



HÉCTOR GARCÍA-MOLINA

1953–2019

Elected in 2003

“For contributions to distributed-information systems.”

BY JEFFREY D. ULLMAN

HÉCTOR GARCÍA-MOLINA, the Leonard Bosack and Sandy K. Lerner Professor of Engineering in the Stanford University Department of Computer Science, died of cancer November 25, 2019, a day short of his 66th birthday. He was a leading figure in research into modern database systems and the involvement of large-scale data in science and technology.

Héctor was born November 26, 1953, in Monterrey, Mexico, to Fernando García Roel and Laura Molina de García Roel. He did his undergraduate work in electrical engineering at the Monterrey Institute of Technology and Higher Education, where his father was president, and in 1975 began graduate studies at Stanford, earning master’s degrees in electrical engineering (1975) and computer science (1977) before receiving his PhD in computer science in 1979. His advisor, Gio Wiederhold, introduced him to distributed database systems, which formed the basis of his early research work. As Wiederhold traces his academic ancestry back to Galileo, thus so does Héctor.

In 1979 Héctor joined the electrical engineering faculty at Princeton University, where he served until 1992, when he returned to the Stanford Department of Computer Science. He was director of Stanford’s Computer Systems Laboratory (1994–97), chair of the Computer Science Department

(2001–04), and at the national level, under President Bill Clinton, served on the President’s Information Technology Advisory Committee (1997–2001). He also was appointed a member of the National Academies’ Computer Science and Telecommunications Board (2002–05).

Héctor’s initial work at Princeton involved distributed databases, including the study of long-duration transactions (“sagas”) conducted in a distributed environment. He was involved in the construction and use of what (in the late 1980s) was called a *massive memory machine*, a device with a quarter gigabyte of main memory—a tiny cell phone by today’s standards, but mammoth at the time. This work opened up the study of main-memory database systems, among other areas. Also connected to this work was an early approach to what is today called RAID (Redundant Array of Independent Disks)—the use of redundancy among several disk units to protect against failures. In 1986 Héctor coauthored a paper presented at the 2nd Second International Conference on Data Engineering that was recognized as significantly influencing the development of RAID storage.¹

After moving to Stanford, Héctor began work on information integration—enabling several related databases to be combined in a useful and consistent way—a challenge that to a large extent persists. He introduced the idea of semi-structured data, which is now reflected in XML, JSON, and other approaches to representing information. In 1994 he began working on digital libraries, a major project from which Google emerged (one of his students was Google cofounder Sergey Brin, who left before completing his doctorate). His interest in data preservation led to the exploration of a number of other directions as well, such as peer-to-peer storage and archiving.

Well before the advent of Facebook, Héctor encouraged one of his doctoral students to develop a social network; it was called Club Nexus and operated at Stanford in 2001. This

¹ Salem K, García-Molina H. 1986. Disk striping. *Proceedings, 2nd Second International Conference on Data Engineering*, pp. 336–42.

work evolved into Google's Orkut social network, which at one point had 300 million users (mostly outside the United States). Another area in which Héctor and his students did pioneering work was the use of crowdsourcing to obtain data. He also led a team of undergraduates to develop a system called CourseRank, an information resource for students planning their academic programs; it was later institutionalized at Stanford and replicated elsewhere.

More recently, recognizing that many traditional issues in database systems research were reaching resolution, Héctor helped turn his department toward more modern approaches to dealing with the largest volumes of data possible, recruiting a number of young faculty to pursue these new directions, including the use of big data in machine learning. He was the founding director of the Stanford Data Science Institute, a cross-cutting effort connecting computer science, medicine, and other application areas as well as inviting industrial participation in Stanford research.

In recognition of his pioneering contributions, Héctor was elected a member of the NAE and a fellow of the Association for Computing Machinery (ACM) and the American Academy of Arts and Sciences. In 1999 he received the ACM SIGMOD (Special Interest Group on Management of Data) Innovations Award, which is presented "for innovative and highly significant contributions of enduring value to the development, understanding, or use of database systems and databases." In 2007 his coauthored article on disk striping was recognized with an International Conference on Data Engineering Influential Paper Award. In 2009 CourseRank was selected for the SIGMOD Best Demo Award, and in 2010 his coauthored paper (with Junghoo Cho) presented at the VLDB conference in 2000, "The Evolution of the Web and Implications for an Incremental Crawler," was selected for the VLDB 10-year Best Paper Award.

Héctor graduated 57 PhD students, many of whom hold important positions in technology and education today. He was greatly loved by his students and colleagues, who recognized him as a truly "nice guy"; he was modest, and he was supportive of all around him.

He encouraged his students to do pioneering work, no matter their particular direction. His secret, he said, was to look for the “90% solution,” his term for an approach to a problem that was simple enough to be realistically implementable and yet served well in the cases that arose in practice. A result of this approach was that he and his students pioneered research in many of the areas of database systems and data management that are important today. The impact of these initial studies is clear: at the end of 2019 Google Scholar showed 37 papers of which Héctor was an author or coauthor with more than 500 citations.

A less technical aspect of Héctor’s interactions with students is illustrated by his “database lunches.” When he arrived at Stanford, he started paying for a lunch for all the database graduate students—his and others. He invited them to share their recent personal experiences, called “trip reports,” and give periodic reports on their recent research directions. Students got to know each other both socially and academically and could be relied on to support each other in important ways. For example, they used the lunches to rehearse conference presentations, and because of the friendships that had developed it became easy for presenters to accept constructive criticism from their peers.

In addition to his substantial technical work, Héctor had a serious hobby: photography. By invitation from the athletic department, he roamed the sidelines at Stanford sporting events, taking photos of the action, and many of the images in the *Stanford Daily* newspaper were his. He loved to share his interest with students, and for many years taught a very popular freshman seminar on digital photography.

Héctor is survived by his ex-wife Alicia Garza Martínez and their son David García, as well as the hundreds of students, friends, and colleagues whom he loved and who loved him.

