

# 数据科学家迎来好日子

## 软件编程、算法开发和其他技能促进科学发展

正如这个数据驱动时代的其他很多科学家一样,Karthik Ram在2009年彻底改变了自己。当他开始进行博士后研究——气候变化如何影响美国怀俄明州黄石国家公园的麋鹿时,他认为自己是一名生态学家。但解释卫星数据以及分析跟踪动物的项圈等工作,让他开始扩展自己的思维定式。

为了了解不断变化的生态系统,他需要不断打磨编程,了解如何管理如山岳般的信息,这些技能已经改变了他的定位自我以及个人职业的方式。“‘生态学家’这个词我用得越来越少。”他说,“现在,我主要称自己是一名数据科学家。”

### 年轻学科

数据科学在2009年是一门年轻的学科,但却快速变得成熟,现在已经与很多学科存在交叉。尽管其定义在变化,但数据科学通常涉及利用计算工具管理和解释大数据集。

现在加州大学伯克利分校伯克利数据科学研究所工作的Ram正在与同样转入数据领域的原神经学家、社会学家和生物学家一道工作。“研究所的每个人都和我一样。”他说,“我们拥有计算技能和统计技能,我们可以将这些用于特定的领域。”

对数据科学家的需求已经从学术界扩展到工业、保健、政府和会产生复杂信息的所有机构。微软公司预测,到2020年美国数据科学和分析领域可能会拥有超过270万个工作岗位,在2015年的基础上增加15%。根据欧洲数据科学学院(一家在欧洲搜寻数据科学技能的识别和收集工作岗位的培训和教育组织)统计,欧洲的相关人才需求量与此类似。2015年以来,学术界已经发布了300多万万个类似的招聘广告,其中包括今年3个月内发布的29万个岗位招聘广告。

对于那些寻求数据科学家角色的人来说,其挑战并不是找工作,而是找到与他们的能力和兴趣相匹配的最佳职位。识别“最正确的岗位非常棘手”,科泉市科罗拉多学院原终身教职数据科学家、现在加州埃默里维尔 Zymergen 公司(研究基因编辑微生物的新用途)做数据科学家的Amelia Taylor说,“数据科学在不同地方看起来会非常不同。有如此多的公司,很难知道要看哪一些。”

当其他领域的博士学位拥有者拥有的选择少之又少时,这个领域拥有“太多选择”是一个好现象。掌握正确的技能、了解自身机遇所在的科学家将会看到一个有回报的、数据驱动的未来。

### 大量角色

数据科学已经导致很多领域“水涨船高”。除了“数据科学家”搜索增多之外,“数据工程师”和“数据分析师”也是工作搜索平台上的热词。这些角色的差异有些微妙但却很重要。“数据工程师的核心技能是打造稳健的不会失败的系统。”英国伦敦 ASI 数据科学企业(提供咨询服务和数据科学奖金计划)首席执行官 Marc Warner 解释说。

“数据科学的定义和期待可能会随着时间而转变,但这一领域将会沉淀下来。”

英国伦敦阿兰·图灵研究所是日益增长的数据科学领域的一个跨学科港湾。

图片来源:Toby Keane



他表示,数据科学家和分析师的一个重要差别是,科学家会跟踪他们所引导的数据,即“数据第一”的方式;而分析师则通常会使用数据验证一个已成立的假设。

在英国伦敦阿兰·图灵研究所(ATI),Mihaela van der Schaar正在让数据开路。她开发了有助于患者个性化治疗、疾病诊断和风险预测的计算机算式。“我认为这样的技术能够改变医学,挽救生命,促进科学突破。”她说。

ATI由英国5所大学与国家工程和物理科学研究理事会于2015年成立,旨在促进产业界和政府的合作,该机构体现了数据科学的跨学科精神,van der Schaar说。她补充说,数据科学领域一些最大、最有趣的问题来自未曾预料到的地方。

“现在,我在ATI参与最多的项目之一是开发更好的方法,了解和治疗囊性纤维变性患者。”她说,“这并非来自于产业界,也并非来自于政府,而是通过与英国囊性纤维化病基金会的合作关系。”

跨学科联系还形成了摩尔·斯隆数据科学环境基金会——已经在加州大学伯克利分校、西雅图华盛顿大学和纽约大学建立了若干数据科学中心的一项倡议。每个中心都将来自各个领域的数据科学家聚集到一个地方。其想法是通过在那些推进数据科学方法的人(数学、统计学和计算科学领域的研究人员)和将其应用于社会、物理和生命科学领域的人之间搭建桥梁,

从而加速促进发现。”

并非所有博士生项目都在为进入数据科学的现实世界准备研究者,因此,短期培训课程正在变得日益受欢迎。Taylor之所以能够进入数据科学大门,是因为她获得了加州帕洛阿尔托一家名为“洞察数据科学”机构7周的奖学金项目,该项目将数据科学家和美国公司连接在一起。其成员已经在亚马逊、脸谱网、JP 摩根以及各种其他大小型技术公司获得职位。

Taylor说,“洞察”的奖学金项目为教导她当前工作所需要的技能起到了非常宝贵的作用。除了掌握熟练的技能之外,培训教会她跳出数据分析的框架思考完成产品的实际应用。她观察到,拥有博士学位的人在进入产业界的工作岗位时通常会碰到过渡期的困境,除非他们已经掌握了一手的产业界经验。“因为我考虑到产品,所以在公司内起步非常快。”她说。

### 医疗帮助

数据科学已经到达医学和医疗中心,为很多研究型科学家提供了使用其技能的另一个出路。作为纽约大学和附近罗切斯特大学的神经科学培训的一部分,“洞察”原参与者 Anasuya Das 学习了 C++ 编程语言,以创建帮助人们从中风中恢复过来的软件,使患者可以利用自己的计算机练习视觉。Das 还参加了两三门计算神

经科学课程,这激发了她将数据科学作为全职工作的兴趣。现在,她正在纽约市纪念斯隆凯特林癌症中心工作。

Das 在利用一个系统将患者和临床试验相匹配。“我每日的工作非常多变,包括从做纯软件工程到与我们正在为其打造产品的医生见面。”她说。

Lazowska 预测,数据科学的崛起最终将会改变“不出版就淘汰”的科学系统。他认为,有一天,编码和数据集将会像目前的出版一样,成为职业发展的先决条件。现在,他表示,他和同事正在鼓励研究人员在简历中列举数据科学相关成就。他们还建议晋升和终身职位委员会将这些技能列为有效指标加以考量。

Ram 还在简历上列出了广泛的数据科学项目。目前,他正在参与一项长期工作,衡量人类活动对 Tahiti 生态的影响。问题已经比他在黄石公园的时候复杂了很多,但工具也变得更加先进。现在,他不需要在一个数据集上埋头苦干数月,而是在数小时内就可以获得结果,这并不是说数据科学已经变得容易了。

“对于很多人来说,一个大挑战是拥有学习这些新工具以及如何正确应用它们的能力。”Ram 说,“随着时间的推移,几乎所有的首席研究员的团队里都会需要一名掌握这些特殊技能的人才。”数据科学的定义和期待可能会随着时间而转变,但这一领域将会沉淀下来。(晋楠编译)

## 科学线人

全球科技政策新闻与解析

## 英国议会科学委员会有了女议员



位于英国伦敦的议会大厦和大本钟。图片来源:Daniel Coomber/Flickr

近日,英国下议院科学和技术委员会在推特网上表示,“非常高兴地宣布我们的组成人员已经确认”。经过最近的选举之后,议会委员会都被解散,现在该委员会推出了新的议员团队。但唯一的问题是:8个几乎都在微笑的面庞均是男性。

这使得该国议会科学委员会的唯一专家组没有任何女性。在此前的议会中,科学委员会成员60%是女性,包括其首席。

很多观察人士对此感到灰心。“我的心沉了下去。”伦敦科学与工程运动组织主任 Sarah Main 在博客中写道。剑桥大学物理学家 Athene Donald 在《卫报》表示:“年轻女性科学家看到这会感到丧气,议会对她们的未来明显不关心,他们甚至没有想到要有一位象征性的女性。”

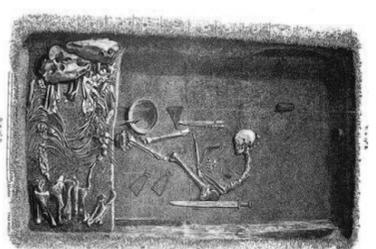
英国大学和科学部长 Jo Johnson 呼吁重新解决这一问题,从而让科学委员会具有“基本合理性”。

该委员会主席、自由民主党的 Norman Lamb 很快便着手解决这一问题。“我认为委员会成员中拥有女性非常必要。”他在推特网上解释说是英国政党负责选举委员会成员。Lamb 近日给“党鞭”写信,敦促他们在剩余3个席位的提名者中纳入女性。

被工党选拔担任科学委员会成员的 Darren Jones 有点“尴尬”地告诉《独立报》:“我通常会拒绝加入全是男性的专家组。”他解释说,工党中没有女性自愿加入科学委员会。“最主要的挑战是大多数在这一领域有经验的工党女性议员已经在忙着做其他的事情。”

不久前,Vicky Ford 宣布她“很高兴入选”。Ford 是来自英国切姆斯福德市的一名保守党议员,此前曾是欧洲议会的一名成员,她曾参与制订欧盟研究资助项目“地平线2020”计划。另一个位置花落格雷夫舍姆的保守党议员 Adam Holloway,他拥有电视新闻方面的背景。(晋楠)

## 北欧女战士身份遭质疑



这张1889年的北欧海盗墓穴遗址绘图表明,她曾与武器、盔甲和战马一起下葬。图片来源:C. Hedenstierna-Jonson

日前,考古学家报告称,一名在19世纪80年代首次被发现后长期以来被认为是男性的北欧海盗,实际上是一名女性。她与战剑、斧头、长矛和两面盾牌被一起下葬,成为目前已知的第一位高级北欧海盗女战士。

然而,一些北欧海盗研究者对于这名女性是否曾是一位身经百战的女战士却存有疑问,或者她可能仅仅是穿着战士的装备被埋葬。《科学》杂志采访了这位研究首席作者、瑞典乌普萨拉大学考古学家 Charlotte Hedenstierna-Jonson,以了解考古学家能够从这名受到质疑的北欧女海盗身上推断出什么,以及女战士遗体否定历史思维定式时所产生的双重标准。

关于这名女性究竟是一名战士,还仅仅是穿着战士的衣服,Hedenstierna-Jonson 表示:“这是关键的问题。一个坟墓代表着什么呢?追问这个下葬者是否曾参加战争是个相关的问题,这里从19世纪80年代以来就被认为是一个战士的坟墓,而且以前也没有人质疑过。”他补充说,“在知道骨骼来自于一位女性之前没有人曾做过这样的评论。考古学并未改变,唯一改变的事情是我们认为那是一名女性而非男性。”

而墓穴内也没有其他的女性特征。Hedenstierna-Jonson 表示,她并未穿戴裙袍,而这对于一名埋葬的女性来说比较常见。所以,一个解释是她活着的时候就拥有战士的身份,而另一种解释是,她象征着并不曾拥有过的角色,那将会比较有趣,也很独特。“换言之,她可能曾经并不是战士,但她却得到了类似的荣誉,所以被当作战士下葬。”

而 Hedenstierna-Jonson 表示,她并未穿戴裙袍,而这对于一名埋葬的女性来说比较常见。所以,一个解释是她活着的时候就拥有战士的身份,而另一种解释是,她象征着并不曾拥有过的角色,那将会比较有趣,也很独特。“换言之,她可能曾经并不是战士,但她却得到了类似的荣誉,所以被当作战士下葬。”

而 Hedenstierna-Jonson 表示,她并未穿戴裙袍,而这对于一名埋葬的女性来说比较常见。所以,一个解释是她活着的时候就拥有战士的身份,而另一种解释是,她象征着并不曾拥有过的角色,那将会比较有趣,也很独特。“换言之,她可能曾经并不是战士,但她却得到了类似的荣誉,所以被当作战士下葬。”

而 Hedenstierna-Jonson 表示,她并未穿戴裙袍,而这对于一名埋葬的女性来说比较常见。所以,一个解释是她活着的时候就拥有战士的身份,而另一种解释是,她象征着并不曾拥有过的角色,那将会比较有趣,也很独特。“换言之,她可能曾经并不是战士,但她却得到了类似的荣誉,所以被当作战士下葬。”

8月29日上午10:24,Jon Cohen 手持冲浪板冲进美国加州太平洋海岸。这次,Cohen 进行了这个月里最成功的一次冲浪。其间的水温约为20.33°C至21.38°C。

而Cohen 知道的数据之所以能精确到小数点之后,多亏了冲浪时与其合作的两位科学家:斯克里普斯海洋研究所工程师 Phil Bresnahan 和海岸生物地球化学家 Tyler Cyronak。

在纽约一个名为“失踪鸟”的非营利组织的支持下,Bresnahan 和 Cyronak 开发了一个冲浪板鳍,包含温度传感器、GPS 设备、微控制器电路板、蓝牙芯片和一个可充电电池。最终,他们还计划为其添加上 pH 值、叶绿素、盐度和氧气传感器。这些技术被打包在一个13厘米高的“智能鳍”中,而这个鲨鱼鳍样的设备被安装在Cohen 的冲浪板上。

不过,科学家的目标不是帮助冲浪者监控他们的冲浪活动,而是收集研究沿海地区的数据。他们希望能够广泛分布这些鳍,以便追踪海洋生物富集的珊瑚礁和海藻类的健康状况,或者监测珊瑚漂白,以及弄清随着时间推移,海洋吸收的来自全球变暖的热量。

通过与冲浪爱好者基金会的合作,Bresnahan 和 Cyronak 在过去3个月里为冲浪爱好者提供了50个“智能鳍”,于是,数据开始涌入。Cyronak 指出,加州海岸有一些防洪堤上安装了温度测量仪,但它们十分少且距离非常远。他说:“科学家想要的是沿海数据,而海岸很难监测。要真正了解沿海地区的情况,你需要进行大量的测量。”

英国普利茅斯海洋实验室卫星海洋学家 Robert Brewin 认为,“利用冲浪者作为平台改善近岸环境取样有巨大的潜力”。Brewin 专门研究浮游植物,并在普利茅斯的寒冷水域上冲浪,他



冲浪板“鳍”能用来收集海洋数据。

图片来源:Kat Hammond

把温度传感器夹在了冲浪者的脚踝上。他还为冲浪者佩戴了一个装着GPS装置的腰包。迄今为止,他的冲浪团队已经进行了400多次观察。

正如该团队最近发表在《科学公共图书馆一综合》和《河口、海岸和大陆架科学》杂志的论文所描述的那样,冲浪者在海上收集到的海面温度显示,近岸海域卫星测量数据并不可靠。

卫星的温度数据由先进型甚高分辨率辐射计(AVHRR)测得,与浮标的测量数据匹配。通常,浮标装备有精密的传感器,位于距离海岸约7到33公里的海域。但AVHRR获得的近海海域

温度并不可靠。“在这一区域,我们不能相信卫星或模型的数据,因为它们尚未被评估。”Brewin 说。

Brewin 将“智能鳍”视为其领域的“高端定制工具”,极大改善了其臃肿且缺乏精确性的系统。“我们使用的这个工具包有些复杂,不太适合公民广泛的科学使用。”目前在测试5个“智能鳍”的Brewin 提到,“这片鳍确实有吸引市民”的潜力。他推算,英国有足够的冲浪者每年进行4000万次的海洋测量。

Andy Stern 是纽约罗切斯特市的一名退休神经学家,他经营着“失去的鸟”项目,并在4年

前构思了“智能鳍”。他表示,该“鳍”的经费主要来自捐助者,并不打算在市场上售卖。而且,目前还没有计划大规模推广这种“鳍”。“成功并不意味着投放了多少‘鳍’。”Stern 说。

目前,制造一个“智能鳍”和电池的费用是300美元。“它将永远不会在冲浪商店里出售。”Stern 说。

Stern 希望,“智能鳍”正在寻求国家科学基金会的资助,他希望该设备能填补沿海研究的“数据空白”,并激发新微型传感器的开发。他补充说,“智能鳍”是一个象征,可能有助于激发和授权当地社区参与气候变化和科学政策相关工作”。

“迷失的鸟”项目将大海雀、拉布拉多鸭、卡罗来纳长尾小鹦鹉、新英格兰黑琴鸡和候鸽的雕塑,放在了它们灭绝前最后一次在北美出现的地点。这些“鸟类纪念碑”希望能让人们意识到自己对环境的影响,或者正如“迷失的鸟”的座右铭所言,“通过艺术让人们与地球更紧密地联系在一起。”

而 Bresnahan 提到,“智能鳍”给更广泛的受众传递了信息,而不仅仅是冲浪者。“我回家过圣诞节,告诉人们我正在为海洋健康做一个冲浪项目时,每个人都围着我转,想要了解更多信息。”

Cohen 知道“智能鳍”可能无法跟踪其冲浪活动,但他还是询问 Cyronak 和 Bresnahan:“其间GPS是否记录了我越过多少海浪?”Bresnahan 发电子邮件说:“我记录的数值是21个,但我们没有很好的方法来确定。”

最后,对于“智能鳍”的价值所在,Bresnahan 说:“人们总是问数据将会被用来做什么。我们认为需要新的传感器和平台,但可能直到推出一段时间后,我们甚至都不知道如何回答这些问题。”(张章编译)

# 用“鳍”感受大海

## 科学家尝试通过冲浪者收集海洋环境信息