

第7篇

赫克托·加西亚·莫利纳访谈录

Interview with Hector Garcia-Molina





赫克托·加西亚·莫利纳简介

赫克托·加西亚·莫利纳（Hector Garcia-Molina）是斯坦福大学计算机科学与电子工程系的列昂纳德·波萨克（Leonard Bosack）和桑德拉·勒纳（Sandra Lerner）教授，从 2001 年 1 月到 2004 年 12 月担任该校计算机科学系主任。从 1997 年到 2001 年，他担任美国总统信息技术顾问委员会（PITAC）成员。1994 年 8 月到 1997 年 12 月，他担任斯坦福大学计算机系统实验室主任。从 1979 年到 1991 年，他在新泽西州普林斯顿大学担任计算机科学系的教师。赫克托·加西亚·莫利纳在 1974 年于墨西哥蒙特雷理工学院获得电子工程专业学士学位，分别在 1975 年和 1979 年于斯坦福大学获得计算机科学硕士和博士学位，并在 2007 年获得瑞士联邦理工学院名誉博士学位。赫克托·加西亚·莫利纳在数据库系统、分布式系统和数字图书馆领域中发表了大量论文，研究兴趣包括分布式计算系统、数据库系统和数字图书馆等。他是 ACM 会士、美国艺术与科学院院士和美国国家工程院院士，并在 1999 年获得了 ACM SIGMOD 创新奖。

本专访主要介绍了初创公司面临的问题，研究工作面临的困难，对改进性论文的看法，墨西哥的计算机科学发展状况，系主任的生活等，更多内容请看访谈录。



玛丽安·温丝特

欢迎来到本期 ACM SIGMOD Record 数据库领域杰出人物访谈。本期访谈对象是赫克托·加西亚·莫利纳（Hector Garcia-Molina），访谈分别在 2001 年 6 月（面对面）和 2001 年 10 月（电子邮件）进行。赫克托·加西亚·莫利纳是斯坦福大学计算机科学与电子工程系的列昂纳德·波萨克和桑德拉·勒纳教授。在去斯坦福大学之前，他曾经在普林斯顿大学工作。赫克托·加西亚·莫利纳是 ACM 会士，并在 1999 年获得 ACM SIGMOD 创新奖。他还是美国总统信息技术顾问委员会成员、美国国家工程院院士和甲骨文公司董事会成员。

在面对面访谈过程中，我们没有足够时间讨论完所有有趣的话题，因此我们之后又进行了一次电子邮件访谈。因此，本访谈录的前一部分是我们面谈的内容，后一部分是通过电子邮件讨论的内容。

玛丽安·温丝特：您认为当今最鼓舞人心的技术发展趋势是什么？

赫克托·加西亚·莫利纳：我有一个很感兴趣的研究课题——我不知道它有多大意义——是点对点系统。我一直对分布式计算很感兴趣。尽管我一直在数据库领域工作，但很大一部分工作和分布式系统有关——当数据存储

在不同的地方，或者数据的备份在不同的计算机上，如何去处理这些数据，如何同步和协调它们。

玛丽安·温丝特：您提到的“点对点”是什么意思？

赫克托·加西亚·莫利纳：点对点（peer to peer）是一个模糊的术语，人们对它有不同的理解。我认为点对点是指在一个计算机集群中，所有的节点具有相同的能力，没有专门的节点控制整个系统或者管理全局索引来供所有节点访问。当然，集中型系统给很多应用带来好处，但是一个所有节点地位相同的系统也有它的优势。研究点对点系统非常有趣，因为它带来很多挑战性的问题。例如：如何在这样的一个系统中寻找资源。在点对点系统中，你不能只是通过 Napster 索引查询“谁拥有这首歌”。你需要通过和其它节点通信得到“谁可能拥有这首歌”的信息。因此，点对点系统中的具体问题非常复杂，但是研究这些问题真的很有趣。

玛丽安·温丝特：有人说如今数据库领域真正激动人心的研究发生在工业界，您对此有何看法？

赫克托·加西亚·莫利纳：我不同意这种看法。在工业界确实有一部分很有意思的工作，但当我去了解公司——特别是初创公司——在做什么时，我发现他们很少有时间去做真正有趣和有意义的事情。他们总是试图用最简便的方法解决问题，越快越好。他们不会花时间考虑为什么采用这种方式解决问题，或者是否还有其它方法能够解决同样的问题。他们只在乎能否不深入问题本质而得到解决方案。而且我认为公司所做的事情不像我们在学校里做的事情那样有趣和有价值。当然，一部分公司或者研究院会花时间做详细的分析。但是，如果没有人花时间去理解的话……换句话说即使我们学术界不提出新的想法，我们只是对工业界的经验进行分析和研究，重新组织这些想法使得人们可以学习和改进，这也是一种贡献。我认为这是很大的贡献。然而，人们在讨论公司所做的研究时忽略了学术界的这种研究。大部分初创

公司并不做真正的研究。他们想出了一些东西，然后迅速实现，但是他们并不做真正的科学研究。

玛丽安·温丝特：工业界和学术界这种功能上的分离是否带来问题，或者这种情况是否正常？

赫克托·加西亚·莫利纳：我认为这不是问题。我们并没有泾渭分明地分开，我们和公司有很多交流。他们来学习我们的工作，我们去和他们讨论。我觉得这很好。但是，我并不同意这样一种越来越流行的观念，就是所有在数据库或者计算机科学领域的研究都是在公司里完成。

玛丽安·温丝特：您认为现在有一些有价值的问题人们还没有给予足够的重视？

赫克托·加西亚·莫利纳：一个重要的但是人们没有花足够时间研究的问题是用户界面问题，或者是一些与用户相关的问题。我们如何使得数据库技术更方便、更容易使用？你怎样用更简单的方式展示数据库里的数据？你怎样操纵数据而不必成为数据库专家？这里面有很多问题需要解决。我想人们会抱怨与用户相关的问题很难发表论文，至少在数据库领域很难发表。用户和用户界面的主题确实难于发表文章，而且很难量化你所做工作的贡献和意义。另一方面，如果你研究的是性能问题，你可以提出一个算法，然后说，好的，使用我的算法一秒钟可以处理十个查询请求，而不使用我的算法一秒钟只能处理五个请求。因此我的算法有意义。但是对于用户界面问题，很难量化你所做研究的贡献。也许你节省了用户的时间，但是如何去衡量呢？

玛丽安·温丝特：用户相关的问题是否更容易在工业界解决，因为他们一定意义上更接近用户？

赫克托·加西亚·莫利纳：同样，我认为工业界没有真正研究性的工作。他们想出一个主意并试图去销售它。如果是个好想法，也许他们会赚钱。即

使是个坏想法，如果有好的销售人员，他们仍然会赚钱。但是我们永远不知道：产品成功是仅仅依靠有良好的销售并且在竞争中获胜，还是因为它确实是一个好的有用的想法？

我认为工业界的人们会在人机交互领域有很多好的想法，但我不认为他们在测试这些想法、性能评价和真正理解什么是正确的技术方面比学术界的人有优势。

玛丽安·温丝特：当您说这些的时候，让我想起了 Mosaic，第一个流行的浏览器。从某种意义上说，那只是一个成功的用户界面工程，但是没有涉及到研究工作。

赫克托·加西亚·莫利纳：是的，这就是它成功的原因，例如还有一个叫 Gopher 的浏览器。而 Mosaic 有一个更友好的用户界面。

玛丽安·温丝特：作为一名教授，您需要同时处理好工作、家庭和社会责任之间的关系。您在普林斯顿大学而不是在斯坦福大学拿到了终身教职的资格，但是我听说在斯坦福大学一些获得终身教职资格的人离婚了。您对这些需要处理工作和家庭关系的人有什么建议？

赫克托·加西亚·莫利纳：我想我是属于在普林斯顿大学获得终身教职资格并且离婚的一类人。

玛丽安·温丝特：我不知道这个情况。

赫克托·加西亚·莫利纳：这两件事不是同时发生的，是不久后。

玛丽安·温丝特：我想这方面您是有经验的。您对要经历这个过程的人有什么建议？

赫克托·加西亚·莫利纳：经历离婚，还是经历终身教职资格？

玛丽安·温丝特：终身教职资格！

赫克托·加西亚·莫利纳：这真的很难给出意见。我给我们系青年教师的意见是：不要把终身教职看得太重。当然，你必须要做很多努力才能获得终身教职资格。但是你决定要做哪方面的研究工作要基于你是否喜欢这个主题，你是否很享受工作的过程。如果你做任何工作仅仅是为了拿到终身教职资格，那是非常令人沮丧的。我认为如果你从事的是喜欢的工作，那么是否拿到终身教职资格不是至关重要的。如果你没有拿到终身教职资格，你可以去其他地方继续做感兴趣的研究。

玛丽安·温丝特：不知道为什么我所认识的科学家中来自墨西哥的要少一些，而更多的是来自于加拿大。

赫克托·加西亚·莫利纳：墨西哥没有做研究的传统。当你去那里的学校，你会发现人们对赚钱更感兴趣。我想这在很多国家都很常见，如果你挣很多钱就会很有威望并且被认为很成功，而在一所大学教书或者做研究则不会。至少当我在墨西哥读书时，那里没有做研究的传统，可能现在情况有所改变。

玛丽安·温丝特：那么您是如何走上研究道路的呢？

赫克托·加西亚·莫利纳：部分原因是我的家庭有做研究的传统。我父亲是一个大学教授，也是为数不多的重视研究工作的人。我从家庭中得到了做研究的概念。

玛丽安·温丝特：您现在能谈谈当初作为一个计算机科学研究新人时最希望知道什么吗？您对那些刚刚起步的人有什么建议吗？

赫克托·加西亚·莫利纳：现在看来做研究变得更困难，特别是在论文发表方面。当我毕业的时候，有一两篇或者三篇论文是很了不起的。现在，如果你没有十篇论文，人们会说，“你怎么回事？”我只是有点夸张，可能不是十篇。

玛丽安·温丝特：我了解，我也发现了。

赫克托·加西亚·莫利纳：我想现在的研究生面临着更多的压力，一方面是因为他们需要发表更多的论文，另一方面是工业界在相关的研究方向也很活跃。原先我认为我们只需要在感兴趣的问题上做一些我们认为有意义并且有趣的工作，如果你发表了论文，那很好，如果没有发表，也可以。但是现在看来情况更加艰难。

玛丽安·温丝特：这种情况是好还是不好呢？

赫克托·加西亚·莫利纳：我不知道，只是更加艰难了。

玛丽安·温丝特：您对数据库研究人员和从业人员有什么建议吗？

赫克托·加西亚·莫利纳：在得到终身教职资格之前，大家会抱怨助理教授的工作太多，然后试图去获得终身教职。这样你就直接或者间接地形成一种思维，认为一旦拿到了终身教职，事情就会变得更好，生活不会再那么繁忙。

但是情况并不是这样，因为一旦你得到了终身教授资格，他们会让你加入委员会，会让你做更多的事情，而且人们对你的期望会更高。因此，生活依然繁忙甚至会变得更糟。

玛丽安·温丝特：您认为我们应该让助理教授了解生活的这种艰难吗？如果我们告诉他们，那么这些助理教授今天都会离任。

赫克托·加西亚·莫利纳：是的，我们不能告诉他们，千万不能。

玛丽安·温丝特：那么，这和研究生完成论文一样，他们并不知道他们的研究工作实际上变得越来越艰难。当您在斯坦福大学做计算机科学系主任时如何保证研究计划正常进行？

赫克托·加西亚·莫利纳：嗯，我们有一个很好的部门工作人员，他负

责所有的日常事务。因此，系主任一周工作两到三天，有时甚至更短。这为我的研究工作留下四天。我的学生可能不是很喜欢我做系主任，因为我经常周末与他们会面，但是我很棒的博士生，他们的研究工作非常出色。因此看起来一切都被安排得井井有条。另外，我的任期只有三年，因此在此期间希望我的头脑不会完全被官僚作风所腐蚀。

玛丽安·温丝特：我知道您在 PITAC 工作，而且最近担任甲骨文公司董事会董事。这些工作都包括什么？

赫克托·加西亚·莫利纳：PITAC 是美国总统信息技术顾问委员会，大概有来自学术界和工业界的 20 个成员。（吉姆·格雷是 PITAC 里数据库领域的另外一名成员。）我们的责任是给国会和总统提供信息技术方面的建议，特别是在研究领域。三年前我们写了一份报告，主要内容是建议研究资金在五年内应该翻倍，即从 1 亿美元增加到 2 亿美元。最初的两年，资金确实有了显著的增加，特别是在 NSF。而 NSF 的资金增长导致了 ITR 方案和其他措施的出现。联邦研究经费的前景未来可能不是很乐观，但我们的委员会继续工作，使政府认识到 IT 研究与发展对国家是至关重要的，同时监督分配的资金用在何处。

在 PITAC 工作给我提供了很多经验。我学到很多政府如何运作的知识。政府部门里有很多杰出的人员，但是他们需要全社会的支持。例如，需要让公众认识到科学技术对国家是非常重要的。同时让社会成员参加到政府活动中也是非常重要的，例如参与咨询委员会，或者在美国国家科学基金会、美国国防部高级研究计划局、美国航天局和能源部等部门工作，例如做一名项目经理。

在甲骨文公司，我刚刚加入其董事会。我觉得这对我来说是一个激动人心的机会，特别是因为甲骨文是我们数据库领域的领先和主导公司之一。（可能因为我现在加入了他们的董事会，我应当说是主导公司！）希望我能够在我们的研究领域和甲骨文公司之间架起一座桥梁，当然这个桥梁是双向的。

玛丽安·温丝特：您从斯坦福大学毕业，然后去了普林斯顿大学工作，最后又回到斯坦福大学。作为一名斯坦福大学的毕业生，回到“家”里工作的感觉如何？

赫克托·加西亚·莫利纳：刚开始的一段时间确实感觉很奇怪。刚到斯坦福时我被安排在原来导师的办公室。（他已经去美国国防部高级研究计划局工作三年了。）因此我在这间办公室总觉得很奇怪，好像坐在了错误的桌子上。另外，我已经离开了很长一段时间，因此没有人把我当一个毕业生看待。我现在的学生跟我开玩笑说，有一天他们将会来斯坦福大学当老师然后把我踢出现在的办公室。当然，我想他们是在开玩笑。

玛丽安·温丝特：有些数据库领域的人认为，现在斯坦福大学做出的工作没有以前的工作那么基础，本质上说是一种改进的工作，斯坦福大学更像一个初创公司的孵化器而不是从事基础性研究的地方。学生还和以前一样聪明，但是他们的工作质量有所下降，这可能是一个时代的标志。您对此有何看法？

赫克托·加西亚·莫利纳：具有这些看法的人是威斯康星大学的吧？开个玩笑，你知道我们在威斯康星大学麦迪逊分校有一个非常友好的“竞争对手”。不管怎样，我认为相对于一些其他部门，斯坦福大学并没有改变。例如，如果你到我们的网站上去看我们的论文发表情况，你会发现我们在原先发表论文的地方仍然保持着令人满意的发表率。

需要被质疑的是整个数据库领域发生了改变，现在更重视应用性的工作。我想这可能是真的，例如我接到一个顶级数据库会议对我们工作的评审意见，认为我们的工作不具有“商业前景”。

我也听有些人抱怨数据库领域流行“改进性工作”。我认为我们的领域和计算机科学的其他领域是一样的，而且这几年的改进性论文的数量也没有发生太大的变化。部分问题是我们的领域已经开始走向成熟，我们不需要太多

的“突破性”论文。你知道，在我们数据库领域，没有人可以发表一篇论文只是说，“嘿，我们重新运行了某某人的实验，只是为了证明他们是否是正确的，或者我们可以做得更好。”今天，评审将给你最低的分：这个工作已经做了，所以谁还在乎？我把这个叫做“刀耕火种”方法。一旦我已经发表了我的想法，没有人可以研究同样的想法，所有的东西都必须是新的。希望这种方法在将来有所改变。

玛丽安·温丝特：我听说一个重要的数据库会议开始采用双盲评审，导致一些著名数据库研究小组论文的录取数量一路下滑。这是一件好事吗？

赫克托·加西亚·莫利纳：我听到的正好相反，其他的一些会议尝试过双盲评审，结果发现没有任何变化，因此放弃了这个方法。我个人没有发现任何特别的东西。我自己不喜欢盲审，一方面不通过作者名字，人们也可以判断出文章来自哪里；另一方面，作为一个评审者，盲审让我感觉会议组织者不相信我的判断。

玛丽安·温丝特：您的 TSIMMIS 工程主要是进行信息集成。在这个领域还有什么未完成的研究工作？在人工智能领域信息集成的有效使用还有什么障碍？

赫克托·加西亚·莫利纳：信息集成确实还有很多难以解决的开放性问题。我相信 TSIMMIS 和现在的其他一些数据集成系统（像 Information Manifold、Garlic 等），都做出了贡献，但是我们对系统的要求越来越高。例如，现在人们希望把更多的语义信息考虑在内。随着 9·11 事件的发生，我发现信息集成的一个更大的作用，它在不同数据源的信息集成方面有着至关重要的作用，可以完成在单一数据库无法实现的“连接”工作。

（史英杰 译，孟小峰 审校）