NSF at work, sometime in the twenty-tens.
By George O. Strawn, CIO

Business and Operations Advisory Committee
November 19, 2008

As always, researchers and educators respond to NSF program announcements and solicitations. (By the way, NSF now has a $15B annual budget, so response to our announcements is greater than ever.) These e-documents are regularly pushed to the community via RSS feeds as well as being readily available on nsf.gov, research.gov and grants.gov. In these days of federated identity management, researchers submitting proposals to NSF are authenticated to NSF by their home institutions and do not need separate NSF identity credentials (one small step for humans, one giant leap for humankind).

NSF now provides more detailed templates for proposals. These templates facilitate proposers providing *annotations* about their proposals, which in turn facilitate *text processing* (as described below). This format also makes reviewer processing more effective and efficient by providing them with more uniform and transparent proposals.

One type of text processing is that the proposals associated with each call are clustered into panel-size groups based on their similarity. Another type of text processing occurs as each proposal is compared to all previously submitted proposals *and* to all published scholarly literature. These comparisons identify the proposals and articles most alike in content to each new proposal.

Then the authors of these similar proposals and articles are identified and their credentials are presented to the cognizant program officer as possible panelists. After review of these and any other candidate panelists' credentials, the program officer indicates which ones to invite, informing them of the panel dates and other details. Panelists who accept may choose to come to NSF for the panel *or* to remain at their institution and utilize telepresence technologies. (Btw, half of the NSF program officers are now "virtual" in the sense that they usually remain at their host institutions rather than come to NSF.)

Scheduling of panels and panelist logistics are now largely automated. This begins with the management of the panelist invitations and rsvp's. As a panel takes shape, travel arrangements are put in place and most conflicts of interest are determined and noted for panel operations.

Panelists have an easier time reading the annotated proposals and they similarly utilize a review template with annotation capabilities, which simplify the subsequent task of the program officer in preparing the review summary (which also has template and annotation aids). It has been recently estimated that the panel process is now twice as effective in producing high quality reviews and three time more efficient in the use of panelist and program officer time.

The scholarly literature resulting from NSF awards has recently been made web-available to the public. This literature is also annotated, making text processing practicable here as well. (Of course, the NSF award abstract public database with award abstracts and pointers to resulting scholarly literature and been annotated for some time now.) As an example of an early use of this annotation, text processing of chemistry articles for pharmaceutical purposes is now in its second decade. Such processing is given credit for first pointing to many compounds that have subsequently become wonder drugs. And text processing is now increasingly used for interdisciplinary research projects.

The annotated proposals and literature also help NSF assess progress at the leading edge of science and engineering research. We are better able to sharpen our program announcements and solicitations with an eye on this evaluation of broad disciplinary progress, which even the specialists on the ground may not have. This broad view also gives us more capability to describe to downtown decision makers and to the public what is happening at the cutting edge science and to explain how NSF can best support transformative science.

Interestingly, all of these advances have been accomplished with an ever-decreasing amount of IT hardware at NSF. "Cloud computing" has matured to the point that all of NSF's software and data are stored on vendors' "clouds". Security and privacy, always good at NSF, have actually been enhanced, NSF floor space has been released for other purposes, and our availability of services (ie, uptime) has been enhanced well beyond what Ballston area electricity transformers could ever do.

Shortly after achieving this "complete automation" of NSF's traditional way of achieving its mission of finding and supporting the most promising s&e projects, a revolutionary idea was presented to senior management for an entirely new way to find and support best projects, enabled by the new IT capabilities that have recently become become available. But I see my time is up, so I'll have to delay discussion of that revolutionary approach for another day.