an ordering of which geoscience sub-disciplines are of the highest immediate priority to the group and which are of intermediate and lower priority) – please specify if this was determined on the collective interest of the parties involved in the workshop, on the importance of the scientific question, or on some other basis.
4. across disciplines ***outside the geosciences*** (if possible with an ordering of which geoscience sub-disciplines are of the highest immediate priority to the group and which are of intermediate and lower priority) – please specify if this was determined on the collective interest of the parties involved in the workshop, on the importance of the scientific question, or on some other basis.

This last item (i.e., c above) is important because most new money at NSF is coming in via inter/cross interdisciplinary initiatives and geoscientists being able to cross disciplines and do data enabled science with data from other fields and on topics of societal relevance is becoming an increasingly more important part of the NSF science portfolio.

1. As complete a list as possible of what participants feel is preventing them, at present, from being able to do the kind of things they want to do in research and education (in terms of data discovery and access; data integration; modeling; differences in scale in terms of time, space, or data density/abundance - e.g., fire hoses of data vs. discrete sample based datasets; difference in types or structure of datasets/bases; user interfaces, shared workspaces, etc.). Examples illustrating the problem(s) would be useful.
2. within the group’s discipline and ***within other geoscience disciplines*** from which the group wants to access data/models/tools, etc.
3. across disciplines ***outside of the geosciences*** from which the group wants to access data/models/tools, etc.
4. As comprehensive a list as possible of all known community data and modeling resources including present databases and where those can be found/accessed; model libraries, if any; essential tools or software already in use for data management and visualization; frequently used web portals where data/tools/utilities can be found; highly used/useful collaboration software and where it can be accessed; other additional frequently used resources for any aspect of being able to do data-enabled science.
5. A list of tools, databases, modeling capabilities, etc. that workshop participants, as representatives of their respective communities, feel need to be developed/created/made easier in terms of cyberinfrastructure and/or data management infrastructure/tools/software/etc. to allow them to do the important science and education that they want to do now and in the future.
   1. ***within the group’s discipline and within other geoscience disciplines*** from which they want to access data/models/tools, etc.
   2. ***across disciplines outside of the geosciences*** from which the group wants to access data/models/tools, etc.
6. Three or more strong use cases with important/transformative science drivers (a template and characteristics for these will come from the funded groups). Also any additional information that presently funded EarthCube groups need from end users. Information on this will be provided to meeting organizers as it comes available over the next month.