



"We have a human problem today. How do we ensure that the best students with the best intellectual minds in the various institutions all over the country see research as a worthwhile career to pursue? How do we encourage them to continue on this path and not just stop after a bachelor's or master's degree and continue on to a Ph.D. program and ideally on to an academic career?"

Jayant Haritsa

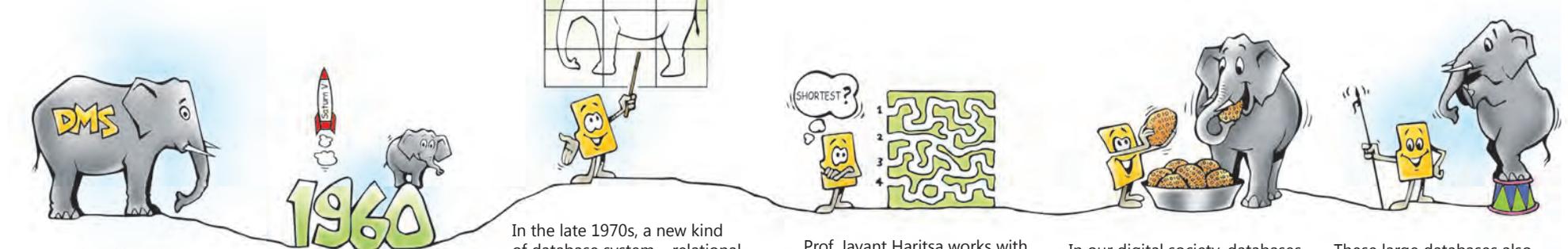
Professor, Supercomputer Education & Research Centre and Chair of Department of Computer Science & Automation, Indian Institute of Science (IISc), Bengaluru

- B.Tech. in Electronics from Indian Institute of Technology, Madras
- M.S. in Computer Science from University of Wisconsin, Madison
- Ph.D. in Computer Science from University of Wisconsin, Madison

Prof. Jayant Haritsa is a pioneer in the design and optimization of database engines that are the core of modern enterprise information systems. His contributions span the full breadth from problem definition to software tools such as the Picasso visualization platform and the CODD metadata processor.



Building robust database systems



Ever wondered how it is that each time you make an online train reservation or bank transaction, the data is processed and you get all the crucial information that you need with just a few clicks of a button? The answer lies in Database Management Systems (DMS). These massive systems organize, store and retrieve vast amounts of data. Database systems are like internal combustion engines in the functions they perform. They 'drive' data just as internal combustion engines help move automobiles.

While societies have always found ways of organizing and storing vast amounts of data, it was only from the 1960s that the concept of Database Management Systems came into vogue. For instance, IBM came up with an Information Management System, initially developed as part of the Apollo program to catalog the materials for the Saturn V moon rocket.

In the late 1970s, a new kind of database system – relational database systems – came to the fore. Here, all information was stored in a set of tables linked through columnar values, and users could easily extract related information across the tables through a powerful query language based on first-order logic. A potent feature of relational systems, which led to their pervasive popularity, was their support for "declarative" user queries. Here, users had to only declare their final objectives and the system took on the responsibility of identifying the fastest execution strategy to achieve these objectives. That is, the system provided the means for the user ends.

Prof. Jayant Haritsa works with relational database systems and works out ways in which they can provide efficient and robust solutions to declarative user queries. Popular databases are used often by people for the information they seek. The trick is to quickly come up with the ideal retrieval process from an extremely large number of alternatives. As an analogy, think of the many different routes one could take in travelling between Bangalore and Delhi, and automatically identifying the ideal choice with regard to metrics such as time, distance or fuel efficiency.

In our digital society, databases store vast amounts of data which need to be retrieved often and as quickly as possible. These include databases of ambitious projects such as the Aadhar unique identity system; biometric databases; genetic databases; health records, and other vital data. Prof. Haritsa's work provides software tools to characterize and significantly improve the behavior of current industrial-strength database engines, and to inexpensively evaluate the design and performance of futuristic Big Data systems.

These large databases also present challenges in the form of security, privacy, distance, safety, connectivity and natural language. Haritsa has been able to come up with solutions for the problem of user privacy that involve mathematical models which randomize customer data in a manner that permits statistical queries to be successfully launched against this data while ensuring the information cannot be traced back to their owners.