

大数据时代的“清华紫”

◆ 特约记者 王睿诗

景昃鸣禽集，水木湛清华。1911年4月29日，清华学堂于清华园建校开学，并于次年更名为清华学校。时隔百年余载，走过世纪风雨的清华，在新的时代焕发出蓬勃的生机。年初上映的影片《无问西东》，引起了社会对清华精神的追寻和思考。水木清华，钟灵毓秀，清华大学秉持“自强不息、厚德载物”的校训和“行胜于言”的校风培育了各界人才。如今，校庆大幕已启，来自世界各地的清华人将汇聚在这个园子里，共话当年，意气风发。此刻，撑起大数据旗帜的清华人也将在这里，共述未来，神采飞扬。现如今，作为更加创新、更加国际化的清华，在这个被互联网、大数据、量子计算、区块链各种技术被重新定义的时代，又孕育了哪些杰出人才？

大数据在金融方向的应用

百融金服 CEO 张韶峰

1996级清华大学电机系本科

张韶峰，清华大学电气工程自动化学士、硕士，清华大学五道口金融学院讲师。毕业最初，以管理培训生身份加入甲骨文大中国区数据挖掘与商业智能软件部门。之后加入全球最大的IT解决方案提供商IBM，负责商业智能与数据挖掘以及智能分析方案，并创建了天才博通公司，他打造了国内第一个SaaS模式的数据挖掘软件系统（Geni-sage Data Mining），时间序列预测效果远超欧洲领先水平。

2014年，张韶峰创建百融金服。作为最早涉足大数据金融信息服务领域的企业之一，百融金服凭借线上线下融合的多维度海量实时数据、服



务银行等金融机构的行业先入优势以及超强的数据处理和建模能力，为信贷行业用户提供包括营销获客、贷前信审、贷中管控以及贷后管理在内的客户全生命周期产品和服务；为保险行业用户提供精准营销、存量客户管理以及个性化产品定制等产品和服务。在技术方面，百融金服积累了深厚的大数据处理技术，基于Hadoop框架开发，具有高效、灵活的特点。云端大数据存储技术具备流式处理实时数据、密集处理活跃数据、存档备份数据的优势。

银联智策 联合创始人兼 COO 呼延如生

1988级清华大学水利系本科

呼延如生，银联智策联合创始人、首席运营官、董事，清华大学水利工程专业，获得美国密苏里大学计算机工程与计算机科学专业及统计学

专业博士学位。历任SAP研究院、第一资本银行、万事达顾问等知名跨国机构高级技术与管理职务，在技术开发、信息决策、金融科技、人工

智能以及大数据应用领域拥有丰富的实战经验。学术研究方面，在自然语言处理领域富有独创性的研究成果。同时，还担任清数大数据产业联盟副秘书长、清大智库创始理事等社会职务。

银联智策成立于 2012 年，作为在大数据和人工智能领域拥有领先技术的金融科技公司，银联智策自成立以来一直致力于以数据和技术驱动基于场景的服务创新与效率提升，业务板块涵盖了信贷、投资以及商业服务的各个行业，已经成

功服务了金融、非金机构、汽车、旅游等行业龙头客户。在技术创新上，银联智策与高校和研究机构深入合作，充分发挥高校在超级计算、人工智能、自然语言处理等专业领域的优势，把数据科学的各个学科组织成一个完整的体系，与银联智策的商业知识与行业经验实现有效融合，目前已经和清华大学、上海纽约大学等高校在经济金融领域的应用研究取得了丰硕成果，在实践教育与人才培养方面也达到了较好的效果。

大数据在图像识别方向的应用

旷视科技联合创始人兼 CEO 印奇

2006级清华姚期智实验班

印奇，清华姚期智实验班本科，从本科开始在 MSRA（微软亚洲研究院）实习加全职工作共四年时间，研发了微软当时核心的人脸识别系统，被广泛应用在 X-box 和 Bing 等微软产品中。2011 年 10 月，他与两位清华同学（唐文斌、杨沐）共同创办旷视科技。为弥补欠缺的技术，2013 年印奇赴美国哥伦比亚大学攻读 3D 相机方向博士学位，两年后中断博士学业，回国创业。

旷视科技旗下第一款产品 Face++ 现已成长



为世界最大的人脸识别技术平台，人脸识别率高达 97.27%，并以云服务方式将技术提供给广大开发者和企业级伙伴；同时，在金融、安防、零售领域分别开始了商业化探索，成功发育出 Face++Financial, Face++Security, Face++BI 等垂直人脸验证解决方案产品。如今，Face++ 平台涉及了 1 亿人的刷脸，图库数量超越 10 亿张，总调用量超过 60 亿次，合作的 APP 有 1.5 万，活跃的开发者的超过 2.4 万人。

大数据在语音识别方向的应用

搜狗 CEO 王小川

1996级清华大学计算机系本科

2003 年王小川从清华大学毕业后，正式加入搜狐公司任高级技术经理，此后陆续晋升为技术总监、高级技术总监。2005 年他晋升为搜狐公司副总裁。2008 年，因搜狗输入法的成功被任命为高级副总裁。2009 年因在推动搜狐技术驱动文化中起到关键作用被任命为搜狐公司 CTO。2010 年，搜狗从搜狐中分拆出来单独运营，王小川兼任搜狗公司 CEO。

王小川是人工智能的坚定支持者，而搜狗

公司也早已在人工智能领域，以“语言”为核心进行布局。在语音方面，搜狗已经有语音识别、语音合成、声纹识别等能力；其中，依赖于搜狗输入法，用户日均语音输入调用次数峰值到 3 亿次，堪称中国最大的语音识别引擎，识别率高达 97% 以上。与语音相比，搜狗在图像方面的能力主要体现在此前已经在世界互联网大会上亮相的唇语识别，它有效地反映了语音与图像的结合。

美国微软研究院孵化研究开发部总经理 黄学东

1982级清华大学计算机硕士研究生

黄学东现为美国微软研究院孵化研究开发部总经理，IEEE 院士，湖南省信息化领导小组顾问，湖南大学客座教授，湖南大学软件学院名誉院长。

1993年，他作为高级研究员加入了微软研究所。2000年到2004年，他担任 Microsoft's 语音识别技术研究团队的总经理，负责微软语音技术和产品的研究、开发和市场拓展。他领导微软语音技术的开发和产品的市场营销，包括获奖的 Speech Server 2004。此前，他在卡内基·梅隆

大学致力于发展 CMU's Sphinx-II 语音识别系统。发表一百多篇论文，并合著出版了两本非常有影响的专著：Hidden Markov Models for Speech Recognition (Edinburgh University Press 1990)（语音识别中的隐含马尔可夫模型（爱丁堡大学出版社 1990））和 Spoken Language Processing (Prentice Hall 2001)（口语处理（普林斯顿出版社 2001））。曾获中国教育部的科技进步奖、IEEE 信息论文奖、Allen Newell 杰出研究奖章，以及 SpeechTek 的语言领域十大领军人物称号。

大数据在自动驾驶方向的应用

Pony.ai 创始人兼 CTO 楼天城

2004级清华姚期智实验班

在清华大学学习期间，楼天城参加了国内外一系列编程竞赛，取得骄人成绩，留下“常以一人之力，对抗一个编程团队”的传奇，被公认为“大学生编程第一人”。

2011年毕业后，楼天城前往美国谷歌就职。在谷歌，做过社交网络项目，主攻无人车算法。2016年，楼天城加入百度（美国），参与无人驾驶项目研制，成为百度历史上最年轻的 T10 技术

高管。2017年初，他从百度辞职，携手前百度无人车首席架构师彭军创办了 Pony.ai。2017年6月，Pony.ai 获得加州路测资质，7月，其无人驾驶汽车就跑上了加州街头。2018年1月，Pony.ai 完成 1.12 亿美元 A 轮融资，创下国内无人驾驶领域最大单笔融资记录。2018年2月2日，Pony.ai 的第一支城区运营白天黑夜全场景无人驾驶车队正式在广州上路。

大数据在 O2O 平台方向的应用

美团创始人兼 CEO 王兴

1997级清华大学电子工程系本科

2003年，王兴中断了在美国特拉华大学电子与计算机工程系的博士学业，回国创立校内网；2007年创办饭否网；2010年创办团购网站美团网；2014年，王兴入围 2014 年度华人经济领袖。

美团是全国最大的生活服务电商平台，拥有目前世界上规模最大、复杂程度最高的人多、多点的实时配送智能调度系统，能够承担每日多达



1800 万次订单。

随着需求不断升级，美团在 2016 年起开始探索机器人补充配送。目前，美团无人配送团队在无人车和无人机领域已申请了超过 60 项技术专利。

同时，美团还为外卖骑手研发了蓝牙智能耳机，使用语音交互对话，解决了骑手在配送过程当中操作手机困难的问题，系统根据骑手骑行状

态自动唤起交通安全提示，减少安全隐患，保障了骑手安全。语音助手已经覆盖 24 万骑手，通过语音接单，80% 的订单可以在 20 秒内完成。

瓜子二手车 CTO 张小沛

1991 级清华大学计算机系本科

张小沛从清华毕业后赴美深造，随后进入微软、HULU 等美国科技公司担任重要职位。2013 年从零组建宜信大数据团队内部创业，并出任宜信 CTO、大数据创新中心总经理。2016 年 3 月底正式加入瓜子，担任公司 CTO。

张小沛带领团队为瓜子建立了基于多种算法模型的智能大数据体系，其中包括 200 多万车辆，1 亿多车主和潜在买家的基因图谱库，还建立了通用的智能平台“瓜子大脑”。该平台会支撑瓜子各条业务线所有需要数据驱动、智能驱动的环节。通过瓜子大脑，瓜子搜集了用户在瓜子



的行为数据，把这些数据同瓜子从第三方合作得来的数据做关联和理解，然后放到基因图谱中，以此来驱动瓜子平台上 C2C 的主场景，金融，保险等附加场景。通过大数据和人工智能深度改造二手车行业的效率，瓜子二手车车辆销售周期由 32 天缩短至 46 天；车源上架成交转化率提高了 30%，买车用户页面转化率提升了 35%。

大数据在传媒领域的应用

大数据文摘 CEO 汪德诚

2011 级清华 MBA

汪德诚，清华大数据产业联盟副秘书长、中关村大数据产业联盟副秘书长、大数据文摘创始人。清华大学 MBA、北京邮电大学通信工程学士，



BIG DATA DIGEST
大数据文摘

之前一直从事电信行业。2013 年 7 月创办大数据文摘，以传播数据文化、推广数据思维为己任，坚持发布高价值文章，已是业内影响力最大的行业媒体。2016 年成立北京惟数科技有限公司，专注数据科学、人工智能，落脚媒体、培训、人才服务。旗下媒体品牌大数据文摘从 2013 年 7 月成立至今，每日坚持发布优质内容从未间断，已有读者超过 50 万，是最有影响力的数据科学、人工智能媒体。培训业务 2016 年 8 月启动，已开设课程覆盖机器学习理论、基础、实践各环节。除此之外，还与网易云课堂独家合作大数据文摘和北邮模式识别实验室联合译制《斯坦福深度学习与计算机视觉》，目前已超过 2200 人报名此课。

数据服务方面的应用

国双科技创始人 祁国晟

2002级清华大学软件学院本科

祁国晟，国双董事长兼 CEO，于 2005 年创立国双，并担任董事长兼 CEO 至今。2016 年 9 月 23 日，祁国晟带领国双成功登陆纳斯达克，成为纳斯达克上市的中国企业中最年轻的 CEO，同年他被清华大学苏世民书院聘为导师，被北京大学授予杰出教育贡献奖。祁国晟于 2016 年、2017 年两度入选《财富》“中国 40 位 40 岁以下商界精英”。

2012 年 12 月，祁国晟荣获由北京中关村高新技术企业协会颁发的“中关村创业之星”奖，后又获选“中关村创业之星”导师。祁国晟 2006 年毕业于清华大学，主修计算机软件。在校期间创立了清华大学学生软件产业研究协会，并曾在微软亚洲研究院实习，参与下一代搜索引擎研发工作。

国双科技自 2005 年公司成立以来，专注于



大数据分析 & 语言处理，在大数据、人工智能领域颇有成果。基于国双大数据平台独有的分布式数据架构和先进的实时、多维度关联性分析技术，国双的解决方案能够使客户充分洞悉数据间的复杂关系，获得全新的商业洞察，以利企业和政府客户作出更好的业务决策。

在技术开发方面，国双科技多与高校展开合作：2013 年，国双数据中心成立；2014 年，成立人民大学—国双大数据科学联合实验室；2015 年，成立哈工大—国双联合自然语言处理实验室；2016 年，国双科技登陆纳斯达克，成为中国首家赴美上市的大数据企业。

博彦科技股份有限公司 CEO 孙铁林

1997 级清华大学经管本科

2007 级清华大学 MBA

博彦科技（深圳）有限公司 CEO、董事孙铁林毕业于清华大学经管学院，同年加入博彦科技团队，负责项目管理及产品研发工作，历任项目执行总监、互联网事业部总经理、集团副总裁。2009 年主持成立博彦科技 -RFID 事业部，且该事业部成为博彦科技垂直业务领域成长速度最快的部门。

大数据是博彦科技近几年转型发展的重要方向之一，博彦科技于 2016 年成立全资子公司——博彦多彩数据科技有限公司，并请孙铁林担任总经理。博彦科技是一家面向全球的 IT 咨询、产品与解决方案服务提供商。公司依托强大的研发与创新



能力，广泛采用基于大数据、云计算和移动互联等新兴技术，提供了包括大数据管理平台、企业风险预警及舆情监测系统、基于 AI 的智能运维服务、智汇云集成平台、智能的自动化测试产品以及国际权威的软件测试资质认证培训在内的多种产品及解决方案，广泛服务于高科技、互联网、金融、医药、汽车、零售、物流、能源、制造、医疗等领域。

北京诺亚星云科技有限责任公司 董事长 汪浩

1993级清华大学电子系本科

1993年，汪浩就读于清华大学电子工程专业攻读本科，2000年硕士毕业，后赴美国赫尔辛基理工大学攻读博士学位。读博期间，同时担任诺基亚中国研究院主任研究员、研究经理；2011年至2013年任联想研究院主任研究员、总监。此后任北京诺亚星云科技有限责任公司执行董事、总经理。

汪浩2012年创立了北京诺亚星云科技有限责任公司。公司拥有优秀的技术团队和国际领先的实

时大数据分析处理技术，在云计算、流处理、智能推荐、移动计算与服务、系统性能分析与优化、服务系统架构等方面具有深厚的学术造诣和多年的技术积累。诺亚星云立足于开发自主知识产权的核心云服务平台技术与实时大数据分析处理技术，打造海量实时数据处理产品，构建衍生的移动互联网服务与移动终端解决方案，并为客户量身打造极速精准的大规模实时数据处理应用平台。

网帅科技（北京）有限公司总经理 胡庆勇

1987级清华大学自动化系本科

网帅科技（北京）有限公司成立于2015年4月，是一家全流程数据科学解决方案公司，通过自主知识产权的算法，知识库和模型生成工具，融合运营商大数据和跨界行业多维数据，标识、分析、挖掘数据中的行业价值信息，使融合后的数据成为服务各行各业的全新数据资产和产品。

总经理胡庆勇是数据科学研究院交通大数据研究中心副主任。中心致力于异构多源交通数据

的采集汇聚和清洗融合技术，多类型交通数据的自动特征提取和结构化转换技术及相关标准体系等。其创新在于利用交通仿真，大数据分析，机器学习和深度学习技术，融合多达近六十类的跨界多类型、多源头大数据，实现交通态势的感知和交通风险的感知，将交通安全管理从被动的事后应对处置转变为交通风险的主动的事前预警处理、事中快速疏导，从而降低交通事故死亡率，优化交通管理和治理。

大数据在工业互联网方向的应用

昆仑数据创始人兼 CEO 陆薇

1996级清华大学计算机系博士研究生

陆薇有19年IT技术前瞻研发及企业服务经验。曾作为IBM物联网重大研发专项（Big Bet）全球技术负责人，团队曾获IBM总裁奖及杰出技术成就奖，成为IBM在中国本土成长出的10个技术高管（Distinguish Engineer）之一，入选IBM技术科学院（Academy of Technology）成员。清华大学计算机系博士，现兼任清华大学工业大数据中心副主任。

陆薇于2014年底创立了昆仑数据，深耕工业大数据领域，为企业提质、增效、降耗和控显。昆仑数据自主研发了偏向底层的KMX机器大数据管理分析平台，并基于前者研发了K2Assets设



备全生命周期管理解决方案，面向装备制造、能源、电子等领域，提升企业的智能化管理水平。KMX是涵盖工业大数据从接入管理、深度分析到

落地应用全流程的一站式、一体化数据平台。目前已支持累计超过200000亿数据点的海量数据、高达8千赫兹频率的高频数据、以及每秒千万数据点实时数据的接入、存储和查询。目前，

K2Assets已服务于全球最大的风电整机制造商金风科技、国内领先的动力装备解决方案提供商陕鼓动力、国家工程机械行业大型骨干企业山东临工，在高频、海量的工业场景下，实现大数据预警。

大数据在 AI 芯片方向的应用

深鉴科技 CEO 姚颂

2011级清华大学电子系本科

2016年，姚颂与汪玉、单羿及韩松三位学长一同创立深鉴科技，并担任CEO。带领团队自主研发高效深度学习平台，致力于为智能安防与数据中心等行业提供算法、软件、芯片一体的人工智能方案。深鉴科技在创立2年的时间里，连续获得了金沙江创投领投的500万美元天使轮融资与Xilinx、联发科、清华控股领投的2000万美元A轮融资，2017年10月完成由蚂蚁金服与三星风投领投，招商局创投与华创资本跟投的4000万美金的A+轮融资。作为一名年轻优秀的创业者，姚颂曾获得福布斯亚洲Under 30、MIT TR35 China（35岁以下科技创新青年）等荣誉。

而深鉴科技也发展为优秀的硬件加速方案和深度学习平台提供者，目前已具备世界顶尖的神经网络压缩编译技术、体系结构与指令集结构设计方法以及FPGA/芯片系统设计集成的研发



能力，形成了以自主研发的深度压缩技术与深度学习处理器（DPU，Deep-Learning Processor Unit）为核心，从算法到软件到硬件的完整解决方案，让所有亟需解决高算力、低功耗需求的智能化设备提供商能够方便使用。

大数据在数据营销方向的应用

北京清数科技发展有限公司 技术总监 黎成权

2001级清华大学工程物理系本科

清数D-LAB成立于2015年底，位于北京市海淀区清华园的清华大学学研综合楼，是清华大学数据科学研究院、清数大数据产业联盟和启迪创业孵化器大力支持下建立的大数据创新创业平台。

清数D-LAB顺应“实施国家大数据战略”

和“大众创新、万众创业”的号召，汇聚了国内大数据领域最优秀的产学研用资源，肩负着引领数据创新，培育数据人才，促进数据开放，打造数据生态的使命。清数D-Lab建立大数据知识积累和人才培养体系，促进业界联合和数据联通，加快大数据与产业的融合从而推动产业创新。