Return to the Boston Area After 50+ Years: What can we learn from VLDB 1975?

Philip A. Bernstein
Microsoft Research
phil.bernstein@microsoft.com

Lei Chen
HKUST (GZ) and HKUST
leichen@cse.ust.hk

David DeWitt
University of Wisconsin-Madison
david.dewitt@outlook.com

Sebastian Link
University of Auckland
s.link@auckland.ac.nz

Volker Markl
TU Berlin
volker.markl@tu-berlin.de

Renée J. Miller
University of Waterloo
rjmiller@uwaterloo.ca

C. Mohan
HKBU and Tsinghua University
seemohan@gmail.com

Michael Stonebraker
Massachusetts Institute of Technology
stonebraker@csail.mit.edu

Divesh Srivastava
AT&T
divesh@research.att.com

ABSTRACT

This article is a written record of the "Return to Boston" event organized at VLDB 2026 to celebrate the return of the VLDB conference to the Boston area, which was the site of the first VLDB conference in 1975.

PVLDB Reference Format:
Philip A. Bernstein, Lei Chen, David DeWitt, Sebastian Link, Volker Markl, Renée J. Miller, C. Mohan, Michael Stonebraker, and Divesh Srivastava. Return to the Boston Area After 50+ Years: What can we learn from VLDB 1975?. PVLDB, 19(12): 4983 - 4988, 2026.
doi:10.14778/3827998.3828155

1 INTRODUCTION

The first VLDB conference was held in 1975 in the Boston area. After travelling around the world for 50 years, the VLDB conference returned to Boston in 2026. To mark this momentous occasion, VLDB 2026 organized a plenary "Return to Boston" event to celebrate the impact of VLDB on the global data management community. The event featured three sessions, with a number of prominent speakers from the data management community. This article is a written record of that event.

2 LEADING UP TO VLDB 1975

The first session had presentations by Philip A. Bernstein and Michael Stonebraker, who recounted the state of data management in the early and mid 1970s, leading up to the first VLDB conference in 1975.

2.1 The Tale of Philip A. Bernstein

In 1975, database management was a new field of computer science. Few computer science professors specialized on the topic, and about half of them were in business schools (due to the connection to business data processing). Commercial database systems were a relatively new product category, with data models that offered record-at-a-time access to a network of records connected by pointers. Record-oriented file systems were evolving into database systems that stored data definitions (i.e., schemas) in addition to the data. The relational data model was a research topic. Recognizing this evolution, the ACM Special Interest Group for data management, called SIGFIDET (for File Description and Translation), was renamed SIGMOD (for Management of Data) in 1975. The first SIGMOD conference was held in May 1975 at IBM San Jose. There were about 75 attendees. I was there, presenting my first major conference paper.

At the request of Dr. Craig Fields, a program manager at ARPA, Dr. Stockton Gaines (at RAND Corporation) and Prof. David K. Hsiao (at The Ohio State University) wrote a research note about very large data bases. (The research note is mentioned in a few places, but it seems that no copy has survived.) To attract attention to the problem, they got funding from Dr. Fields for a conference on the topic, namely The International Conference on Very Large Data Bases [19]. Since Prof. Hsiao was visiting Prof. Stuart Madnick at MIT, and since SIGMOD 1975 was held on the west coast, they decided to hold the VLDB conference in the Boston area, which was home to many computer companies at the time [17]. The conference took place September 22-24, 1975, at the Sheraton Tara Hotel in Framingham, Massachusetts. I was there, my first conference post-Ph.D. Dr. Gaines and Prof. Hsiao were conference co-chairs, and Prof. Peter Chen was program chair. There were 29 paper presentations. Since Prof. Hsiao had recently been appointed editor-in-chief of the new journal ACM Transactions on Database Systems, they selected best papers from the conference to be published in the journal's first issue.

The first VLDB Conference program covered topics that are familiar to everyone working on database management today. They included access methods, data models, data integrity, data security, database design, database machines, database system architecture, distributed databases, index selection, query languages, recovery, and transactions.

Some amusing tidbits: At the time, the field hadn't decided whether to use the term "data base", "data-base", or "database."

Proceedings of the VLDB Endowment, Vol. 19, No. 12 ISSN 2150-8097.
doi:10.14778/3827998.3828155

VLDB guessed wrong (Data Base); SIGMOD sidestepped the issue (Management of Data). Conference Registration was $40 for ACM & SIG members, and a single hotel room was $24/night (approximately $148 in 2026 dollars). The conference ended up with a $2000 surplus [17], which the conference organizers decided to use to fund a standing committee to run future conferences. To avoid competing with the SIGMOD Conference, a U.S.-based conference, they decided future conferences should be held overseas, alternating among different regions. Hence, the second VLDB was held in Brussels in 1976 and the third in Tokyo in 1977. After 10 years roaming around the world, it returned to the U.S. in Los Angeles in 1988.

2.2 The Tale of Michael Stonebraker

In my opinion the "high level bit" in the DBMS community in the mid 1970s was the "data model wars", touched off after Codd's pioneering paper in 1970 [11]. Like many others at the time, I closely followed Codd's work and read (multiple times) the CODASYL proposal and was familiar with IBMs IMS. On the one hand, there were a collection of (mostly) academics who firmly lined up behind Codd in the relational camp. On the other hand, there was Charlie Bachmann, inventor of IDS, on which the CODASYL proposal was based, and his followers. In the mid 1970s the two groups largely talked past (yelled at) each other. The exemplar of this effect was the "Great Debate" at SIGFIDET (now SIGMOD) in 1974. I will go through my impressions of the two communities, as well as my own point of view.

I will then continue with what happened in the latter part of the decade: how CODASYL attempted to adapt, why the relational camp won, and culminating in a little known news article from (I believe) 1980 titled "The Eagle Has Landed". This is partly a history lesson (people who do not understand history are doomed to repeat it), but has a few important takeaways which everyone in the DBMS community should take to heart today. The first one is the answer to "Why did Codd invent the relational model?" in his own words. I will touch upon other important takeaways during my presentation.

3 MEMORIES OF A FEW VLDB 1975 PAPERS

The second session had presentations by David DeWitt, Sebastian Link and C. Mohan, who provided their personal insights on the technical impact of a few papers from VLDB 1975 that are relevant even today,

3.1 Memories of David DeWitt

In 1975 it was not settled fact that the relational model would emerge as the defacto standard for our field. At the time of Codd's 1970 paper [11], most commercial products were based on either a hierarchical data model (e.g. IBM's IMS) or the CODASYL standard created by the DBTG Task Group (e.g. Cullinane, Honeywell, DEC, HP, Univac, and ICL). Many (most?) people doubted that the relational model could ever be implemented efficiently and it became a period of intensive debate within the field – especially between Charlie Bachman (for the CODASYL camp) and Mike Stonebraker.

The VLDB 1975 Proceedings had three papers of particular significance to this debate. The first is the paper titled *PRTV: An Efficient Implementation for Large Relational Data Bases* [35]. While its impact was eventually eclipsed by the successes of the INGRES [29] and the System R [2] projects, it is generally recognized as the world's first operational relational database system and included a primitive, rule-based query optimizer.

Two other papers from the VLDB 1975 conference are of particular significance to those who have worked in the area of scalable database systems. These include: *The Design of a Rotating Associative Array Memory for a Relational Database Management Application* [21] and *CASSM: A Cellular System for Very Large Data Bases* [32].

Both systems, as well as the 1975 RAP paper [23], explored the use of parallelism as a solution to the challenges of implementing the relational data model effectively. Like the RAP project, each of these two systems associated a small microprocessor with one head of the fixed head disk or drum. Unlike modern disk drives, the heads of these storage devices did not move. Rather, each track of the drive had its own read/write head. By combining a simple processing unit capable of filtering rows with each head (and hence each track) the entire database could be "searched" in a single revolution. The attractiveness of such a design allowed data to be filtered in place, avoiding the need to transfer unnecessary rows into the memory of the host computer. Joins in such a design employed what is basically a nested loops algorithm requiring one full scan of the inner table for each row in the outer table. These three systems represent the very first attempt to use parallel query execution to obtain scalable performance.

The CASSM project tried to hedge its bets on which data model would eventually dominate the field and attempted to support all three data models equally well. In the case of the relational model, tuples were stored as (attribute name, value) pairs and rows were stored sequentially along a disk track. In attempting to support all three data models the design of CASSM was frankly far too complex.

The RARES project [21], like RAP, focused only on supporting the relational data model. Like RAP and CASSM, its design associated a simple LSI processor with each read/write head of a fixed head disk. Unlike RAP, which laid out rows along each track, RARES spread the attributes of a single row across multiple tracks allowing some filters involving multiple attributes to be processed in a single revolution. An external bit array was used to handle more complex filters. Joins were implemented in the host computer. After the qualifying rows of the outer table were filtered and retrieved into the host's memory, they were sorted and divided into equal-sized buckets. Then, the upper and lower values of each bucket were used as a filter condition to retrieve matching rows of the inner relation where they were joined with the current outer bucket.

After the publication of the RAP, CASSM, and RARES papers in 1975 the field of "database machines" simply exploded. There were dozens of academic projects world-wide involving numerous exotic storage technologies including charge-coupled devices and bubble memories as a replacement for fixed head disks and drums. Most never actually ran a SQL/Quel query and the over reliance on custom hardware proved to be a losing long-term strategy. By 1983 the field had collapsed [7]. However, the goal of using parallelism to query relational database systems continued and eventually succeeded with the adoption of commodity hardware

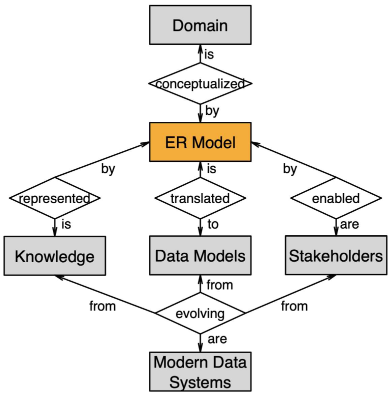

Figure 1: Impact of the ER-model as ER-diagram

based, shared-nothing designs as exemplified by the Teradata [33] and Gamma [13] systems. Today, the use of parallelism for processing petabyte-sized relational data sets is a given. Not so in 1975.

3.2 Memories of Sebastian Link

Peter P. Chen introduced the Entity–Relationship (ER) model in his landmark paper *The Entity–Relationship Model: Toward a Unified View of Data*, first presented at VLDB 1975 [8] and published as the inaugural scientific article in *ACM Transactions on Database Systems* in 1976 [9]. Over the following five decades, the paper has become one of the most influential publications in computer science, accumulating more than 14,000 Google Scholar citations and over 5,400 ACM Digital Library citations by August 2026, alongside tens of thousands of downloads. Its impact extends far beyond a single publication, giving rise to the *International Conference on Conceptual Modeling* (running since 1979) and the *Data & Knowledge Engineering* journal (since 1985). The ER model fundamentally transformed database design by introducing entities and relationships as intuitive yet mathematically rigorous abstractions [34] that bridge the gap between application requirements and database implementation [36]. Because these concepts closely resemble the way people naturally describe real-world domains [10, 18], they enable effective communication between technical and non-technical stakeholders while supporting the systematic translation of conceptual models into diverse logical data models, including relational, hierarchical, network, object-oriented, graph, spatio-temporal, and big data systems [15]. As a result, the ER model has had enduring influence on conceptual modeling, database design, software engineering, modeling standards, education, and ontology development.

Today, the ER model remains central to the evolution of data management and conceptual modeling [30]. As organizations increasingly rely on data-driven decision-making, conceptual models are becoming more tightly integrated with artificial intelligence [5], enabling automatic generation of ER models from natural language [22] and the discovery of conceptual schemas directly from existing datasets [20]. At the same time, collaborative cloud-based modeling platforms [4, 24], stronger links to data governance and regulatory compliance [25], and growing demands for explainability and transparency [31] are expanding the strategic importance of conceptual modeling beyond database design alone. Recent developments, including the standardization of GQL (Graph Query Language) [1] and the emergence of E/R graphs [27, 28] that unify conceptual schemas with property graphs [6], further strengthen the connection between conceptual, logical, and graph data models. These advances move closer to Chen's original vision of a unified view of data, suggesting a future in which ER modeling serves not only as a design methodology but as the foundational abstraction underlying modern database systems, knowledge graphs, and AI-driven information management [3, 12, 34]. This is illustrated by the ER-diagram in Fig. 1.

3.3 Memories of C. Mohan

Work done in multiple pioneering IBM database projects of different sizes were described in 4 papers and an extended abstract in the proceedings of VLDB 1975. The projects covered were - an extended abstract on IBM UK Peterlee Scientific Center's Peterlee Relational Test Vehicle (PRTV) [35], 2 papers on IBM San Jose Research Laboratory's System R [14, 16], and one paper each on IBM Thomas J. Watson Research Center's DIAM (Data Independent Accessing Model) II [26] and QBE (Query-By-Example) [37]. Together those publications covered numerous significant database topics: concurrency control (locking) aspects of transaction management, data integrity assertions/constraints, query optimization and processing, and query specification interfaces and languages.

Among the above set of publications, the System R paper by Gray, Lorie and Putzolu has had the most long-term impact of great significance on the database industry [16]. It dealt with one very important component of transaction management, namely the orderly execution of concurrent transactions via more sophisticated locking modes than before to support simultaneously multiple granularities of locking being done by different transactions. It comprehensively dealt with topics like implicit and explicit locking, intention locks, lock compatibilities and conversions, efficient management of lock requests, deadlock implications, index range locking and differences compared to predicate locking. During the subsequent decades after System R, and numerous IBM and non-IBM database products implemented those ideas, many of us have extended that work by defining new lock modes to improve concurrency, novel usage of the initial set of lock modes for index range locking concurrency improvements (ARIES/KVL and ARIES/IM), locking overhead reduction by leveraging recovery state tracking metadata (Commit_LSN) and data versioning (SQL Server), and reducing global locking overheads in shared disks environments by leveraging dynamically adaptive hierarchical locking (DB2 Data Sharing).

While the PRTV prototype described by Todd was a more narrowly focused effort with a limited set of accomplishments and less manpower, compared to System R, it was a pioneering early relational DBMS which showed the viability and the superiority of relational technology [35]. It represented a historic milestone in the evolution of relational technology. Since it was a single user system with batch updating of data, it didn't deal with the very complicated algorithms needed to support concurrent transactions and sophisticated crash recovery. It implemented some sophisticated query optimizations like rewrites, intelligent way of choosing a query execution plan and query processing efficiency gains via pipelining. Since it supported relational algebra as the query language, it wasn't as user friendly compared to other query languages or interfaces like IBM's QBE, System R's SQL or UC Berkeley Ingres project's QUEL. A product named BS12 was released based on PRTV, but it didn't have the long-lasting impact of products based on System R like DB2 and SQL/DS.

Zloof's paper on QBE describes a system that was developed in parallel with System R, as a competing RDBMS [37]. Instead of a fully textual query language, QBE supported a very novel graphical user interface for constructing queries via examples, thereby targeting non-programmers. It is based on the concept that the user formulates his query by filling in the appropriate skeleton tables through a keyboard display with an example of a possible answer. It allowed updates also to be performed using the same interface. While QBE was developed as a full-fledged standalone DBMS, its strength was in the user interface area. At one time, internal to IBM, a contest was held to choose between System R and QBE as the DBMS to focus on for the long term. System R won and that resulted in QBE never being released as a DBMS by itself. When DB2 was released on the mainframe operating system MVS, some of the QBE query construction functionality was productized as a layer top of DB2 and it was called QMF. As RDBMS functionality became more sophisticated and hence SQL was extended extensively, supporting the paradigm of query by example became very difficult and the support for QBE-like features was terminated. Different ways of helping naïve users emerged in the form of 4GL tools. QBE's most important impact on the world was bringing the focus on ways to help non-programmers.

The paper by Eswaran and Chamberlin was focused on data integrity as a distinct topic compared to security, consistency and reliability that had been worked on extensively by others [14]. Various dimensions of that problem were discussed like tuple versus set orientation, immediate versus delayed checking, state versus transition orientation, and soft versus hard enforcement of integrity constraints. The definition of integrity constraints, data item compatibility and null values were also explored. While the ideas described were proposed in the context of System R and were preliminary in nature, my impression is that in the final version of System R implementation they were missing since the product version of the research prototype which was named SQL/DS didn't have them. Most of the ideas from this paper would have impacted the subsequent worldwide work on referential and other types of integrity constraints, triggers, event-condition-action (ECA) rule systems, etc.

4 PERSPECTIVES ON THE IMPACT OF VLDB

The third session had presentations by Renée J. Miller, Volker Markl and Lei Chen, three recent VLDB Endowment Presidents from around the world, who discussed the pioneering innovations and global influence of the VLDB conferences over the years.

4.1 Perspectives of Renée J. Miller

VLDB's global impact is rooted in its endowment bylaws that ensure that no one geographic region has majority control of the endowment, that trustees and endowment executives have strict term limits, and that election of new members is done transparently. These simple democratic principles have allowed VLDB to evolve and innovate to remain a flagship publishing venue while maintaining the trust of its diverse community.

In its early days, VLDB championed geographic affirmative action. There were three separate PCs for the three VLDB regions: the Americas, Europe/Africa, and Asia/Pacific. The PCs coordinated to promote underrepresented regions by favoring acceptance of papers from them. As the regions equalized in strength and representation, this policy was replaced by a new broadening strategy, led by John Mylopoulos and the endowment's future directions working group, which had concluded that database research had become too narrow and incremental, and that the distance to application and impact had grown substantially. The geographic PCs were replaced by two thematic PCs (Core DB Technology and Infrastructure for Information Systems) and a third, Industrial Applications and Experience. This cultural change took time and diligence to enact, but it was very successful. Another major innovation was spurred by a dwindling confidence in review quality and a feeling that there was no memory in the review system. The endowment was also responding to researchers who were in old-school institutions that did not value conference publications, only journals. The resulting innovation was PVLDB, a journal that publishes VLDB conference papers, using journal style reviewing.

All of these innovations are rooted in having an endowment in which board diversity and renewal are institutionally mandated, ensuring the injection of new ideas. These innovations have often been controversial and disruptive. They have succeeded, and VLDB has remained strong in large part due to its foundation in accountable, representative, and transparent governance.

4.2 Perspectives of Volker Markl

VLDB has been an important innovator in the conference and dissemination culture in the field of database systems. Key innovations out of VLDB that have impacted the database field and beyond are (1) the establishment of the Proceedings of the VLDB Endowment (PVLDB) and (2) the Experiments and Analyses track.

PVLDB was conceived and implemented by H. V. Jagadish, who served as its founding Editor-in-Chief from 2008 to 2014. Volume 1, Number 1 appeared in August 2008 for the 34th VLDB conference in Auckland. The idea was to dissolve the distinction between a conference and a journal that had long constrained our field. Authors submit year round rather than against a single annual deadline, papers are reviewed by a standing review board in a journal style process that allows revision rather than a binary verdict, and accepted papers are published in a monthly issue and presented at the following conference. The transition was completed with Volume 4, from which point every paper at the VLDB conference was accepted through the year round process, with no traditional conference program committee route remaining, and PVLDB moved to twelve issues per year. This solved two problems at once. It restored the possibility of revision, which conferences had lost and journals had never given up, and it removed the artificial scarcity created by a single annual deadline, which had made the reviewing load impossible to distribute and the outcome for authors needlessly arbitrary. The model has since been adopted well beyond our community, and in 2026 Jagadish received the ACM SIGMOD Edgar F. Codd Innovations Award.

The Experiments and Analyses (E&A) track, which I conceived and chaired over its first two years, was introduced in 2008 at VLDB as a new type of submission. The E&A track sought papers such as experimental surveys, the verification or refutation of previously published results, and the deeper analysis of a problem or phenomenon. Establishing it was more of a cultural change than an organizational one. The DM community had to accept that rigorous evaluation is a scientific contribution in its own right, and reviewing had to be recalibrated, since reviewers who were primarily trained to reward novelty would otherwise reject such work by default. To distinguish these contributions from traditional papers introducing novel work, the E&A track initially operated with its own program committee and its own assessment criteria. Over time, a shared understanding emerged, and the track transitioned into the regular research track as a category assessed by the standard review board. Today, Experiment, Analysis and Benchmark is one of four PVLDB paper categories, it carries its own best paper award, and analogous categories exist in both EDBT and SIGMOD.

VLDB's capacity for innovating follows directly from how the Endowment is governed and from its diversity. The Endowment is a non-profit organization with equal representation of professionals from industry and academia across the three regions "Americas", "EMEA", and "Asia/Pacific". It takes a truly holistic view of promoting and exchanging scholarly work in databases and related fields throughout the world. Because the Endowment answers to the community alone, a new form of contribution can be proposed, calibrated and, where warranted, retired within one or two cycles. The same responsiveness extends well beyond publication formats. The Endowment introduced the Women in Database Research Award, first conferred in 2016, to recognize female role models in the community, and in October 2018 the Endowment and the SIGMOD Executive Board jointly started DBCares, which committed our community to zero tolerance for abuse, discrimination and harassment. Through its travel grants, the Endowment ensures that community members from underrepresented regions can attend the conference, and it strives to regularly organize the conference in underrepresented regions as part of academic outreach.

I have had the honor to serve VLDB as Trustee, Secretary and President across a decade. I saw how much this governance matters in practice in order to shape and advance the community. I congratulate VLDB on its anniversary and hope that it will continue to address the new challenges that we face in the era of AI and geopolitical tensions. Two major challenges to tackle in the near future will be to (1) ensure paper quality and relevance in the time of AI-generated content and (2) maintain a truly global conference at times of nationalism and fragmentation of research and communities.

4.3 Perspectives of Lei Chen

Since its beginnings in 1975, VLDB has always aimed for a global reach—visiting cities like Tokyo, Rio, and Singapore in its first decade—but the core of the community remained largely rooted in Western academia. Over the decades, that landscape has transformed completely. These days, Asia isn't just participating in VLDB — in many ways, it's leading the charge.

I still remember VLDB 2014 in Hangzhou. That was a real turning point. Having the conference in China brought together researchers and practitioners from across the region in a way we hadn't seen before. Since then, VLDB has become a genuine bridge between East and West, which matters enormously when you're trying to solve data problems at the scale we face today.

I was General Chair for VLDB 2024 in Guangzhou, and the numbers tell a remarkable story. We had over 1,400 attendees — the largest VLDB ever. More than 900 of them came from Asia, which was another record. And regional sponsorship hit over 400,000 USD, again driven largely by Chinese companies.

We're here in Boston today celebrating VLDB's return to where it all began. But looking at those figures, it's hard to miss the reality: Asia has become central to this conference's future. The community there isn't just growing — it's deeply invested, both in terms of people and resources.

5 CONCLUSIONS

VLDB has thrived in its first 50+ years. The future is in the hands of the next generation of leaders and we are confident that they are up to the challenge of ensuring a bright future for the data management community and VLDB.

REFERENCES

[1] Information technology — database languages — gql. Technical Report ISO/IEC 39075:2024, International Organization for Standardization and International Electrotechnical Commission, April 2024.

[2] M. M. Astrahan, M. W. Blasgen, D. D. Chamberlin, K. P. Eswaran, J. Gray, P. P. Griffiths, W. F. K. III, R. A. Lorie, P. R. McJones, J. W. Mehl, G. R. Putzolu, I. L. Traiger, B. W. Wade, and V. Watson. System R: relational approach to database management. *ACM Trans. Database Syst.*, 1(2):97–137, 1976.

[3] A. Badia. Entity-relationship modeling revisited. *SIGMOD Rec.*, 33(1):77–82, 2004.

[4] C. Batini and M. Lenzerini. A methodology for data schema integration in the entity relationship model. *IEEE Trans. Software Eng.*, 10(6):650–664, 1984.

[5] L. Bellomarini, D. Fakhoury, G. Gottlob, and E. Sallinger. Knowledge graphs and enterprise AI: the promise of an enabling technology. In *35th IEEE International Conference on Data Engineering, ICDE 2019, Macao, China, April 8-11, 2019*, pages 26–37. IEEE, 2019.

[6] A. Bonifati, G. H. L. Fletcher, H. Voigt, and N. Yakovets. *Querying Graphs*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2018.

[7] H. Boral and D. J. DeWitt. Database machines: An idea whose time passed? A critique of the future of database machines. In H. Leilich and M. Missikoff, editors, *Database Machines, International Workshop, Munich, Germany, September 1983*, pages 166–187. Springer, 1983.

[8] P. P. Chen. The entity-relationship model: Toward a unified view of data. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, page 173. ACM, 1975.

[9] P. P. Chen. The entity-relationship model - toward a unified view of data. *ACM Trans. Database Syst.*, 1(1):9–36, 1976.

[10] P. P. Chen. English, chinese and ER diagrams. *Data Knowl. Eng.*, 23(1):5–16, 1997.

[11] E. F. Codd. A relational model of data for large shared data banks. *Commun. ACM*, 13(6):377–387, 1970.

[12] A. Deshpande. Beyond relations: A case for elevating to the entity-relationship abstraction. In *15th Conference on Innovative Data Systems Research, CIDR 2025, Amsterdam, The Netherlands, January 19-22, 2025*. www.cidrdb.org, 2025.

[13] D. J. DeWitt, R. H. Gerber, G. Graefe, M. L. Heytens, K. B. Kumar, and M. Muralikrishna. GAMMA - A high performance dataflow database machine. In

W. W. Chu, G. Gardarin, S. Ohsuga, and Y. Kambayashi, editors, *VLDB'86 Twelfth International Conference on Very Large Data Bases, August 25-28, 1986, Kyoto, Japan, Proceedings*, pages 228–237. Morgan Kaufmann, 1986.

[14] K. P. Eswaran and D. D. Chamberlin. Functional specifications of subsystem for database integrity. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 48–68. ACM, 1975.

[15] C. Fahrner and G. Vossen. A survey of database design transformations based on the entity-relationship model. *Data Knowl. Eng.*, 15(3):213–250, 1995.

[16] J. Gray, R. A. Lorie, and G. R. Putzolu. Granularity of locks in a large shared data base. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 428–451. ACM, 1975.

[17] T. Haigh. David K. Hsiao Interview: June 12-13, 2009; Mountain View, California. In *ACM Oral History Interviews*, New York, NY, USA, 2012. Association for Computing Machinery.

[18] S. Hartmann and S. Link. English sentence structures and EER modeling. In J. F. Roddick and A. Hinze, editors, *Conceptual Modelling 2007, Proceedings of the Fourth Asia-Pacific Conference on Conceptual Modelling (APCCM2007), Ballarat, Victoria, Australia, January 30 - February 2, 2007, Proceedings*, volume 67 of *CRPIT*, pages 27–35. Australian Computer Society, 2007.

[19] D. K. Hsiao. Will the real VLDB conference please stand up? *SIGMOD Rec.*, 12(3):8–14, 1982.

[20] H. Koehler and S. Link. Entity/relationship profiling. In *40th IEEE International Conference on Data Engineering, ICDE 2024, Utrecht, The Netherlands, May 13-16, 2024*, pages 5393–5396. IEEE, 2024.

[21] C. S. Lin, D. C. P. Smith, and J. M. Smith. The design of a rotating associative array memory for a relational database management application. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 453–455. ACM, 1975.

[22] J. O. K. Menezes and C. E. C. Campelo. Database modeling automation from natural language requirements. In C. E. S. Pires, E. P. M. de Sousa, and M. A. Brandão, editors, *Proceedings of the 40th Brazilian Symposium on Databases, SBBD 2025, Fortaleza, CE, Brazil, September 29 - October 2, 2025*, pages 658–670. SBC, 2025.

[23] E. A. Ozkarahan, S. A. Schuster, and K. C. Smith. RAP - an associative processor for database management. In *American Federation of Information Processing Societies: 1975 National Computer Conference, 19-22 May 1975, Anaheim, CA, USA*, volume 44 of *AFIPS Conference Proceedings*, pages 379–387. AFIPS Press, 1975.

[24] A. Pan, X. Wang, and H. Li. Conceptual modeling on tencent's distributed database systems. In J. Trujillo, K. C. Davis, X. Du, Z. Li, T. W. Ling, G. Li, and M. Lee, editors, *Conceptual Modeling - 37th International Conference, ER 2018, Xi'an, China, October 22-25, 2018, Proceedings*, volume 11157 of *Lecture Notes in Computer Science*, pages 12–24. Springer, 2018.

[25] T. Priebe and S. Markus. Business information modeling: A methodology for data-intensive projects, data science and big data governance. In *2015 IEEE International Conference on Big Data (IEEE BigData 2015), Santa Clara, CA, USA, October 29 - November 1, 2015*, pages 2056–2065. IEEE Computer Society, 2015.

[26] M. E. Senko. Specification of stored data structures and desired output results in DIAM II with FORAL. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 557–571. ACM, 1975.

[27] P. Skavantzos and S. Link. Entity/relationship graphs: Principled design, modeling, and data integrity management of graph databases. *Proc. ACM Manag. Data*, 3(1):40:1–40:26, 2025.

[28] P. Skavantzos and S. Link. When entity/relationship models meet graph databases. *Proc. VLDB Endow.*, 18(12):5444–5447, 2025.

[29] M. Stonebraker, E. Wong, P. Kreps, and G. Held. The design and implementation of INGRES. *ACM Trans. Database Syst.*, 1(3):189–222, 1976.

[30] V. C. Storey, R. Lukyanenko, and P. P. P. Filho. Conceptual modeling: Topics, themes, and technology trends. *ACM Computing Surveys*, 55(14s):317.1–317.38, 2023.

[31] V. C. Storey, J. Parsons, A. Castellanos, M. C. Tremblay, R. Lukyanenko, A. Castillo, and W. Maaß. Domain knowledge in artificial intelligence: Using conceptual modeling to increase machine learning accuracy and explainability. *Data Knowl. Eng.*, 160:102482, 2025.

[32] S. Y. W. Su and G. J. Lipovski. CASSM: A cellular system for very large data bases. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 456–472. ACM, 1975.

[33] Teradata Corporation. *DBC/1012 Data Base Computer Concepts and Facilities*. Teradata Corporation, 1983.

[34] B. Thalheim. *Entity-relationship modeling - foundations of database technology*. Springer, 2000.

[35] S. Todd. PRTV: an efficient implementation for large relational data bases. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 554–556. ACM, 1975.

[36] M. Winslett. Peter chen speaks out: on paths to fame, the roots of the er model in human language, the er model in software engineering, the need for er databases, and more. *SIGMOD Rec.*, 33(1):110–118, Mar. 2004.

[37] M. M. Zloof. Query-by-example: the invocation and definition of tables and forms. In D. S. Kerr, editor, *Proceedings of the International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA*, pages 1–24. ACM, 1975.