Cost-Benefit Analysis of the RFA

THE U.S. NATIONAL INSTITUTES OF HEALTH issues requests for applications (RFA) to solicit proposals on a specific topic. While a well-designed RFA can have significant benefit to the scientific enterprise and support the mission of the Institute, a poorly designed RFA can produce a significant loss of scientific effort.

The benefit of an RFA can be estimated by multiplying the number of teams funded by the duration of support. The costs associated with an RFA are more diffuse, but just as real. Teams prepare proposals, diverting effort from ongoing projects. Reviewers evaluate proposals, again diverting effort. University staff review budgets and deal with regulatory approval, increasing overhead rates. Institute staff attend study sections and prepare summary statements, consuming resources. Our conservative estimate is that each proposal costs two months of team effort in preparation and review.

Unfortunately, some RFAs have much greater cost than benefit. As a recent example, the National Institute for Biomedical Imaging and Bioengineering issued a “Quantum Projects” RFA. In response, 89 proposals were received with only one grant funded. The benefit of this RFA was support for three years of scientific effort. The cost of this RFA was nearly 15 years of lost effort. This RFA resulted in a net loss of 12 years of scientific effort.

Institutes can improve the cost-benefit ratio of RFAs. Sufficient resources must be invested in a net loss of 12 years of scientific effort.

Finally, Institutes should employ pre-proposals to screen applications and minimize the number of full proposals required for preparation and review. (Pre-proposals are used by NSF and also a few NIH programs. They are much shorter and require much less effort than a full proposal.) Finally, Institutes should publish the number of proposals received and the number of grants funded to guide response to future RFAs.

Data Mining on the Web

WE READ WITH GREAT INTEREST THE PERSPECTIVE “Creating a science of the Web” by T. Berners-Lee et al. (11 Aug., p. 769). We agree that evolving Web technologies enable the creation of novel structures of information, whose properties and dynamics can be fruitfully studied. More generally, we would like to point out that the Web is a specific phenomenon associated with the increasing prevalence of information being digitized and linked together into complicated structures. The complexity of these structures underscores the need for systematic, large-scale data mining both to uncover new patterns in social interactions and to make discoveries in science through connecting disparate findings. For this vision to be realized, we have to develop a new science of practical data mining focusing on questions answerable with the existing digital libraries of information. In particular, today, free-text search (as embodied by Google) is the primary means of mining the Web, but there are many kinds of information requests it cannot handle. Queries combining general, standardized annotation about pages (such as from the semantic Web) with free-text search within them are often not supported—e.g., doing a full-text search of all biophysics blogs emanating just from governmental institutions within 100 miles of Chicago. Furthermore, it would be useful to develop ways of leveraging the small amounts of highly structured information in the semantic Web as “gold-standard training sets” to help bootstrap the querying and clustering of the large bodies of unstructured information on the Web as a whole. Thus, the science of the Web should enumerate the range of information requests that can be fruitfully made and the kinds of information infrastructure and data mining techniques needed to fulfill them.

ANDREW SMITH1 AND MARK GERSTEIN2

1Department of Computer Science, Albert Williams Professor of Biomedical Informatics, Yale University, New Haven, CT 06520, USA.
2Department of Chemistry, University of Washington, Seattle, WA 98195–1700, USA.

Response

WE AGREE WITH SMITH AND GERSTEIN’S VIEW that data mining is among the many important areas of research that are considering the Web as an object of scientific inquiry. They are correct in pointing out the importance of “text mining,” the basis of current Web search, for providing new Web capabilities.

However, with the increasing amount of directly machine-readable data that are available on the Web (coming from, for example, database-producing equipment such as modern scientific devices and data-oriented applications), it is also clear that text mining needs to be augmented with new data technologies that work more directly with data and metadata. Data mining is also an excellent case in point for the main focus of our Perspective in relation to the interdisciplinary nature of the emerging science of the Web. Analytic modeling techniques will be needed to understand where Web data reside and how they can best be accessed and integrated. Engineering and language development are needed if we are to be able to perform data mining without having to pull all the information into centralized data servers of a scale that only the few largest search companies can currently afford. In addition, data mining provides not just opportunities for better search, but also real policy issues with respect to information access and user privacy, especially where multiple data sources are aggregated into searchable forms.

TIM BERNERS-LEE,1 WENDY HALL,2 JAMES HENDLER,3* NIGEL SHADBOLT,2 DANIEL J. WEITZNER1

1Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. 2School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK. 3Computer Science Department, University of Maryland, College Park, MD 20742, USA.

*To whom correspondence should be addressed. E-mail: hendler@cs.umd.edu

Using Models to Manage Carnivores

THE NEWS FOCUS ARTICLE “THE CARNIVORE comeback” (M. Enserink, G. Vogel, 3 Nov., p. 746) illustrates the difficulty of conserving free-ranging predators in highly anthropic landscapes such as Europe. Because large carnivores can cause heavy damages to livestock as well as threaten human beings, it is critical that management policies are flexible enough to allow for some removals while keeping...
Management recommendations would be much improved and accepted by the public if they were based on population modeling rather than on expert opinion consensus. Because models are logical constructions based on falsifiable assumptions, their recommendations can be invalidated, whereas expert opinions are verbal constructions difficult to refute. Fisheries management has made an extensive use of population models, but it has not been widespread (2). Designing efficient adaptive management schemes—i.e., implementing policies as experiments—should be achieved through a wider use of such models.

Because models are logical constructions based on patterns at the individual level (3), designing efficient adaptive management schemes—i.e., implementing policies as experiments—should be achieved through a wider use of such models.

We suggest that parthenocarpic or fertile fig branches were planted along with stipules like wild barley in the early Neolithic villages of Gilgal and Netiv Hagdud. In contrast to the repeated sowing of wild barley, we argue that planting branches of selected fig trees constitutes a form of domestication. The simplicity of fig tree propagation likely contributed to its domestication before cereal crops.

Full text at www.sciencemag.org/cgi/content/full/314/5806/1683a

RESPONSE TO COMMENT ON “Early Domesticated Fig in the Jordan Valley”
Mordechai E. Kislev, Anat Hartmann, Ofer Bar-Yosef

We suggest that parthenocarpic or fertile fig branches were planted along with stipules like wild barley in the early Neolithic villages of Gilgal and Netiv Hagdud. In contrast to the repeated sowing of wild barley, we argue that planting branches of selected fig trees constitutes a form of domestication. The simplicity of fig tree propagation likely contributed to its domestication before cereal crops.

Full text at www.sciencemag.org/cgi/content/full/314/5806/1683b

CORRECTIONS AND CLARIFICATIONS


Policy Forum: “Genomics and medicine at a crossroads in Chernobyl” by G. S. Ginsburg et al. (6 Oct., p. 62). In the first paragraph, in line 11, the phrase “1.1-billion-ton temporary ‘sarcophagus’” should instead read “1.1-million-ton temporary ‘sarcophagus.’”

Special Section on Migration and Dispersal: News: “Follow the footprints” by K. Unger (11 Aug., p. 784). In the article, tapirs are described as “piglike.” Although to the uninitiated observer, tapirs seem piglike, they are actually more closely related to horses and rhinos.