This is the first report of time-dependent seismic tomography applied to an erupting volcano. It builds on earlier work of the same kind done in geothermal areas in California and Iceland and the Long Valley Caldera, California. But the seminal example of major changes in V_P/V_S comes from The Geysers geothermal area in northern California.

During the 1980s and 1990s, some 13,600 tons of steam per hour were extracted from The Geysers to generate electricity. As a result of this overexploitation, the reservoir became progressively depleted as pore water was replaced by steam. Repeat seismic tomography showed the steady growth of a reservoir-wide negative V_P/V_S anomaly that coincided with the steam-production zone. This anomaly was caused by the combined effects of the replacement of pore liquid with steam, the resulting decrease in pressure, and the drying of clay minerals. A remarkable series of snapshots showed the relentless growth of a volume of heavy depletion (3, 4). The work helped to increase awareness of the nonsustainability of such high rates of fluid withdrawal. Production at The Geysers has now been reduced to sustainable levels. Time-dependent tomography is currently used to monitor the Coso Geothermal Area, southern California (5).

Time-dependent seismic tomography was first applied to a volcano in a study of Mammoth Mountain, a volcano on the rim of Long Valley Caldera, California. In 1989, an intense swarm of hundreds of earthquakes accompanied an injection of new magma into the roots of this volcano, and triggered the outpouring of some 300 tons of CO_2 per day from the volcano's surface. Several broad swaths of trees died as a result of high levels of CO_2 in the soil, and the CO_2 also presented an asphyxiation hazard to humans. A comparison of V_P/V_S tomographic images calculated for 1989 and 1997 showed changes that correlated well with areas of tree death on the surface above, and were attributed to migration of CO_2 in the volcano (6).

By showing that time-dependent seismic tomography can be used to monitor structural changes directly associated with a volcanic eruption cycle, Patanè et al. take a critical step toward developing a useful volcano-hazard-reduction tool based on seismic tomography. As with all good experiments, however, it ushers in new challenges. V_P/V_S is affected by several factors, including pore fluid phase, pressure, mineralogy, and fracture density. However, determining how each

of these has changed when changes in only two quantities (V_P and V_S) have been measured is not possible and requires the addition of other kinds of data. Both theoretical advances and more data from different volcanoes are needed before the potential of the method can be fully assessed.

At present, monitoring of active volcanoes still rests mostly on relatively unsophisticated seismic networks and the monitoring of simple parameters, such as the numbers of earthquakes and the amplitude of harmonic tremor. Patanè et al. show that much more sophisticated methods can now be used. Some of these methods only need to be automated—a critical factor if they are to be useful in situations where information is needed on an hourly basis. It is hoped that this automation work will be pushed forward rapidly in the near future, putting us on track to realizing technological capabilities resembling those of the fictional Virtual Geophysical Laboratory by 2025.

References and Notes
1. D. Patanè, G. Barberi, O. Cocina, P. De Gori, C. Chiarabba, *Science* **313**, 821 (2006).
2. The compressional and shear waves are the fastest and second-fastest waves to be radiated from an earthquake source, so they arrive first and second on seismograms. Their ratio provides information about pressure and about the presence of gas and liquid in the study volume. Thus, changes in their ratio can tell us about changes in pressure and gas/liquid, which are thought to accompany the buildup and occurrence of a volcanic eruption.
3. G. R. Foulger, C. C. Grant, A. Ross, B. R. Julian, *Geophys. Res. Lett.* **24**, 135 (1997).
4. R. C. Gunasekera, G. R. Foulger, B. R. Julian, *J. Geophys. Res.* **108**, 2134 (2003).
5. G. R. Foulger, B. R. Julian, *Geotherm. Resour. Counc. Bull.* **33**, 120 (2004).
6. G. R. Foulger et al., *J. Geophys. Res.* **108**, 2147 (2003).

COMPUTER SCIENCE

Creating a Science of the Web

Tim Berners-Lee, Wendy Hall, James Hendler, Nigel Shadbolt, Daniel J. Weitzner

Understanding and fostering the growth of the World Wide Web, both in engineering and societal terms, will require the development of a new interdisciplinary field.

Since its inception, the World Wide Web has changed the ways scientists communicate, collaborate, and educate. There is, however, a growing realization among many researchers that a clear research agenda aimed at understanding the current, evolving, and potential Web is needed. If we want to model the Web; if we want to understand the architectural principles that have provided for its growth; and if we want to be sure that it supports the basic social values of trustworthiness, privacy, and respect for social boundaries, then we must chart out a research agenda that targets the Web as a primary focus of attention.

When we discuss an agenda for a science of the Web, we use the term "science" in two ways. Physical and biological science ana-

Enhanced online at
www.sciencemag.org/cgi/
content/full/313/5788/769

lyzes the natural world, and tries to find microscopic laws that, extrapolated to the macroscopic realm, would generate the behavior observed. Computer science, by contrast, though partly analytic, is principally synthetic: It is concerned with the construction of new languages and algorithms in order to produce novel desired computer behaviors. Web science is a combination of these two features. The Web is an engineered space created through formally specified languages and protocols. However, because humans are the creators of Web pages and links between them, their interactions form emergent patterns in the Web at a macroscopic scale. These human interactions are, in turn, governed by social conventions and laws. Web science, therefore, must be inherently interdisciplinary; its goal is to both understand the growth of the Web and to create approaches that allow new powerful and more beneficial patterns to occur.

Unfortunately, such a research area does not yet exist in a coherent form. Within computer science, Web-related research has largely focused on information-retrieval algorithms and on algorithms for the routing of information through the underlying Internet. Outside of computing, researchers grow

T. Berners-Lee and D. J. Weitzner are at the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. W. Hall and N. Shadbolt are in the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK. J. Hendler is in the Computer Science Department, University of Maryland, College Park, MD 20742, USA. E-mail: hendler@cs.umd.edu
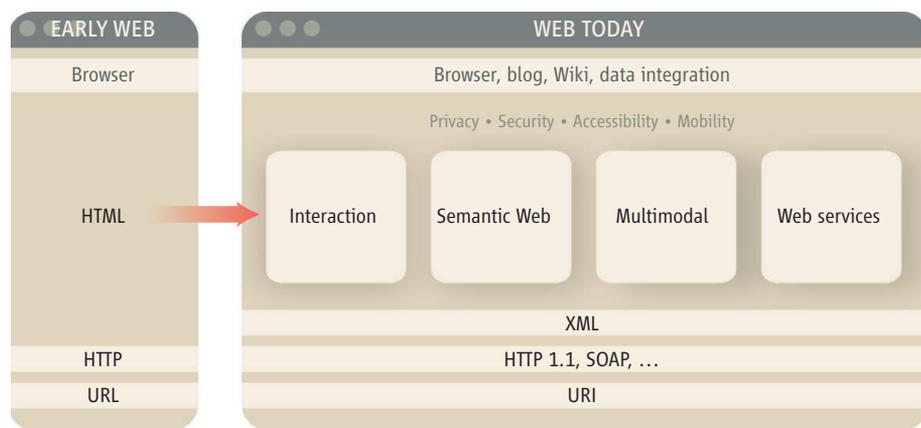
ever more dependent on the Web; but they have no coherent agenda for exploring the emerging trends on the Web, nor are they fully engaged with the emerging Web research community to more specifically focus on providing for scientists' needs.

Leading Web researchers discussed the scientific and engineering problems that form the core of Web science at a workshop of the British Computer Society in London in September 2005 (*1*). The participants considered emerging trends on the Web and debated the specific types of research needed to exploit the opportunities as new media types, data sources, and knowledge bases become "Webized," as Web access becomes increasingly mobile and ubiquitous, and as the need increases for privacy guarantees and control of information on the Web.

The workshop covered a wide range of technical and legal topics. For example, there has been research done on the structure and topology of the Web (*2, 3*) and the laws

given topic. Conventional information-retrieval techniques are insufficient at the scale of the Web. However, it turns out that human topics of conversation on the Web can be analyzed by looking at a matrix of links (*7, 8*). The mathematics of information retrieval and structure-based search will certainly continue to be a fertile area of research as the Web itself grows. However, approaches to developing a mathematical framework for modeling the Web vary widely, and any substantive impact will, again, require a new approach. The process-oriented methodologies of the formal systems community, the symbolic modeling methodologies of the artificial intelligence and semantics researchers, and the mathematical methods used in network analyses are all relevant, but no current mathematical model can unify all of these.

One particular ongoing extension of the Web is in the direction of moving from text documents to data resources (see the figure).

modeling. Imagine being able to query the Web for a chemical in a specific cell biology pathway that has a certain regulatory status as a drug and is available at a certain price. The engineering challenge is to allow independently developed data systems to be connected together without requiring global agreement as to terms and concepts. The statistical methods that serve for the scaling of language resources in search tasks and the data calculi that are used in scaling database queries are largely based on incompatible assumptions, and unifying these will be a major challenge.

Despite excitement about the Semantic Web, most of the world's data are locked in large data stores and are not published as an open Web of inter-referring resources. As a result, the reuse of information has been limited. Substantial research challenges arise in changing this situation: how to effectively query an unbounded Web of linked information repositories, how to align and map between different data models, and how to visualize and navigate the huge connected graph of information that results. In addition, a policy question arises as to how to control the access to data resources being shared on the Web. This latter question has implications both with respect to underlying technologies that could provide greater protections, and to the issues of ownership in, for example, scientific data-sharing and grid computing.

The scale, topology, and power of decentralized information systems such as the Web also pose a unique set of social and public-policy challenges. Although computer and information science have generally concentrated on the representation and analysis of information, attention also needs to be given to the social and legal relationships behind this information (*9*). Transparency and control over these complex social and legal relationships are vital, but require a much better-developed set of models and tools that can represent these relationships. Early efforts at modeling in the area of privacy and intellectual property have begun to establish the scientific and legal challenges associated with representing and providing users with control over their own information. Our aim is to be able to design "policy aware" systems that provide reasoning over these policies, enable agents to act on a user's behalf, make compliance easier, and provide accountability where rules are broken.

Web science is about more than modeling the current Web. It is about engineering new infrastructure protocols and understanding the society that uses them, and it is about the

The Web yesterday and today. (**Left**) The World Wide Web circa 1990 consisted primarily of text content expressed in the Hypertext Markup Language (HTML), exchanged via the hypertext transfer protocol (HTTP), and viewed with a simple browser pointing to a Universal Resource Locator (URL). (**Right**) Users of the Web now have a variety of top-level tools to access richer content including scalable vector graphics, the Semantic Web, multimodal devices (e.g., voice browsers), and service descriptions. These are expressed in extended markup language (XML), exchanged by newer protocols [e.g., HTTP 1.1 and SOAP (simple object access protocol)] and are addressed by uniform resource identifier (URI) schemes.

of connectivity and scaling to which it appears to conform (*4–6*). This work leads some to argue that the development of the Web has followed an evolutionary path, suggesting a view of the Web in ecological terms. These analyses also showed the Web to have scale-free and small-world networking structures, areas that have largely been studied by physicists and mathematicians using the tools of complex dynamical systems analysis.

The need for better mathematical modeling of the Web is clear. Take the simple problem of finding an authoritative page on a

In the Web of human-readable documents, natural-language processing techniques can extract some meaning from the human-readable text of the pages. These approaches are based on "latent" semantics, that is, on the computer using heuristic techniques to recapitulate the intended meanings used in human communication. By contrast, in the "Semantic Web" of relational data and logical assertions, computer logic is in its element, and can do much more.

Researchers are exploring the use of new, logically based languages for question answering, hypothesis checking, and data

creation of beneficial new systems. It has its own ethos: decentralization to avoid social and technical bottlenecks, openness to the reuse of information in unexpected ways, and fairness. It uses powerful scientific and mathematical techniques from many disciplines to consider at once microscopic Web properties, macroscopic Web phenomena, and the relationships between them. Web sci-ence is about making powerful new tools for humanity, and doing it with our eyes open.

References

1. Workshop on The Emerging Science of the Web, British Computer Society, London, 12 to 13 September 2005. See www.cs.umd.edu/~hendler/2005/WebScienceWorkshop.html
2. R. Milo *et al.*, *Science* **298**, 824 (2002).
3. R. Milo *et al.*, *Science* **303**, 1538 (2004).
4. A.-L. Barabási, R. Albert, *Science* **286**, 509 (1999).
5. J. M. Kleinberg, *Nature* **406**, 845 (2000).
6. S. H. Strogatz, *Nature* **410**, 268 (2001).
7. S. Brin, L. Page, in *Proceedings of the 7th International World Wide Web Conference* (Elsevier Science, Amsterdam, 1998), pp. 107–117.
8. Z. N. Oltvai, A.-L. Barabási, *Science* **298**, 763 (2002).
9. L. Lessig, *Code and Other Laws of Cyberspace* (Basic Books, New York, 1999).

10.1126/science.1126902

ASTRONOMY

Testing Star Formation Theory

Richard M. Crutcher

The importance of magnetic fields for the formation of stars, such as the Sun, is supported by measurements of polarized radio waves from dust particles near a newly forming star.

Understanding how stars form is one of the outstanding challenges of modern astrophysics. It has become clear that stars form from dense interstellar clouds of gas and dust, called molecular clouds because gas in such clouds is predominantly in molecular rather than atomic form. However, despite substantial progress in recent years, there remain fundamental unanswered questions about the basic physics of star formation. In particular, it remains unclear whether molecular clouds undergo rapid gravitational collapse as soon as sufficient matter accumulates to make the clouds gravitationally bound, or whether there is some mechanism resisting collapse that delays the process and introduces new star formation scenarios. The observational result reported on page 812 of this issue by Girart *et al.* (*1*) provides new data regarding this important scientific question.

The "standard" model for the formation of low-mass stars such as our Sun has been that interstellar magnetic fields provide support against gravity in dense molecular clouds (*2*). In this picture, interstellar magnetic fields are "frozen" into interstellar matter by the small fraction of the gas and dust that is ionized. As material accumulates (due to the driving of flows by galactic spiral-arm shocks, supernovae explosions, the gravity of a galaxy, etc.), the magnetic field increases in strength
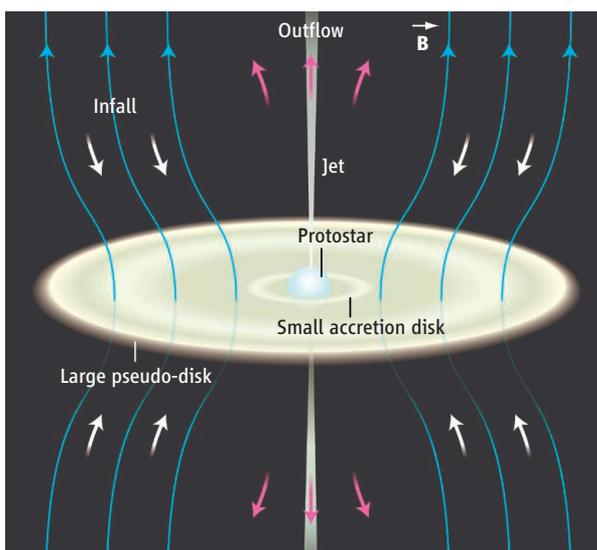
Shaped by magnetism. Schematic diagram of a collapsing molecular cloud core with a strong magnetic field (**B**) showing the characteristic hourglass shape. [Adapted from (*4*)]

as the gas density increases. After a molecular cloud accumulates sufficient mass to become self-gravitating, it will still not collapse and form stars because gravity is balanced by magnetic pressure.

If there were no other forces operating, molecular clouds would persist indefinitely and star formation would not occur. However, magnetic fields are frozen only into the ions of molecular clouds, not into the neutral gas and dust. The neutrals are therefore free to respond to gravity and collapse to form a much denser, gravitationally unstable core to the molecular cloud and eventually to form stars. However, as neutrals collapse through the ionized gas and dust, collisions with ions will occur. These collisions will greatly slow down the collapse rate, leading to molecular cloud lifetimes typically several orders of magnitude longer than the gravitational free-fall lifetime of a cloud.

In contrast to magnetically dominated star formation, the other extreme point of view is that magnetic fields are too weak to provide support against gravity. In this model, molecular clouds are intermittent phenomena, and the problem of cloud support for long time periods is irrelevant (*3*). Supersonic flows in the low-density turbulent interstellar medium produce regions of enhanced density. Star formation does not occur in every location where the gas is dense, but only in small volumes within clouds where sufficient mass accumulates to become self-gravitating. Collapse and star formation then proceed in that small fraction of the total cloud mass at a very rapid, free-fall rate.

In both models, the rate at which low-density interstellar gas is turned into stars is consistent with the observed star formation rate in the Milky Way Galaxy, about one solar mass per year. The strong magnetic field model achieves this result by setting the time scale for collapse of a dense molecular cloud much longer than the gravitational free-fall time. In the turbulent, intermittent model, only a small fraction of each molecular cloud actually becomes self-gravitating and forms stars. But the physical principles behind the two models are fundamentally different.

As a result, the two models make very different predictions that can be tested observationally. Simulations of molecular cloud formation and evolution carried out with weak magnetic fields show that the fields have a chaotic morphology, because the field lines are twisted by turbulence in the clouds. On the other hand, turbulence cannot twist field lines very much if the field strength is sufficiently strong. Magnetic field lines in dense, strongly magnetized clouds would then be roughly parallel. Collapse along the magnetic field is not impeded by the field, so cores are predicted to have a disk morphology. However, perpendi-

The author is in the Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA. E-mail: crutcher@uiuc.edu

771