

序

与传统的谚语“三十年河东、三十年河西”相比，计算技术的变革突飞猛进，似乎呈现着“十年河东、十年河西”的规律，微软 δ 谷歌 δ 脸谱的发展大致可以印证这一事实。

因此从事这个领域的研究以十年一个周期布局自己的研究方向应较为合理，变换周期太快不宜成果的积累，变化周期太长容易跟不上发展的节奏，从而与主流技术脱节。这与“*Outliers: The story of success*”一书中的10年一万个小时专注做一件事的秘诀是一致的。

新世纪以来，数据库界普遍面临的一个问题是，在传统的数据库技术成熟之后，数据库研究应向何处去？凭借自己对当时技术趋势的判断，我将研究目标定位在解决数据库技术与 Web 计算和移动计算交叉结合所产生的挑战性问题上，即结构多样的 Web 数据管理，半结构化 XML 数据的管理，以及移动环境下的数据管理问题，并创立了“网络与移动数据管理实验室（**Web and Mobile Data Management**）”，致力于这方面的研究，取得了一些国内外所共知的研究成果。我把这一阶段的研究概括为创新数据管理研究 1.0。

今年是又一个十年的伊始，我一直在思索实验室下一个十年的研究布局。我们不难发现数据库技术的变革（其实任何信息技术亦如此）主要来自三方面的驱动力，即：计算模式，硬件技术，应用模式的不断创新。基于新的三方面驱动力的需求，把对下一个十年的研究概括为创新数据管理研究 2.0，具体包含如下的研究方向：

1) 闪存数据库系统的研究。它来自硬件技术变革的驱动力，其研究目标是针对闪存硬件特性、灵活的应用模式和传统数据库技术的不足，研究全新的闪存数据库管理技术；

2) 云数据库系统的研究。它来自计算模式变革的驱动力，其研究目标是实现一种具有灵活配置、高可用性、高容错性、可扩展性和高性能的云数据库系统；

3) **Web** 与社会计算的研究。它来自应用模式变革的驱动力，其研究目标是把社会计算的方法引入 Web 数据管理，解决 Web 信息的可信和隐私保护问题；

4) **Mobile** 与隐私保护研究。它来自应用模式变革的驱动力，即 **Mobile Web** 需求日益迫切，其研究目标是解决移动搜索、隐私保护等关键问题；

5) 纯 XML 数据库系统研制。过去 10 年我们系统研究了纯 XML 数据库技术，获中国计算机学会“王选奖”，我们将寻求技术转移的途径，进行产业化尝试。

数据库系统发展经历了三十年，大致呈现出了“分久必合、合久必分”规律。六七十年代广泛的应用需求的出现促成了各类数据库系统的产生。八九十年代大型网络分布计算环境的普及使得政府、企业的应用需求趋同，导致几大数据库系统的“大一统”局面出现。当下互联网特别是云计算的出现，使得应用需求再趋多样化，人们更期盼与自己的需求功能相宜的数据库系统，而不是面面俱到的“大拼盘”系统，多样化时代重新到来。最近日渐火爆的“NoSQL”运动正是迈向这一目标的尝试。我们在本年度报告里试图把这些我们观察到的、看明白或没看明白的一些问题总结成短文，与大家交流，抛砖引玉。

本年度，实验室开了一个好头。新获国家自然科学基金项目 1 项，IBM 和 MSRA 资助项目各 1 项，出版英文专著一部，国际会议论文集两部，组织期刊 JCST 专辑“Trends Changing Data Management”，发表相关方向高水平科研论文 20 余篇（包括 IEEE TKDE、Pervasive and Mobile Computing 等），获专利授权 2 个，新申请专利 6 个。在系统开发方面，我们进一步深化原有的中文文献集成系统 C-DBLP，试图构筑一个“ScholarSpace”，它将由三部分组成：SearchScholar+EasyScholar+SocialScholar。这里将充分利用我们 Web 数据集成的技术，并融入社会网络的思想。希望这一成果能为学界同行提供更好的服务。

作为大会主席在人大成功举办第 27 届中国数据库学术会议（NDBC），特别是我们邀请到二十多位七八十岁的老专家共聚“中国数据库历史回顾和萨师焯教授追思会”，是特别感到欣慰的一件事情。感觉能为同行提供有价值的交流服务是一种享受，感谢实验室老师同学为此付出的巨大辛劳。

这里谨以此集感谢来自学校方方面面的支持，感谢国家自然科学基金委和 863 计划的资助，感谢所有关心和支持过我们的人们。

孟小峰

2010 年 12 月 24 日于北京