A "Gap Bridger"

Richard Rashid
Microsoft
One Microsoft Way
Redmond, WA 98052
(01) 425-706-9080
rashid@microsoft.com

Abstract

I knew Jim Gray as a colleague, a friend and an employee. He created the first Microsoft Research group outside of our initial research lab in Redmond, WA. Jim's impact on industry and science was not limited to his own area of research. He championed new ways of thinking about how computer science and information processing could be integrated into other areas of research. In so doing; he will have a lasting impact on the rate of progress in many disciplines. His success stemmed not just from what he knew and what he could do with his knowledge, but also from who he was as a person. Jim was a "Gap Bridger" – someone who could connect people, groups, companies and disciplines.

1. A Special Kind of Person

"Jim Gray is the kind of person that you would want your child to grow up to be."

Over the years I frequently made this comment in my speeches about Microsoft Research and the leadership of our research group in the Bay Area. Many of us, I would guess, would like to see our kids grow up to accomplish something and to make an impact in the world. But we also want them to exemplify human values like kindness and humility. Jim was incredibly smart and accomplished but also thoughtful, unassuming in person and generous toward others. I think it is indicative of who he was that most of the conversations I remember having with Jim over the years were about the accomplishments of other people or the problems that other people had that needed to be solved. Jim's willingness to give of himself, educate, mentor, support others, solve the hard problems people bring to him and reach out to make the world better made him one of the field's most loved individuals.

That love was evident in the extraordinary worldwide effort to find Jim when he went missing last year. So many people cared so much and tried so hard to find him. The search transcended anything that had come before using new tools and new technologies and diverse communities of people -- many of whom never even knew Jim. Boats, submarines, planes and satellites were all deployed. Computers around the world were brought to bear as well as armies of people looking at images. That the search ultimately ended without discovery is a pain still felt today…

Jim Gray Tribute, May 31, 2008, Berkeley, CA, USA.

2. Smart, Compassionate and Independent

If you go online you can learn a bit about Jim's early life. He grew up in a working-class neighborhood on San Francisco's south side. By all accounts I've read he was as a child growing up not so different than he was as a grown man – smart, compassionate and independent. Reporter Brier Dudley of the Seattle Times interviewed a friend of Jim's family that told stories of Jim as a child going out by himself to see bears at Yosemite and swimming out past the breakwater in Santa Cruz.

His professional career is much better documented.

Jim earned his PhD at Berkeley in 1969 – the first student to get a PhD at that institution. He did a short stint working at Bell Labs but then moved to IBM Research in San Jose. Jim went on to be a pioneer in the field of databases and transaction processing leading some to call him the "father" of the field. Everywhere he went his impact was felt: IBM, Tandem, DEC and Microsoft. Jim's impact is measured not just in his technical accomplishments but also in the number of people around the world whose work he inspired.

3. Creating a New Lab

I honestly can't remember when Jim and I first met – undoubtedly at a conference somewhere. I first got to know him well, though, when Jim joined David Gifford and me in teaching a one-week summer systems course at Stanford University. I was very impressed by his command of the subject, his approach to teaching and his empathy for his students. That experience made me jump at the chance to hire Jim when he first let me know he was "available and interested" even though it meant setting up a new lab in San Francisco.

I fondly remember the day I spent with Jim back in 1995 "sealing the deal" for the new lab. We walked around Market Street in San Francisco, had lunch at a table outside and talked about our dreams for the field. Neither of us could have guessed how much would be accomplished over the next 10 years.

4. Bridging Gaps

One of Jim's enduring qualities was that he was able to bridge the gap between people, between companies and between disciplines. It didn't matter if he was working at Microsoft and someone else was working at Oracle. He could work as well with competitors as friends.

Jim had a way of listening to others problems and finding way of bringing people together. Over the last 10 years, that often meant bringing together computer scientists with scientists from different disciplines to accelerate research. The first step along this path was the creation of the TerraServer in 1997/1998 [Barclay 1998].

I well remember the conversation I had with Jim that led to the creation of the TerraServer. Jim had stopped by my office to talk about the idea of a putting together a terabyte scale database system that would be accessible over the internet. We talked about what kind of data we could get that would both be useful and at sufficient scale and hatched the notion of looking at satellite imagery of the earth. Jim and Tom Barclay ran with that idea and the TerraServer was born – in its day the largest online database in the world and the granddaddy of online aerial and satellite imagery.

In the beginning, the TerraServer was a tool for people – a way for individuals to see their world and often their own backyards as they had never seen them before. It was also a learning experience for Microsoft's product teams and for our image information partners such as the US Geological Service. For the first time vast collections of data on "dusty tapes" could be integrated into a single database and accessible online 24/7 [Barclay 2000].

Over time, the presence of so much data online opened up new possibilities for use. Early "mash-ups" were created as Jim and his team made the TerraServer accessible as a web service and opened it up to applications from groups as diverse as the USDA that used it for soil reports and school children doing class projects.

Even as the uses of TerraServer expanded, Jim began working with other communities of scientists to bring to other disciplines the power of global scale databases, common schema and powerful analysis tools. Jim believed that data management was critical to long term success in many fields [Gray 2005].

The TerraServer work led directly to collaborations with the astronomy community and the creation of the SkyServer. The SkyServer (www.skyserver.sdss.org) brought together data from top telescopes around the world into a federated infrastructure [Szalay 2000]. It was a tool that could be used both by ordinary people who simply wanted to see the great images of the sky taken by the best telescopes in the world and by scientists who could use it to find information or perform data mining [Singh 2006]. A school child could do an online lesson and learn about Hubble's Constant and at the same time a scientist could write 3 lines of SQL and find all the quasars in the Sloan Digital Sky Survey. SkyServer was a Virtual Observatory that had a view of the sky 24 hours a day. The SkyQuery Portal (www.skyquery.net) added the capability of doing queries across federated astronomical data sources from around the world.

Of his work on SkyServer, Jim was quoted as saying "Astronomers have this great big shoebox of data that they can't find things in very easily. We're trying to make it easier for them."

Jim went on to work with other scientific groups and communities in geography, hydrology, oceanography, environmental monitoring, biology, and health care. This included work with the National Center for Bioinformatics and the National Library of Medicine in creating a portable and internationally federated version of PubMed Central (www.pubmedcentral.nih.gov).

5. Lasting Memories

My last personal memory of Jim dates back to the 2006 21st Century Computing Symposium in China where we lectured together. I have a great mental image of Jim Gray sticking his head through the Great Wall, ringing a giant bell "10 times for luck" and riding a toboggan down a mountain from – of all things –a Buddhist temple.

In many of my speeches I was fond of saying that the only reason I started a research group in San Francisco was because of Jim. "If Jim wanted to live in Monaco", I would say, "we would have a group in Monaco."

And we would have.

6. References

[Gray] http://research.microsoft.com/~Gray/

[Barclay 1998] Tom Barclay, Robert Eberl, Jim Gray, John Norlinger, Guru Raghavendran, Don Slutz, Greg Smith, Phil Smoot, John Hoffman, Natt Robb III, Hedy Rossmeissl, Beth Duff, George Lee, Theresa Mathesmier, Randall Sunne, Lee Ann Stivers, Ken Goodman, *The Microsoft TerraServer*, MSR-TR-98-17.

[Barclay 2000] Tom Barclay, Donald R. Slutz, Jim Gray, *TerraServer: A Spatial Data Warehouse*, SIGMOD Conf. 2000: pp. 307-318.

[Szalay 2000] Alexander S. Szalay, Peter Z. Kunszt, Ani Thakar, Jim Gray, Donald R. Slutz, *Designing and Mining Multi-Terabyte Astronomy Archives: The Sloan Digital Sky Survey*, Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data: pp. 451-462.

[Gray 2005] Jim Gray, David T. Liu, Maria A. Nieto-Santisteban, Alexander S. Szalay, Gerd Heber, David DeWitt, *Scientific Data Management in the Coming Decade*, MSR-TR-2005-10.

[Singh 2006] Vik Singh, Jim Gray, Ani R. Thakar, Alexander S. Szalay, Jordon Raddick, Bill Boroski, Svetlana Lebedeva, Brian Yanny, *SkyServer Traffic Report – The First Five Years,* MSR-TR-2006-190.