

The mystery of the slowing space probes

Find out how big data preservation helped to solve the Pioneer anomaly

FEATURE | JANUARY 16, 2013 | BY STEFAN JANUSZ

It could have been the slowing of their on-board clocks. They could have been feeling the effects of dark energy. Or, perhaps, they provided just the evidence needed to support a theory of modified Newtonian dynamics – proposed to explain why spiral galaxies don't lose their shape as they spin.

In the end, though, the explanation as to why [NASA probes Pioneer 10 and 11](#) were slowing down more quickly than expected as they traveled through space turned out to be much more simple. [Thermal radiation](#) was emanating from the decaying radioisotopes which serve as the probes' power sources and this was producing a small amount of thermal recoil. Thermal recoil is a miniscule force which results from the emission of thermal photons from a surface. If the emission of these photons is unevenly distributed across the surface of a spacecraft, they could cause an imbalance in the forces acting on different parts of the spacecraft. This is exactly what happened in the case of Pioneer 10 and 11, due to each probe's radioactive power source being held on the end of a long boom, so as to prevent it interfering with sensor equipment. Consequently, even though the thermal recoil effect was roughly equivalent to the slowing of a car by the photons from its headlights, the imbalance in photon emissions turned out to be subtly diminishing the crafts' velocities as they ventured ever further into deep space. In fact, the discrepancy between the predicted and actual velocities of the Pioneer space probes was so minute that that it caused a [Doppler shift](#) of only 1.5Hz in 1.4GHz radio signals from the craft several billion kilometers out from Earth (indicating a slowing of roughly just one billionth of the crafts' original velocities). As such, despite the appeal of the array of exotic explanations on offer for the slowing effect observed, the cause was discovered thanks to the kind of conscientiousness and rigor that marks out a great scientist – or at least gives them a significant head start.



Artist's conception of the Pioneer 10/ 11 probe. Image courtesy NASA Ames Research Center.

Hiding under the stairs

[Larry Kellogg](#), a former Pioneer team member, had a hunch that the answer lay in careful reanalysis of the data. He had been preserving it for years, scrupulously transferring it from magnetic tape and magneto-optical disks that had been abandoned under a staircase at the NASA Jet Propulsion Laboratory onto a modern hard disk, where it could be more easily accessed. 40Gb of Doppler data, and some meticulous computer modeling of the craft (there were no [CAD models](#) for Pioneer 10 when it was launched 40 years ago), eventually identified thermal recoil as the most likely candidate for the slowdown. It's thanks to Kellogg that we have the data, but it was data that was the hero here, helping to solve a mystery that had lasted decades.

40Gb of basically the same measurement might sound like a lot, but experiments in the physical and life sciences can routinely produce hundreds of terabytes to several petabytes of data. The question is: should scientists really be going to the trouble of storing all this data? The story of the solution to the Pioneer anomaly suggests that they should. Other stories of old data throwing up new discoveries are now coming to the fore too, such as [the discovery of a rocky 'Super Earth' orbiting within the habitable zone around its parent star, HD40307g](#), made by analyzing old data with new methods. Scientists are rapidly developing better ways to analyze large data sets, the storage and processing of which is only really possible thanks to modern computers. If we do keep all of the data, how do we manage and categorize it? How do we avoid the risky business of having to transfer it from old storage formats, if it's thousands or millions of times larger than the data Kellogg was trying to manage? Most crucially of all, will scientists in the future be able to access and understand it? All of these questions have led to the concept of 'big data', which neatly summarizes both the challenges and the means of tackling the vast data sets being produced by modern-day science.

Rotting data

"Without context, data rots!" said the [Australian National Data Service's Ross Wilkinson](#) at [the EUDAT conference in Barcelona last year](#). "It needs to be integrated into other datasets and publicized... data coming off an instrument that is publicly funded is intended to be shared." Tagging and categorizing data with metadata – 'data about data' – helps scientists sort and locate it in the future, just as you can sort through your computer's music library to look specifically for tracks tagged with the genre 'rock' or 'classical', for example. It's also reminiscent of the [Dewey Decimal System](#) loved by librarians, and indeed it is from libraries that some of the ideas of how to deal with big data are coming, especially when it comes to uniquely identifying data.

[Peter Wittenburg from the Max Planck Institute for Psycholinguistics](#) is EUDAT's scientific coordinator. One of the projects he works on is [CLARIN](#), a linguistics project that – through its [Virtual Language Observatory](#) – aims to preserve corporuses of existing and endangered languages, while simultaneously building a foundation for something even more far-reaching: e-humanities. This would allow social sciences researchers to search for specific information using natural language terms – a crucial part of what is termed the semantic web. The challenges facing big data research in the humanities are not all that different from those in the physical sciences, maintains Wittenburg: "We're realizing now that there are common components – building blocks – for global data infrastructures. A crucial one we're working on now is a worldwide registration and resolution system for persistent identifiers. Every piece of data would have a PID, just as computers connected to the internet have a unique IP address."

Stopping the rot with standards

Indeed, standards (both in terms of file formats and for metadata) form a basic building block of dealing with big data. For EUDAT, the European Data Infrastructure, an agreement has been made for all metadata for European projects to be in English. But working towards standards is important even for very similar research communities, such as those working on mouse models of disease, and those researching the same diseases in humans; the two communities can often use different terminology for the same thing. [Stephanie Suhr of BioMedBridges](#) explains: "We are linking diabetes and obesity-related data from human patients and from mouse models, which requires translation between the terms used by both research communities. Systematic use of extensive mouse data resources by clinical researchers will be an extremely powerful tool for new scientific discoveries and therapies."

Some researchers believe that big data could mean an even more fundamental shakeup of science. [Matthew Dovey, from the UK body JISC](#), says: "In some branches of applied science, being able to make accurate predictions can be of more practical importance than understanding the underlying models. For example, determining future weather patterns, or choosing between different but established medical treatments based on a patient's lifestyle. Here, big data can be used to identify trends and patterns with improved reliability. Ever increasing sophistication of analytical tools may even one day replace the role of the theoretical scientist in hypothesizing new models. Scientists then have the task of devising experiments to challenge and test these computer-generated models."

Average:

Your rating: None Average: 4.4 (17 votes)

About the Author »

Stefan Janusz

e-Science Impact Reporter, e-ScienceTalk

RELATED TERMS: [big data](#) [Data management](#) [NASA](#) [Pioneer](#) [probe](#) [spACE](#) [North America](#) [data management systems](#) [information science](#) [physics and astronomy](#)