My 10 Fears about the Future of the DBMS Field

(with apologies to David Letterman)

By

Michael Stonebraker

Outline

- The 3 biggies (3)

- Five more that are a direct result of the biggies (5)

- The big enchilada (2)

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Fear #1 – The Hollow Middle

- Consider the DBMS field as it was defined 40 years ago (in 1977)
  - storage structures
  - query processing
  - security
  - integrity
  - query languages
  - data transformation,
  - database design
  - data integration

Fear #1 – The Hollow Middle

• Now count the percentage of SIGMOD papers that deal with this "core" over time

|	
1977	(100%)
1987	(93%)
1998	(68%)
2008	(52%)
2017	(47%)

The Conclusion

- We are drifting into "applications"
  - As the core becomes increasingly well understood

- Which have little-to-nothing to do with each other:
  - E.g. NLP
  - E.g. autonomous vehicles
  - E.g. complex analytics

- Effectively a "multi-furcation"
  - Reminiscent of the PODS/SIGMOD bifurcation 35 years ago

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Pictorially – The Hollow Middle

- 1977
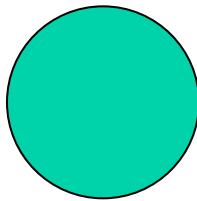
- 2017
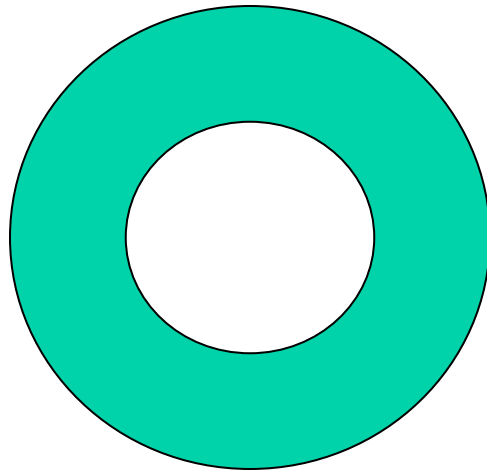
The Conclusion

- We are moving to a world where nothing binds us together except 200 researchers looking favorably on each other's papers

- Time to formally multi-furcate!
  - Multiple PCs

- Otherwise DBMS Systems may well declare a divorce
  - And start their own conference
  - May start a trend

Fear #2: We Have Been Abandoned by our "Customer"

- 40 years ago

  – "Industry types" came to our conferences

  – Who were "pathfinders" from financial services, oil&gas, insurance, etc. looking for DBMS solutions

  – They participated (in panels, program committees, etc.)

  – I.e. they tended to keep us "honest", i.e. working on relevant problems

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Fear #2: We Have Been Abandoned by our "Customer"

- Now

  — These industry types have largely disappeared, presumably because our research conferences are no longer relevant to their needs

  — They have been replaced by either:
    - "the whales", who represent 0.01% of DBMS users
    - Or nobody (we are talking to ourselves)

  — I.e. we have become disconnected from the real world
    - A very dangerous state of affairs, indeed!!!!
    - Obvious solution: get him/her back!

Fear #2: We Have Been Abandoned by our "Customer"

- The whales

  – Are inhabited by IT professionals way smarter than the average
  – Have a NIH syndrome (build, rather than buy)
  – And their own unique needs
  – (Until recently) knew next to nothing about DBMS technology

- Not an exemplar of a DBMS customer…..
- But the alternative is "nobody" -- i.e. we make stuff up

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Fear #3: Diarrhea of Papers

- I got a PhD in 1971 from Michigan
  - With zero publications

- I got tenure from Berkeley in 1976
  - With about 5 papers

- Others in my peer group (e.g. Dave Dewitt) report similar numbers

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Fear #3: Diarrhea of Papers

- Now both numbers are up by an order of magnitude
  – A PhD applicant who wants to get a good academic job should have ~10 papers
  – A tenure case should have ~40 papers

- And there are around 10X the number of researcher
  – I.e. 100X the number of papers
  – Nobody can keep up with this deluge

Fear #3: Diarrhea of Papers

- Also, everybody divides their papers into least publishable units (LPU's)

  – To generate the 40 tenure papers….
  – The result is few seminal papers; lots of LPUs (which contribute to the deluge)
  – If you know a seminal paper in the last 10 years, please tell me…

Fear #3: Diarrhea of Papers

- A student has to grind out ~10 papers in 3-4 years
- Any serious implementation is impossible in this climate
  - Takes too much time
  - Focus on "quickies"
  - i.e. incrementalism on steroids!!!
  - And this tilts us toward theory papers (which are quicker to write)
- Postgres would be impossible in this climate
  - Way too much work for too few publications

Fear #3: Diarrhea of Papers

- Where did this come from?
    - Lazy Deans and Department Chairs, primarily from the Far East
    - Western universities also share some of the blame
- My fear: it is getting worse!
- Solution:
    - Require Asst. Prof applicants to submit resumes with at most three papers on them
    - Tenure cases: say 10
    - Get top tier universities to adopt this restriction, and others will follow

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Five Ancillary Fears –

- Which follow from the 1st three biggies

Fear #4: Reviewing is Getting Very Random (Follows from Hollow Middle and Diarrhea)

- Quality stinks
  - Huge program committees
  - Overwhelming crush of papers
- Leads to a revise and resubmit (R&R) mentality
  - Which is killing both the submitters and the reviewers
- Solution
  - Address Fears #1 and #3
- My real fear
  - If current system persists, it will collapse under its own weight

Fear #5: Research Taste Has Disappeared (Follows from Customer Leaving and Whale Replacement)

- I am amused at the acceptance and subsequent rejection of:
  - OleDB,
  - MapReduce,
  - Eventual consistency,
  - The semantic web,
  - Object databases and
  - XML
- Just to name a few

Fear #5: Research Taste Has Disappeared

- We are way too uncritical of the whales

- Otherwise, "people who don't understand history will be condemned to repeat it"

- We need desperately to reconnect with the "real world"
  - Requires recruitment
  - And free registrations···.

Fear #6: We are Polishing a Round Ball (Follows from Diarrhea)

- Fear #3 drives required paper production
- Which drives quickies
- Which drives incrementalism
  - 10% papers are not worth reading!
  - Drives us to be disconnected from the real world

- An unrealistic solution
  - Require every Asst Prof to spend a year in industry

Fear #7: Irrelevant Theory is Taking Over (Driven by Diarrhea and Customer Abandonnement)

- Fact: It is very difficult to get systems papers accepted
    - Seemingly because they have no theory
    - Which will probably cause those folks to "divorce" the traditional conferences
- When we stop being driven by the real world
    - Then problems become arbitrary
    - And the best theory wins
- Also driven by quickies

Fear #7: Irrelevant Theory is Taking Over

- Real world doesn't share data
    - Considers it proprietary
    - So do the whales

- Systems papers are becoming more artificial
    - Driven by artificial benchmarks (e.g. YCSB)
    - Or "off the beaten path" benchmarks (e.g. Wikipedia)

- Very ominous situation

Fear #8: We are Ignoring the Most Important Problems
(Driven by Diarrhea and Customer Leaving)

- In favor of ones that are easy to solve
  – Driven by the need for quickies
- Examples
  – Data integration
  – Data base design and evolution
  – Tuning a DBMS application is way too difficult
  – Average DBMS takes $20M to get to production readiness.

Fear #8: We are Ignoring the Most Important Problems

- A consequence of being detached from the real world
- A field is defined to have "lost its way" when it forgets who the customer is.

Fear #9 and #10: The Big Enchilada
(A USA Perspective)

- Most research is in universities
- Quality of life in universities driven by
  - Money
  - Student overload

Fear #9: Research Support is Disappearing (USA perspective)

- NSF success rate down to 7%
  - And the number of mouths to feed is increasing

- American industry is not picking up the slack
  - Foreign industry is taking over
  - Thereby endangering US preeminence

- We are not likely to remain at the top of the heap
  - Without significant intervention from our government
  - Which is turning a blind eye

- Reminiscent of the Roman Empire

DBg Database Group
MIT Computer Science and Artificial Intelligence Lab

Fear #10: Student Load
(A USA Perspective)

- ## CS departments are overrun with students

  – 40% of MIT undergraduates are majoring in CS!

- ## Will likely make academic life less and less attractive off into the future

  – Best people will depart for greener pastures

  – And a lot more money

- ## With unknown impact on our DBMS field

  – What happens if the whales hire the "best and the brightest?

  – The Donald Kossmann effect….

Summary

- ## Most of my fears are addressable
  - But require hard decisions
  - By inspired leaders
- ## This paper is a call to action to the field leaders
  - Now is the time to fix stuff
  - Diarrhea and fragmentation are the most important ones