

生提供学科和研究知识自不必说，企业导师的存在可以为教学培养提出各种需求和建议，从而培养出真正适合企业的人才，让人才走出校园就能快速步入社会。第二个意义，就是企业专家的参与可以为数据院提供一些业界资源，比如跟美团、百度等成立联合实验室，或者给学生提供企业真实数据和需求去做项目，再如为学生提供实习机会等。这样的设置是非常务实的。

记者：数据院成立之初，并没有任何经验可借鉴参考，在您参与建设期间，遇到过哪些困难？

刘震：数据院做到现在这一步，已经非常出色，也非常不容易。特别是作为一个虚体机构，数据院背负着“跨界融合”的使命，需要联合不同院系的师生资源，需要打通学界和业界的研究合作，这当中要耗费大量的精力去动员、协商。实现了

“跨界”之后，怎样“融合”又是一道坎，数据院当时完全是摸着石头过河，从一开始数据院没有一个自己的老师、学生、科研人员，到如今发展为联合经管、计算机、交叉信息、社科等多院系的老师授课，招生覆盖全校所有院系，成立医疗健康、科技、交通等十一个领域的科研中心，其中的艰难不言而喻。

但是，时代变化是很快的，如今各高校也开始做大数据教育，数据院就需要不断创新，不断在探索中寻求新的路径。例如数据院目前还是一个非实体的研究院，我曾经好几次跟学校提出希望清华考虑将数据院转为一个实体的院系，因为就目前数据院的探索经验，很多事情是一个虚体院系比较难推进的，这应该是数据院和学校都要考虑的新方向。☞

数据科学的积累：“海平面下的冰山”

——访清华大学信息技术研究院语音和语言技术研究中心主任郑方

◆ 本刊记者 杨璐

人工智能、大数据、云计算是技术发展的新趋势。这三者既有相互的联系，也有相互的支撑。人工智能研究的是如何用计算机对人的智能进行模拟、延伸和扩展，进而让机器帮助人类做更加复杂的工作。大数据关心的则是如何从海量数据中挖掘出有价值的知识，而这些数据用传统方法难以找到规律，需要用新方法来分析，比如通过例子学习（learning by examples），就像人小时候学习语言，虽然不会语法，但能通过一个个例子

学会说话。通俗地讲，大数据提供“训练的数据”；人工智能提供“训练的技术”；云计算提供“训练的平台”，既提供底层技术的支持，也有超算的能力，还提供安全的存储。

从AlphaGo出现之后，人工智能和数据科学变得非常热，然而现在的“热”属于外部的热。其实在很早的时候，内部（学术研究者）就一直在积累，一直在发热。以前，人们在怀疑AI为何研究了六十多年，一直没有成果；AlphaGo

出现之后，人们又觉得 AI 无所不能，这是从一个极端走到了另一个极端。其实现在 AI 的发展水平，一直都在按照科学的规律在螺旋式地逐步发展，一点一点积累，AlphaGo 并没有使其在技术层面出现突变，其贡献在于把 AI 拉进人们的视野。人工智能有很多不同的分支，诸如自然语言处理、计算机视觉，还有我的研究方向语音信号处理，每个分支都有不同的特点，规律并不完全相同。然而现阶段并没有一个能够解决所有问题的通用的理论和工具，因此比较理性的做法，是具体问题具体分析，以应用为驱动，从问题出发，用合适的理论、模型、方法等，才能找出比较好的解决办法。

在人工智能和数据科学前进的路上，我认为清华人一直是一个开拓者的形象。清华人做事有一种情怀，为了解决一个实际问题，可以十年如一日、非常专注地做研究，而不受外界的诱惑和影响。人工智能的很多问题，靠大量的资金不一定能解决，靠市场上的经验积累也不一定解决。那么凭借什么？从根源上讲，无论是数据科学还是人工智能，其发展必须有深厚的基础研究作为支撑，而清华在这一点很有优势，有很多老师在基础研究方面耕耘了至少二三十年。市面上的一些产品和应用，背后是清华这座“海平面下更巨大的冰山”，不断地为这些应用提供原动力和技术积累。

而在这个过程中，数据科学研究院则处在一个纽带的角色，连接校内研究者和校外产业界的纽带。数据院既了解老师们的输出，也知



清华大学信息技术研究院语音和语言技术研究中心主任郑方

道社会的需求，既能为老师创造成果转化的条件，又能帮助企业和社会解决实际的问题。这就是产学研的融合，不仅仅局限在表层的合作，而是更紧密的耦合，我把它称作“化合态”的“产学研”。刚才讲到清华人的情怀，老师们在这个研究方向深扎几十年，无问西东，所以才有很多的成果、很多的输出。我们需要给他们一个契机、一个平台、一个机制把它们成功地转化出来，落地到社会服务上面。“得意音通”就是一个例子，为什么它在语音这个领域竞争力特别强？这跟资金没关系。很多国际公司包括国内的上市公司，远远比它有钱。得意音通的竞争力就体现在“技术积累”方面，这个不是仅仅靠钱能砸出来的。这个才是关键。

说到技术积累，我想谈一谈“实践教学”这个教学中很重要的环节，当然这也是数据科学研究院在人才培养方面的一大特色。实践课我也认为很有必要。信息社会发展迅速，一定不能关在象牙塔中不接触外界，要适当地进行实践，从而更深入地理解所学的知识，从而进一步理解和体会学习的目的。把基础打牢，有扎实的能力，毕业后即便做新东西也得心应手，后劲非常强。诸如人工智能发展的新技术、大数据发展的新技术，其中很多技术的基本原理是一样的，数学、物理、信息领域等很多基础都会用上。

随着大数据和人工智能等技术的发展，相信清华将在行业中扮演越来越重要的角色。肩负着责任感和使命感，师生们也将共同在学科交叉、技术融合的舞台专注耕耘，解决新时代的问题，更好地服务社会。❏