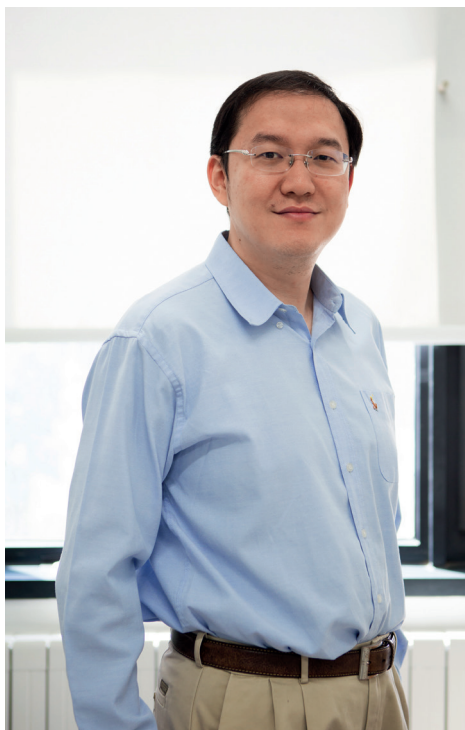


万事皆起于微，教学尤是

——访清华大学交叉信息研究院助理教授、
数据科学研究院 RONG 兼职教授徐葳

◆ 本刊记者 关悦

“世界上百年名校无非育人，天下第一等职业还是教书。”正值清华107年校庆，我们在清华园的蒙民伟科技楼采访到了交叉信息研究院助理教授徐葳，希望他与我们聊聊这几年来，他与数据科学的故事。说起来，徐葳与数据科学研究院也“相识已久”，从数据院筹办第一次会议，他就已加入其中。而大数据能力提升项目课程设置、教学目标的建立，他也一路见证、相伴参与。“清华每年能培养多少大数据人才？从数量上看，并不多。那我们对国家的贡献体现在哪里？如果我们的学生能把大数据思维带去他们的行业，星星之火，一点点影响和改变身边的同事，最后变革传统企业，那才算是我们的成功。”



清华大学交叉信息研究院助理教授、数据科学研究院 RONG 兼职教授徐葳

好雨知时节，当春乃发生。随风潜入夜，润物细无声。

在北京2018年的第一场春雨后，我们见到了徐葳。伴随着空气中还有些泥土与水汽的味道，听着徐葳老师娓娓道来的理想与实践，脑海里倏地浮现出“润物无声”的诗句。恰巧，他与我们聊起的第一个话题有关教学。教学这个词很大，如果认真说起来恐怕几天几夜都聊不完，但如果只讲数据科学这一个领域，徐葳用了简短的一句话概括：“在学习的过程中，不断解决问题，发现价值。”

缘何而来：新故相推，日生不滞

严谨、有条理，无论要阐述多少观点，理性的逻辑链条都时刻在线。和徐葳有过交流的人，大抵都会给出这样的评价。说起为何参与到数据院的建设，徐葳采用了一种倒推的方式开始讲述。“最初我参与了数据院的筹建会，是研究生院的老师叫我去的。而接触到研究生院，则是因为跟他们一起做在线教育。”理性的表述方式特别符合大家对他的评价。然而他的下一句话就变得感性起来：“但之所以参与得这么深，是因为这件事儿有意思！”

2014年9月，在研究生院的牵头下，大数据能力提升项

目正式招生,开设了《大数据系统基础》《数据分析》等课程。徐葳作为以上课程的授课教师,开始了一种新的尝试。“2014年的时候,数据科学这个概念在业界刚兴起不久,大家都在探索,数据科学的教学到底该如何开展?应该面向什么样的学生?基于这些问题,我们做了一些相关的调研。”

徐葳在当时调研了很多国外的院校,其中部分院校开设了“大数据专业”,也有部分院校把“大数据”作为人才培养项目来发展。而其中比较成熟的,则是在美国颇受关注的伯克利 MOT (Management of Technology, 即“技术管理”证书项目)。而这种模式也恰好与学校一直支持的学科融合、学科交叉有很强的相关性。清华大学大数据能力提升项目的雏形也就此开始显现。

“我觉得所谓跨学科的教学,其实就是一种语言课。”提及跨学科学习,徐葳颇有一些“执念”。这个执念,来自于对大数据真正作用的探寻:

“这主要分为两方面。一是教大家说这个学科的语言。比如我太太是学 MBA 做金融的,她做的金融模型我做不出来,但是她的思维方法和她用的描述语言我能看懂,知道它要干什么,这就是理解这门学科的语言;二是你能理解别人的需求,尤其是对于比你更不懂的这个领域的人。我觉得这才是重要之处。”而也恰恰因为这股执念,徐葳开始琢磨如何让这门“语言课”充分发挥价值。于是我们又聊回了开篇的那个话题:为什么数据学院正在做的事儿有意思。让学校里不同专业的学生一同上课,促进交流,相互学习。这种新型的教学方式徐葳觉得:“是个新鲜事儿!”

然而,由于不同背景的学生对于数据科学的理解程度也参差不齐,对于基础薄弱的同学来说,有些内容十分吃力。“同学们叫苦连天。”所以,课程最后还是分 A/B 班上, A 班为非信息类的学生,讲得相对基础一些; B 班是信息类的学生,讲得比较深入。而徐葳觉得这离“跨学科”的初衷有些距离,因为这失去了把不同专业的学生融合在一起的机会。后来他设计出了一个新的方法:上课分开上,但是课程项目 A/B 班一起做!这就使得不同学科背景的学生有机会进行深度的思想

交汇,非信息类的同学可以组织信息类同学做数据抓取、建模、做前端界面等工作;而数据抓取后的分析、发现及解决问题,则是非信息类的学生所擅长的。

“为了更好的结果,不仅要有取有舍,还要适时变化,方能长久。”徐葳说道。然而,大数据作为一门新兴学科,其人才培养之路却道阻且长。

从零到一:石以砥焉,化钝为利

2017年7月,徐葳辅导的第一位博士生毕业了,这件事让他感慨很深。为此,他还专门写了一篇题为《建立与学生的信任——写在第一个博士生毕业之际》的文章。文中提到:我从来不是天才儿童型的学生,幸运的是,我一路碰到了很多好老师,教育不仅教给了我计算机科学知识,更是改变了我整个人。我自己的经历让我相信,教育是可以(正面地)影响一个人的整个人生。

1999年初入清华之时,徐葳还是懵懵懂懂的本科生。曾在教室听课记笔记,也曾在西操场飞奔。可能当时的他并没有想过,会在十几年后,以另一种身份回到这里。在回到清华之前,徐葳曾在 Google 总部工作,负责基础架构的研发。以往的工作经验让他更加明白“实践”的重要性。因此当他来到讲台时,不仅成为了知识的传授者,也带着学生真刀真枪搞实践,变成了学生与企业之间的沟通者。

“我觉得大数据最重要的是应用。大数据的学习绝不是躲进小楼自成一统,在实践的过程中才能碰撞出灵感。”大数据能力提升项目中的《大数据实践课》由此进一步发展。

然而实践说起来容易,做起来难。每个企业所处的行业、细分领域都不一样,这就导致了有些企业能够理解“什么是大数据”,而有些企业对数据几乎没有概念。“这些企业确实拥有很多数据,但是说不清自己的需求到底是什么,或者提出的需求没有价值。”徐葳说道。

而另一方面,学生们也经常会出现“小状况”。为了让学生们进一步深入到企业,了解企业的真

实需求，徐葳要求学生们尽量跟不同类型的人沟通，建立开放的思维。然而有些学生面对企业“没有价值”的需求，会产生一些抵触心理。“有的学生跟我说，‘这个企业的人不懂大数据，不懂就没办法指导我’，或者‘企业没给我合适的的数据，没数据我就绝对做不出来。’我经常收到类似于这样的反馈。”说到这个，徐葳有些无奈。但是他随即给我们讲起了另一个故事：“我在上学的时候，是很畏惧跟人交流的。那时候我成绩不错，但是由于很少说话也很少参与集体活动，很多老师和同学都不认识我。当时我的导师给了我很多机会去锻炼，比如每年两次给业界人士演讲。我去参会之前他一遍一遍帮我修改讲稿，帮我排练。我当时还挺烦他的，觉得他逼我，然而，在我做了几次成功的演讲之后，我感觉我走出了这个怪圈，之后我再演讲慢慢没有那么紧张了。”

回想起自己导师当时的做法，徐葳向学生们表述了两个理念：一是相信自己能做到；二是适时改变、因地制宜。“为了解决这个问题，我就跟学生说，给企业提出你能做出来的需求，并向他们解释清楚为什么他们提的需求没有意义或者为什么你做不出来。对于我来说，我要求学生们阐述，通过有限的数据分析，你们能得出什么样的结论，能挖掘出什么样的价值。有时候不需要做到100%，70%甚至20-30%就足够产生价值了。”

数据院教学部主管罗国荣在大数据能力提升项目中与徐葳合作紧密。据她回忆：“碰到问题了，有时候徐葳老师会带着学生们亲自跟企业谈判。他有项目管理的经验背景，学术积累又很丰富，所以跟企业对接起来也更高效。探索的过程中，困难一定会有，但关键看如何解决它。”

从最初的10个小组，到如今的30多个小组，越来越多的学生参与到实践的环节中。回首这两年走过的足迹，徐葳说道：“不仅学生们通过实践收获了很多，我也学到了很多。”

路在前方：星星之火，可以燎原

2016年9月9日，徐葳获得了清华大学第

15届“良师益友”特别奖。在颁奖典礼上，他说了这样一句话：“我要对我的学生负责任，和他们一同进步成长。”作为2017年度图灵奖得主David A. Patterson教授的高足，交叉信息研究院助理院长，徐葳的学术能力自不待言。此外，他为人风趣幽默，一直备受学生们的敬重和爱戴。然而，与这些相比，更重要的是一种“情怀”。

“教学生们有用的东西，让他们做有价值的事情。”徐葳将教学过程中的价值一分为二，一方面对于学生们的价值：“学生们不断去适应新的教学方式，然后慢慢去消化，能够真正学到东西。”

而另一方面价值，则更偏向于结果。“如果生物材料专业的学生学了我们的项目，最后去BAT了，对于他个人来说可能是成功的，但那是我们的失败，因为那并不是我们的初衷。我们不是让大家改行去学计算机，我们希望看到的是，他去了他所在领域的企业从事数据相关的工作。如果回到原点，我们来探讨清华能培养多少人？其实没有办法看绝对数量，那么我们对国家的贡献在哪？就是把这些具有跨学科学习及应用能力的人撒到各个领域。他既有数据思维，又有行业背景，又是清华毕业的，他能够潜移默化地影响带动周围的人有数据思维。尤其是传统行业，逐步地给这个行业带来一些小变化，就是我们的成功。”

“星星之火，可以燎原。”

徐葳，清华大学交叉信息研究院助理教授，助理院长，清华大学数据科学研究院管理委员会委员、RONG兼职教授，清华大学金融科技研究院副院长。专注于交叉学科的分布式系统和机器学习方面的研究。美国加州大学伯克利分校计算机硕士、博士，师从2017年图灵奖获得者David Patterson教授。在宾夕法尼亚大学计算机获得学士学位（在清华计算机系本科学习两年）。2013年入选“青年千人计划”，曾获得谷歌、IBM的教授科研奖，获得清华大学“良师益友”特别奖，清华大学先进工作者等称号。