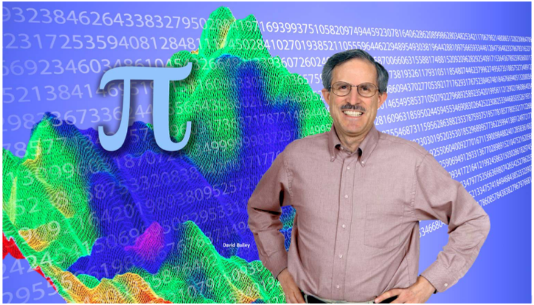
Mathematics, computing and the acceleration of scientific progress

David H. Bailey
Lawrence Berkeley Natlional Laboratory (retired)
University of California, Davis, Department of Computer Science



Did an 18th century mathematician foresee modern computing?

In 1773, Pierre-Simon Laplace wrote:

> *An intelligence knowing all the forces acting in nature at a given instant, as well as the momentary positions of all things in the universe, would be able to comprehend in one single formula the motions of the largest bodies as well as of the lightest atoms in the world, provided that its intellect were sufficiently powerful to subject all data to analysis; to it nothing would be uncertain, the future as well as the past would be present to its eyes.*

We now know that this is a bit unrealistic: from quantum theory, one cannot know both positions and velocities to arbitrarily high accuracy; from chaos theory, small errors are magnified exponentially. Thus predicting the future exactly is not possible.

Nonetheless, much of our universe — physical and mathematical — can now be explored computationally.

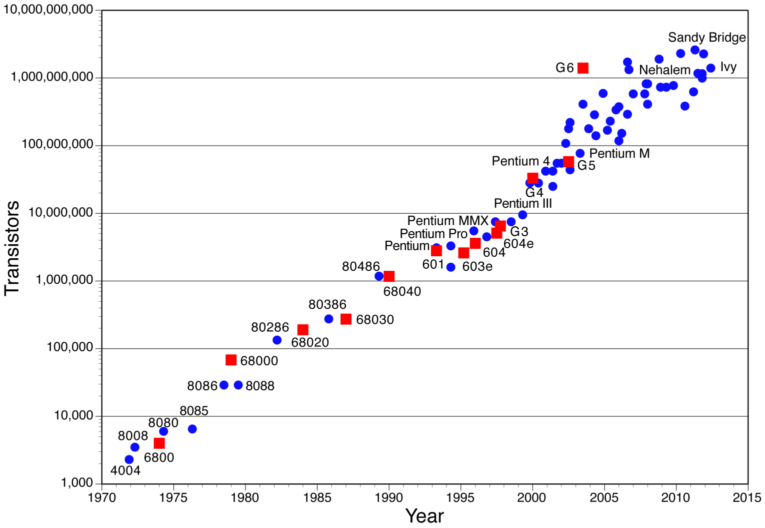
Ada Lovelace: The first computer programmer

Ada Lovelace, daughter of British poet Lord Byron, worked closely with 19th century inventor Charles Babbage in his attempts to construct his pioneering computers. She is often credited as being the first computer programmer.

She realized early on that the potential of these devices extended far beyond scientific computing:

> *[The Analytical Engine] might act upon other things besides number, were objects found whose mutual fundamental relations could be expressed by those of the abstract science of operations, and which should be also susceptible of adaptations to the action of the operating notation and mechanism of the engine.*

Moore's Law: 2.5 million times progress in 40 years, with no end in sight

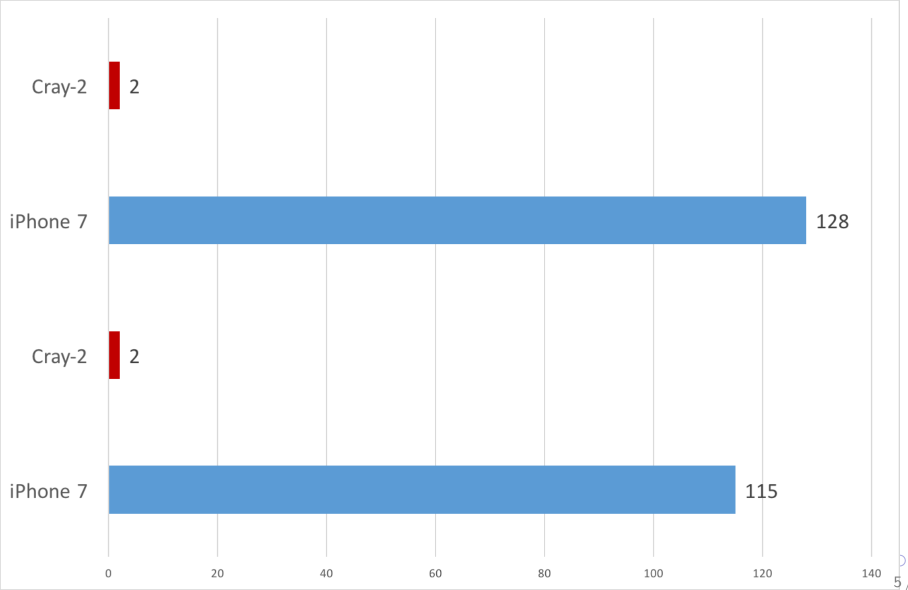


http://pointsandfigures.com/2015/04/18/moores-law/

Cray-2 (most powerful supercomputer in 1986) versus iPhone 7 (2016)

Memory (Gbytes)

Performance (Gflop/s)



Present-day petascale supercomputers

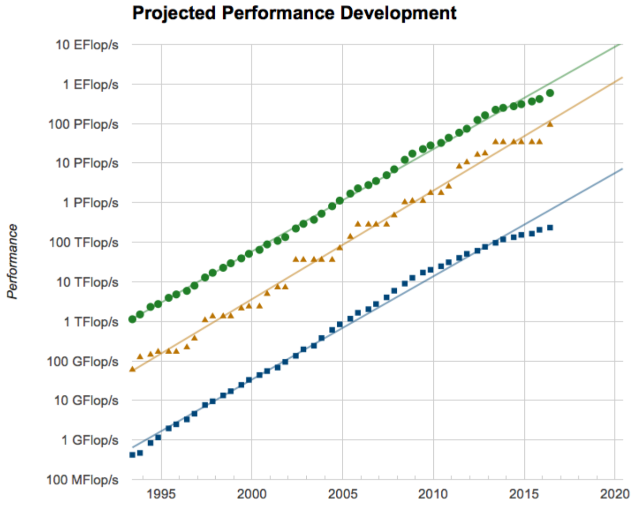
Memory capacity of present-day supercomputers is measured in Pbytes or "petabytes" (quadrillions of bytes).

Performance is measured in Pflop/s or "petaflops" (quadrillions of 64-bit floating-point arithmetic operations per second).

How big is a quadrillion?

- The volume of Lake Ontario is 1.64×10^{15} liters, or roughly 1.6 quadrillion liters.
- The distance to Alpha Centauri is 4.15×10^{16} meters, or roughly 42 quadrillion meters.
- The time since the big bang is 7.25×10^{15} minutes, or roughly 7 quadrillion minutes.

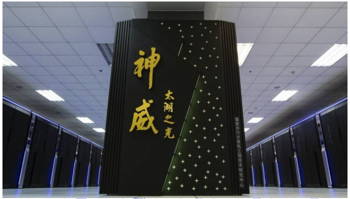
Performance of the world's top 500 supercomputers



Green = Sum #1 thru #500
Orange = #1
Blue = #500

Source: `http://www.top500.org`

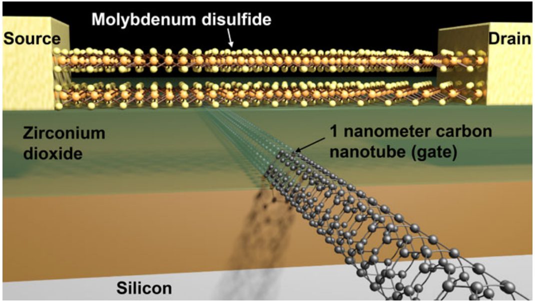
The Sunway TaihuLight system runs at 93 petaflops (93 quadrillion operations per second).



Is Moore's Law ending?

Ever since the dawn of computer technology, many have predicted that Moore's Law will "soon" end. Is today really the end?

October 2016: Berkeley Lab researchers demonstrated a transistor only 1 nanometer in size (1/20th of current state-of-the-art), constructed from a carbon nanotube and a commonly used engine lubricant.



http://newscenter.lbl.gov/2016/10/06/smallest-transistor-1-nm-gate/

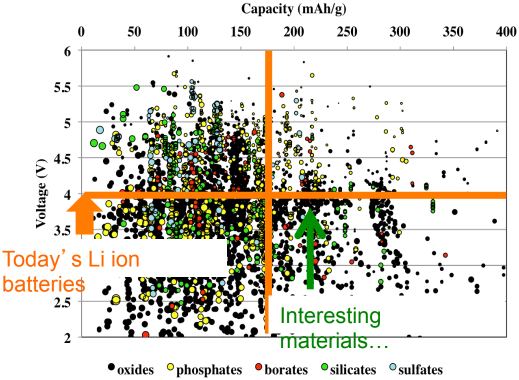
Advanced computation in scientific and mathematical research

- Large-scale computer simulations are being used to understand climate change.
- Computer searches for Type 1A supernovas led to the discovery that the universe is accelerating.
- Analyses of the cosmic microwave background are used to study the big bang.
- Computations were essential to the discovery of the Higgs boson.
- Proteins simulations are used in the development of new drugs.
- Computations are used to develop new materials for solar cells and batteries.
- Computations are now being used to discover new mathematical formulas.

Computing in the Materials Project

Researchers at Harvard and the Berkeley Lab have started the "Materials Project":

- Objective: Reduce lag time between materials science advances and real-world commercialization.
- Methodology: Perform extensive calculations of tens of thousands of potentially interesting compounds, then save in searchable database.
- Now in operation at http://www.materialsproject.org.



Voltage vs. capacity for over 20,000 Li-ion cathode compounds using high-throughput *ab initio* methods.

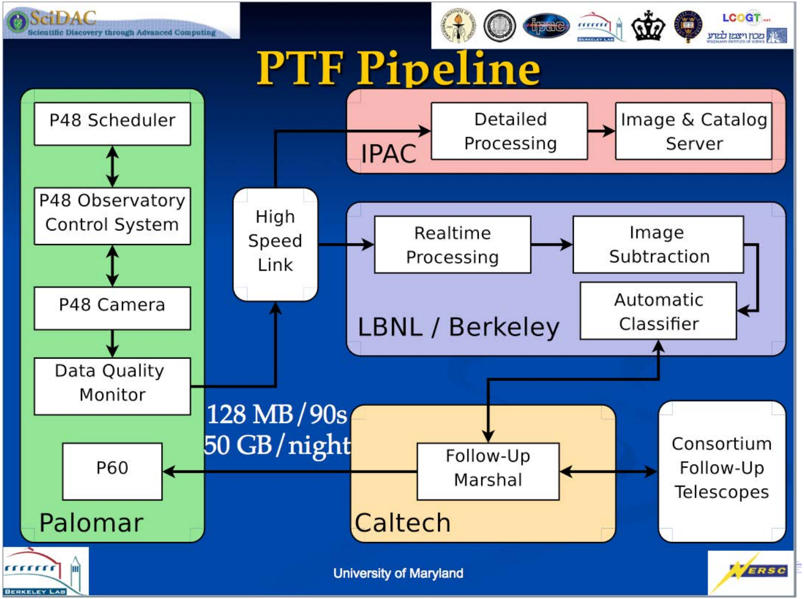
Discovery of a supernova in the Pinwheel Galaxy

In 2011, Peter Nugent of the Berkeley Lab, working within a worldwide consortium of astronomers, discovered a Type Ia supernova in the Pinwheel Galaxy, which is only 21 million light-years from earth. It is the closest and brightest supernova of this type seen in the last 30 years.

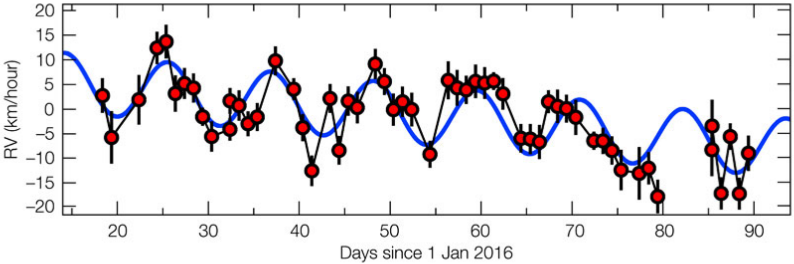
Nugent and his team utilized the Palomar Transient Facility (PTF), a robot-controlled telescope combined with a worldwide network of data processing and storage facilities.

http://newscenter.lbl.gov/2011/08/25/supernova/

The Palomar Transient Facility (PTF) data pipeline



Exoplanet in orbit around Proxima Centauri (4 light years away)



This plot shows the motion of Proxima Centauri, either towards and away from Earth, deduced from very small Doppler shifts in the wavelength of the star's light.

Computer analyses, essential to the discovery, show data are consistent with a planet roughly 1.3 times the mass of Earth, orbiting at a distance of roughly 7 million kilometers from star.

http://www.eso.org/public/usa/news/eso1629/

The Starshot Initiative

April 2016: Russian billionaire Yuri Milner and several Silicon Valley leaders propose to send a fleet of “nanocraft” to explore the Alpha Centauri system and its planets:

- Thousands of credit-card-sized spacecraft, quickly accelerated to 20 percent of the speed of light by light sails powered by an array of Earth-bound laser beams.
- Upon arrival in the Alpha Centauri system in 20 years or so, the nanocraft will send back photos and other data, which will arrive at Earth four years later.

There are daunting technical hurdles, including:

- Fabricating megapixel cameras, computer processors and batteries for the nanocraft, all weighing less than one gram and able to survive 20 years in space.
- Maintaining integrity of the light sail and nanocraft during acceleration by lasers.
- Producing sufficient laser power and maintaining the focus of the laser array.
- Detecting the very weak images and data sent back to Earth.

http://breakthroughinitiatives.org/Initiative/3

Humans to Mars and beyond

September 2016: SpaceX, headed by Elon Musk, announced commercial plans:

- Spacecraft will ferry 100 or more people to Mars.
- The trip will initially take 80 days, but eventually just 30 days.
- Cost: $200,000 or less per person.
- Ultimate goal: A self-sustaining city on Mars in 40 to 100 years.
- Not to be outdone, Boeing is planning a competitive system.

Key technical challenges:

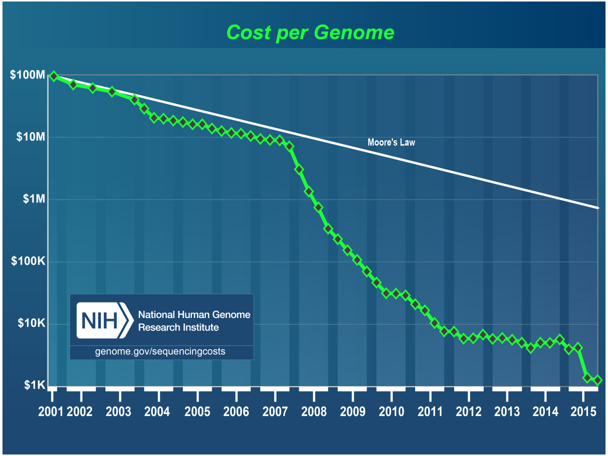
- Advanced artificial intelligence — key decisions need to be made instantly, on-site.
- High-tech robots that can function on the surface of Mars.
- High-tech, robotic medicine and remote drug synthesis.
- Computer-controlled, 3-D manufacturing of everything from common tools to computer equipment.

http://www.nytimes.com/2016/09/28/science/elon-musk-spacex-mars-exploration.html

https://www.bloomberg.com/news/articles/2016-10-04/boeing-ceo-vows-to-beat-musk-to-mars-as-new-space-race-beckons

DNA sequencing technology: Even faster than Moore's Law

The original human genome project took 15 years and cost hundreds of millions of dollars; now it can be done quickly for a mere $1000.



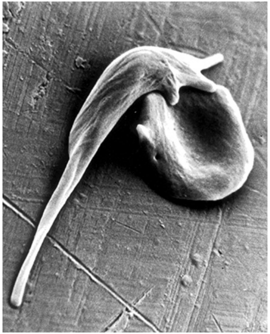
https://www.genome.gov/sequencingcostsdata/

DNA editing technology paves the way for cure to sickle cell disease

October 2016: Researchers have combined cheap DNA sequencing technology with the new "CRISPR" DNA editing to fix the mutated gene responsible for sickle cell disease.

- The team used CRISPR-Cas9 to correct the disease-causing mutation in hematopoietic stem cells, which are precursor cells that mature into red blood cells.
- The correct cells producted healthy hemoglobin.
- After additional tests and analyses, clinical trials could start within five years.

This is an important advance because for the first time we show a level of correction in stem cells that should be sufficient for a clinical benefit in persons with sickle cell anemia. [Mark Walters, UCSF]

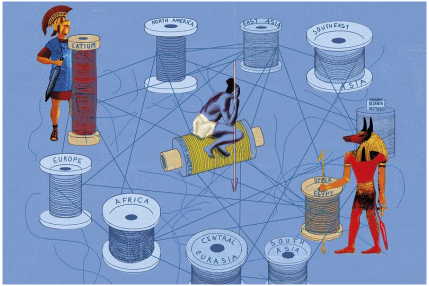


http://news.berkeley.edu/2016/10/12/genome-engineering-paves-way-for-sickle-cell-cure/

Computing in historical and archaeological research

October 2016: A team of researchers at Oxford, University of Connecticut, Exeter and others are assembling a vast database of historical and archaeological knowledge that can be explored to test historical theories.

- The world is divided into ten regions, with three sample areas within each, according to geography and level of social development.
- Researchers have entered 130,000 facts, spanning from the dawn of agriculture to 1900.
- Goal: make history more evidence-based, and detect historical patterns not otherwise visible.
- Sample research topic: What causes societies to collapse?

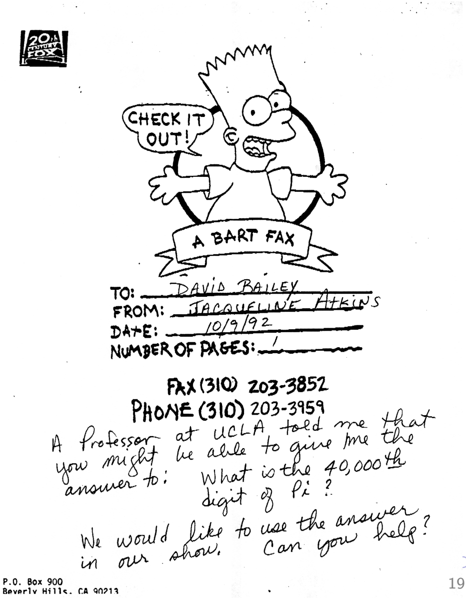


https://www.newscientist.com/article/mg23230950-600-the-database-that-is-rewriting-history-to-predict-the-future/

A fax from the Simpsons

In October 1992, I received this fax from the Simpsons TV show, requesting the 40,000th digit of π. I computed the first 40,000 digits, and faxed the result back (noting that the 40,000th digit is a 1).

In the episode airing May 6, 1993, Apu, the manager of a convenience store, was challenged by Marge's attorney in a courtroom. He replied that he has a perfect memory. For example, he said, he can recite 40,000 digits of π, and the last digit is a 1.



Computing in mathematical research

Computation has long been an essential tool for applied mathematics, but now is being applied to problems in pure mathematics as well.

Example: In 1996 a computer program discovered this formula for π:

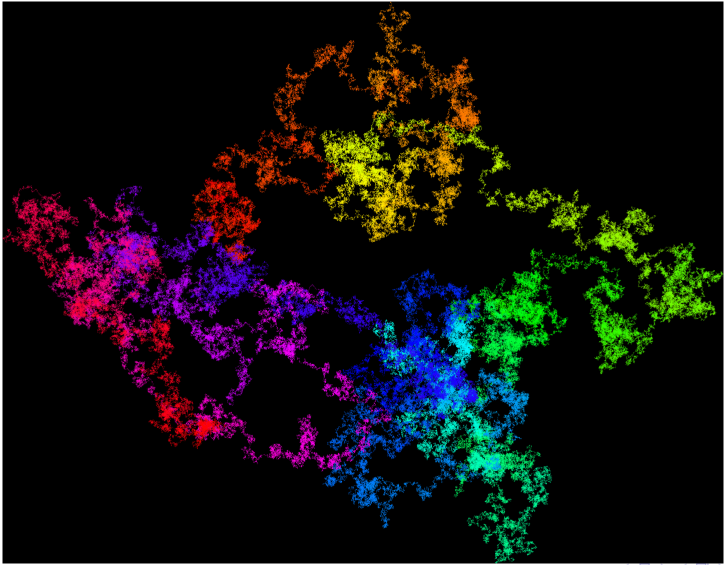
$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n} \left(\frac{4}{8n+1} - \frac{2}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right)$$

This formula has the remarkable property that it permits one to calculate binary or base-16 digits of π beginning at an arbitrary starting point, without needing to calculate any of the preceding digits. Prior to this discovery, mathematicians had presumed that this was not possible.

This formula was discovered using high-precision software and the "PSLQ" algorithm of mathematician-sculptor Helaman Ferguson, a graduate of Hamilton College.

Helaman Ferguson's "Syzygy" sculpture at Hamilton College

A random walk on the first 100 billion base-4 digits of π



http://gigapan.com/gigapans/106803

The excitement of modern high-tech science

Modern society today faces many serious challenges: poverty, hunger, inequality, climate change, terrorism, crime, racism and many more.

Science and technology is one big bright spot:

- Computer technology continues to advance at a dizzying pace.
- We are on the verge of human travel to Mars and beyond.
- Cures of many previously untreatable diseases and conditions are now within reach.
- Advances in astronomy may one day soon answer the question of whether we are alone in the Milky Way.
- We may be on the verge of finding the "final theory" — an elegant set of laws that describes how the universe operates from the big bang to the present day.

We can all share in these exciting developments, which give meaning to human existence and hope for a better future.

The war on science

In his new book *The War on Science*, Shawn Otto describes a three-way “war” on modern science in present-day America:

- *Some sectors* of the academic left, which have asserted that science has no claim to objective truth.
- *Some sectors* of the religious right, which have fought evolution and more under the banner of biblical literalism.
- *Some sectors* of the industrial world, which have fought scientific findings in the area of health and environmental protection.

What is Otto’s proposed solution?

Otto calls for mathematicians and scientists to abandon their decades-old habit of keeping their heads buried in their offices, and instead take to the streets — dramatically improve their outreach to the public.

The growing gulf between scientists and the public

- 40% of Americans do not believe that the Earth (or even the entire universe) is more than a few thousand years old.
- 40% do not believe in evolution, saying instead that humans have existed in their present form from the beginning of time.
- 30% do not accept that the climate is changing.
- 52% do not accept that human activity is a key factor in climate change.
- 17% say that vaccines are unsafe or cause autism.
- Political candidates often reject conclusions of widely accepted scientific consensus.

http://www.pewresearch.org/fact-tank/2016/02/12/darwin-day/

http://www.pewinternet.org/2016/10/04/public-views-on-climate-change-and-climate-scientists/

Why doesn't the public trust science?

Numerous reasons have been proposed, including:

- The public fears that advanced technology will steal their jobs.
- The credibility of science has been questioned in the wake of several widely publicized lapses in reproducibility.
- Scientists are perceived as out of touch with common people.
- Scientists have failed miserably to communicate the wonder and excitement of modern science to the public.

http://www.realclearscience.com/articles/2016/08/20/why_doesnt_the_public_trust_science_109728.html

Reproducibility crises in biomedicine, psychology, economics, finance

- In 2011, Bayer researchers reported that they were able to reproduce only 17 of 67 pharma studies.
- In 2012, Amgen researchers reported that they were able to reproduce only 6 of 53 cancer studies.
- In 2015, Virginia researchers reported that they were able to reproduce only 39 of 100 psychology studies.
- In 2015, the U.S. Federal Reserve was able to reproduce only 29 of 67 economics studies.
- In 2015, "backtest overfitting" emerged as a major problem in computational finance.

Virginia Reproducibility Project

http://www.nytimes.com/2015/08/28/science/many-social-science-findings-not-as-strong-as-claimed-study-says.html

Reproducibility in mathematical and scientific computing

A December 2012 workshop on reproducibility in mathematical and scientific computing, held at Brown University, noted that

> *Science is built upon the foundations of theory and experiment validated and improved through open, transparent communication. With the increasingly central role of computation in scientific discovery this means communicating all details of the computations needed for others to replicate the experiment.*

Issues identified in the ICERM report and other studies include:

- The need to carefully document the full context of the computational experiment: underlying mathematics, algorithms, system environment, input data, source code, computed results, timings, etc.
- The need to save the code, input data and output data in a permanent repository.
- The need to preserve numerical reproducibility by controlling round-off error.
- The need for reviewers, research institutions and funding agencies to recognize the importance of reproducibility in computing.

http://www.davidhbailey.com/dhbpapers/icerm-report.pdf

Winning the battle, but losing the war

Mathematicians and scientists have been very successful in their battles to make discoveries, analyze data, write journal articles and obtain grants.

But we are losing the war for the hearts and minds of the public.

What can we do?

- Start a blog.
- Visit schools.
- Give public lectures.
- Write articles for science news and information forums.
- Write books describing scientific developments for the general public.
- Study creative writing, arts and humanities to sharpen communication skills.
- Recognize those who do reach out in hiring, promotion, tenure and research funding decisions.
- Promote interdisciplinary coursework and studies at universities that combine the arts with science, working in synergy rather than in competition or opposition.

Science versus humanities versus the public

Given the growing tensions in society, and the enormous impact of future rapidly changing technology, we can no longer afford a three-way war between science, humanities and the general public.

- Scientists and engineers should study arts and humanities, to better communicate to the public, and to better understand what it all means to society and humanity.
- Artists, musicians and writers should learn math, computing and science, not just to gain skills, but also to better participate in dialogue on emerging issues.

It's in Apple's DNA that technology alone is not enough — that it's technology married with liberal arts, married with the humanities, that yields us the result that makes our hearts sing. [Steve Jobs]

https://www.scientificamerican.com/article/stem-education-is-vital-but-not-at-the-expense-of-the-humanities/

They should have sent a poet

In one memorable scene from the movie *Contact*, Jodi Foster views a galaxy from her spacecraft, and is so overcome with awe that she exclaims,

> They should have sent a poet. So beautiful.
> So beautiful... I had no idea.

In a similar way, those of us involved in scientific research are often stunned by the beauty and elegance of science and the mathematics behind it, along with the remarkable (and quite mysterious) fact that we humans are able to comprehend these laws through diligent effort.

So why don't we do more to share this wonder? Why don't we write some poetry?

Thanks! This talk is available at
`http://www.davidhbailey.com/dhbtalks/dhb-hamilton-plant.pdf`