



关于大数据,除去技术层面的研究和应用,实际上在与历史、社会、人文等学科领域交叉的过程中,也能引发出很多有趣有益的话题——

大数据可能沦为“大忽悠”

编者按

近日,腾讯网举办了一场主题为“大数据”的“夏季思享会”。此次分享会请到了来自各个领域的学者和专家,堪称是一场跨学科的思想碰撞。由于是跨学科,因此让人耳目一新的观点不少,其中有些还是目前国内关注大数据的人们不太了解的观点,既思辨又妙趣,可谓是对大数据的一次反思,打开了它的另一扇窗,让人看到了大数据的别样“风景”。

在一般人看来枯燥乏味的大数据,究竟深藏着怎样有趣的内涵和观点?我们不妨随着这些大方之家的思路行进,从多维的角度来了解。

□储棕荷

众所周知,大数据就是海量数据。目前,海量数据的应用正在改变我们现有的生产方式,不仅成为一个重要的生产要素,还在一定程度上提高了生产效率。这时,海量数据应用带来的信息风暴不仅影响了我们的工作、生活,还可以逐渐影响我们的思维方式。

有人说,大数据本身就是一场革命,这场革命就是把“量化”这个概念推进到我们生活中的方方面面,角角落落,几乎无人可以幸免。

大数据的挖掘过程需要注入思想,这种思想体现为“人性”

在北大新闻与传播学院的刘德寰教授看来,即使是像BAT这样的互联网巨头,它所

拥有的数据也只是一座“数据孤岛”。因此,很多大数据其实是断裂且封闭的。在这座座孤岛被打破之前,大数据都无法称“全”。这种大而不全的数据,可能导致我们无法真正获得想要的规律,也可能沦为“大忽悠”。

虽然关于大数据的介绍和分析都告诉我们,大数据注重的是“相关关系”,而非“因果关系”,但是刘德寰指出,如果只注重相关关系,数据量的增加将不可避免地带来伪相关的现象。比如,当数据量达到几十万之巨之后,只会导向“万物相联系”这一结论。

这时,所选取的变量都会呈现为统计显著,而这样的数据分析无疑就是失败的。因此,即使大数据现在备受推崇,但传统的抽样调查和实验等社会学研究方法也不该被抛弃,而应结合使用。

《大数据时代》的译者周涛也表明自己的观点:从技术层面上来看,我们的确需要找到因果关系。一方面,有的关联根本没有因果在背后是很难被发现的;另一方面,因果能够帮助我们更好地解释和分析相关。

刘德寰强调,大数据的挖掘过程需要注入思想,这种思想体现为“人性”。因为,统计数据不能代替对人性的理解,真正进行跟人相关的大数据挖掘的时候一定要关注人性,这个时候的算法和建模才有针对性,才有意义。

怎样在数据中注入人性?刘德寰指出,观察、实验、调查等传统的研究方法都是很好的选择。比如,乔布斯就经常躲在苹果体验店外人看不见的地方,观察体验店里的人的行为——这是最早期的研究方法,在拥有大数据的今天还在使用,是有一定原因的。说白了,诸如“体会、体验、直觉、灵机一动、内省”这些看似与大数据无关的东西,有

可能恰恰就是大数据的核心。

大数据也可以小应用,即把大数据应用到各个独立领域中去。怎样做好大数据的小应用?刘德寰认为要从两个方面着手:一是大数据与小数据结合,这个小数据就是基于抽样调查的数据;二是来自各个独立领域的专业知识非常重要,比如以心理学、法学、社会学、营销学等混合知识作为背景,再去跟数据结合,就能够探索出一个更科学的大数据分析和挖掘的方式。

基因研究是一项很少被人了解的产业研究,但却是真正的大数据应用

怎样有效地预测疾病,降低传染病的发病率,这些都需要依赖于基因研究。人体中的细胞数量有10的14次方,而一个细胞的癌变就可能导致生命的终结。这一个癌细胞分解为RNA后就是10的9次方,变成蛋白质则是10的19次方。

华大基因董事长汪建告诉大家,这么大的数据量足以构成大数据,而基因研究面对的就是如此量级的数据研究和应用。

以对肿瘤细胞的定性、定量研究为例,汪建分析说,首先,要知道肿瘤细胞现在的基因是怎样的,而当出现一两个基因变化的时候,细胞又会变成什么样;其次,当众多肿瘤细胞的基因变化构成了相当量级的数据之后,能通过数据来展现细胞癌变的动态变化过程,从而进一步预测疾病。

汪建认为,这样的研究如果用在先天性疾病的防治上将会更有意义,比如有出生缺陷的婴儿就会越来越少。华大基因目前正在推进的“百万人基因计划”,就是一个名符其实的大数据应用。

因此,汪建从自己擅长的基因研究

出发,表达了自己对大数据的理解:从大目标出发,践行大数据,实现大科学,从而将研究成功运用到各个领域中去,形成大产业。

大数据让每个人变得更加透明,因此普通百姓是无秘密而言的

大数据更高级的阶段应该是数据的集成与共享,并把集成和共享本身作为一种商业模式——这是周涛的另一个观点。为此,他举了一个通过大数据帮助全球能源更好地进行配置的例子。

周涛所在的电子科大与电力集团达成合作,通过大量遥感装置和本地勘探数据了解到各个国家的能源供需情况;其次,通过与气象局合作获得气象数据,包括日照、风速风向、降雨等,从而对太阳能、风能、水能等新能源做到短期预测,并实现能源负载平衡。周涛认为,这些数据若再与电力集团的数据进行集成和分享,就能将各项能源更有效地调配使用。

社科院农发所社会问题研究中心主任于建嵘曾经说过,大数据让每个人变得更加透明,我们生活中的各种行为将通过数据被采集,并且分析出来,因此普通百姓是无秘密而言的。然而,这些数据却不是普通百姓可以看到或使用的,真正拥有这些数据的是部分企业、机构和政府,普通用户只是大数据的提供者,这就是“阶级”的产生。

社会学教授郭于华所关心的,是社会中的普通人怎样面对和使用新技术。譬如,新技术

是不是会缩小信息鸿沟和社会鸿沟?

郭于华曾经做过一个有关新生代农民工怎样使用信息技术的研究。她对研究结论的预判是:新技术将有助于新生代农民工的社会融入和城市融入。事实上,新技术确实给他们的生活带来了改变;其一,农民工是一个公共生活缺失的人群,信息技术给他们相互之间的交流带来了很多便利;其二,新媒体给了他们表达的渠道,这是以往所不曾有的;其三,他们可以通过新媒体即时调用各种所需要的信息,更加了解所生活的这个世界。

但即便如此,郭于华认为,我们也不能过分夸大信息技术对弱势群体的作用,因为其实呈现出比较强的马太效应:强者者可以更好地、更有利地利用这些新技术,同时他们也会对其他人加以控制。所以,新技术有没有可能穿透社会结构的屏障才是关键的问题。

像郭于华教授这样的调研,其实也正是创业者们所需要的。比如“豌豆荚”就曾经派人去广东东莞的企业做调研,了解打工族们对智能终端和移动应用的真正需求,体现了大数据背景下的一种人文关怀。



3D 打印进普通家庭 尚需技术改进和提升

头”匀速挤压出来,不需要使用模具和手工辅助工具,便可打印出由操作者设计的种种可以食用的巧克力。

据介绍,3D打印技术在重建物体的几何形状和机能上已经获得了一定的水平,几乎任何静态的形状都可以被打印出来,但是那些运动的物体和它们的清晰度难以实现了。“3D中国行”项目有关人士表示,3D打印是一个颠覆性的技术,对大多

数传统制造业来说,它带来的是技术的改进和提升,所以,“想要进入普通家庭,每个人都能随意打印想要的东西,还有很多技术需要提升。”

目前,我国正在制定的相关发展规划,明确提出将信息技术与制造技术深度融合的数字化、智能化制造作为今后发展的主线,中国将在2020年广泛推行数字化制造,并在2020年后重点推广智能制造。

最新科研动态

眼睛会说话确有其事

据新华社电 (记者蓝建中)俗话说,“眼睛是心灵的窗户”,“眼睛会说话”。日本一项新研究发现,人类和黑猩猩会因“同伴”的瞳孔放大或缩小而改变自己瞳孔的大小,从而证实眼睛在交际过程中确实发挥重要作用。

人类在交际时会根据各种因素不断调整自己的动作,比如对方点头时自己也会点头,在无意识情况下就会与对方动作保持同步。此前有研究显示,有的人在相互交流时,其眨眼也会保持同步,这种同步行动被认为是人类保持顺畅社交的基础。

人类和黑猩猩的瞳孔会根据周围的亮度放大或缩小,以调节到达视网膜的光线量。此外,瞳孔在兴奋时也会扩张。日本京都大学灵长类研究所副教授友永雅己等人,为了调查瞳孔大小的变化是否会保持同步,他们让18名学生和8只黑猩猩参与试验。

研究人员让学生和黑猩猩分别通过录

像观看人类和黑猩猩的瞳孔扩大和缩小,并用特殊摄像装置记录其瞳孔变化。他们发现,学生和黑猩猩在看到同类录像时,会瞬间改变瞳孔的大小,而且学生瞳孔的变化参数更大。不过,如果让学生看黑猩猩或让黑猩猩看人类瞳孔变化的录像时,两者的瞳孔都没有明显变化。

此前,研究人员曾发现黑猩猩之间会“传染”打喷嚏。这项新研究则显示,黑猩猩对于“同伴”瞳孔的变化也非常敏感,这对于了解灵长动物的社交演化特点是个重要线索,而且,人类和黑猩猩只与“同伴”的瞳孔尺寸保持同步,这可能与共鸣等高层次社交有关。人类瞳孔的变化参数更大,可能意味着人类的交际层次更高。

友永雅己指出,“与‘同伴’的瞳孔变化保持一致,可以更容易领会对方的情绪,这有可能成为交际的润滑油。这说明‘眼睛会说话’确有其事。”



植物也有喜怒哀乐

人有七情六欲,动物也有喜怒哀乐。科学家发现,与人和动物一样,植物也拥有思维,同样“情势汹汹”。

最先发现植物懂得“高兴”的人是美国情报专家巴克斯特。1966年2月的一天,正在庭院给花草浇水的巴克斯特看到植物,突然蹦出一个念头,这些可爱的家伙是不是在装深沉呀?也许是经常与间谍、情报打交道的原因,巴克斯特竟想一探植物的“内心隐秘”。他把一台测谎仪安装到一株干早已久,叶片发蔫的矮叶桃上,然后给它浇水。当水缓缓到达矮叶桃的根茎部时,测谎仪出现了奇妙的曲线变化图形,居然跟人在激动时测到的曲线图形很相似。

巴克斯特决定再测一下植物会不会“发怒”。这一次,他选择的植物是一株石榴树。他依然将测谎仪绑在石榴树身上,所不同的是这次选用火药——点燃一根沾上油的火把,凑近石榴树的叶片,还没有接近叶片,测谎仪的指针即发生剧烈摆动。当火把燃烧叶片一分钟时,测谎仪上的曲线竟然跟人绝望发怒时的曲线一模一样。

之后,巴克斯特又做过一次有趣的实验,对象依然是石榴树,但只是假意地用火把攻击它,而不是真正地接近。让人惊奇的是,测谎仪上的曲线无动于衷。原来,植物也能识破人的“诡计”。

20世纪70年代,前苏联科学家维克多也发现了植物的“哀”与“乐”。他向一位病人施用了催眠术,控制其感情,然后将一盆水晶兰放在病人的床边。维克多用了一台脑电仪,把病人的手和植物的叶子连接起来。之后,维

人类已约有300万年的悠久历史,远在人类未出现以前,地球上已布满了各种各样的植物,大约90%以上的漫长岁月人类主要靠采集野生植物的果实、根、茎、叶和渔猎为生。在我国远古神话中就有“神农氏尝百草,一日遇七十毒”的描写。神农氏尝百草,实际上就是我们的先民尝百草。在长期的实践中,人们逐渐了解到哪些植物可食,哪些植物不可食。对于可食的植物,人们由采集逐步发展为栽培,形成了今天丰富的蔬菜种类。

作为人们日常食用的200多种蔬菜中,绝大多数是安全无毒的。但有少量蔬菜在不同的生长期、不利的贮藏条件下会产生一些毒素,人们不慎食用后会对身体产生不良影响。

●毒蕈中毒 蕈类又称蘑菇,毒蕈是指食后可引起中毒的蕈类。毒蕈在中国有100多种,对人有生命威胁的有20多种,其中含有剧毒可致死的有10种。

每年夏季会陆续发生毒蕈中毒,并有多人死亡。原因是采摘蘑菇的人不会识别毒蘑菇,自食或卖出后引起中毒。比如牛肝菌,目前国内已发现有20多个品种可食用,但其中的粘盖牛肝菌(滑肚子)、褐绒盖牛肝菌(松毛菌)、细网牛肝菌(魔王牛肝菌)、红网牛肝菌、黄粉牛肝菌等品种有一定毒性,误食后会中毒。

●豆角中毒 俗称的“毒扁豆”中毒多因

海王星今天“冲日”

据新华社电 (记者周润健)8月29日,海王星将上演“冲日”表演,此后20多天中,由于海王星离地球的距离最近,在透明度较高的晴朗夜空,有兴趣的天文爱好者可借助望远镜一睹这颗淡蓝色神秘星球的风采。

海王星是太阳系八颗行星中距离地球最远的,也是看上去最暗的。由于海王星是依据纯天文理论计算而被发现的行星,在天文学史上其又被称为之“笔尖上的发现”。

中国天文学会会员、天津市天文学会理事史志成说,由于海王星本身并不发光,完全靠反射太阳光,因此亮度不高,冲日时也仅为7.8星等,肉眼无法看到。

天文专家说,要想更容易找到海王星在天空中的位置,除了使用天文望远镜,还需借助电子星图的帮助,才能更好地一睹这颗淡蓝色星球的风采。

“藏”在蔬菜中的有毒有害物质

没有炒熟而发生。未断生的豆角含有毒蛋白和皂素,毒蛋白具有凝血作用,皂素是一种能破坏红细胞的溶血素并对胃肠有强烈的刺激作用。豆角如果没有煮熟煮透,人吃了之后1-5小时就会引起中毒,轻者头痛头昏,重则恶心、呕吐、腹痛。

食用豆角要将豆角两头含毒素较高的部分去掉,在烹饪时炒熟煮透,至豆角原有的深绿色消失为止。

●未成熟西红柿中毒 未成熟的西红柿为青绿色,含有一种叫龙葵素的生物碱,食用后可导致中毒,轻者口腔感到苦涩,严重者会出现中毒现象。

青西红柿变红以后,就不含龙葵碱了,所以一定要吃已经红透的。目前市场上有一种叫“绿宝石”的小型西红柿,成熟后颜色也是绿色,但可放心食用。

●鲜黄花菜中毒 人们日常食用的是干制后的黄花菜,但也有人喜食鲜黄花菜。鲜黄花菜中含有秋水仙碱,一个成年人如果一次摄入50~100克即可中毒,严重者可出现血便、血尿、尿闭等现象。

要防止出现鲜黄花菜中毒,可将鲜黄花菜在沸水中稍煮片刻,再用清水浸泡,就

可将大部分水溶性秋水仙碱去除。也可将鲜黄花菜煮熟、煮透,再烹调食用。

●发芽土豆中毒 土豆中含有的龙葵素对人体有害。平时土豆中龙葵素含量极微,一旦见光变绿、发芽、溃烂,其含量急剧增加,可高出平时的40~70倍,吃了之后,轻者恶心、呕吐、腹痛、腹泻,重者可出现脱水、血压下降、呼吸困难、昏迷、抽搐等症状,严重者还可因心肺麻痹而死亡。

为了防止土豆中毒,可将土豆贮藏于干燥阴凉的地方,防止发芽。发芽或皮肉呈黑绿色的土豆,最好不要食用。

●蚕豆过敏 蚕豆中含有巢菜碱苷,葡萄糖-6-磷酸脱氢酶(G-6PD)有遗传缺陷者应慎食蚕豆。该人群在食用新鲜蚕豆后皆会发生急性溶血性贫血症——蚕豆病,致病机制尚未十分明了。

有上述遗传缺陷的人,大多食蚕豆后1至2天发病,早期症状有厌食、疲劳、低热、恶心、不定性的腹痛,接着因溶血而发生眼黄染及全身黄疸,出现酱油色尿和贫血症状。严重时会有尿团、休克、心功能和肾功能衰竭,如不及时抢救可于一至两天内死亡。



日照流行 新能源公交车

近日,山东日照市交通部门新购入150台新能源电动公交车,将全线更新到14条公交线路中。更新后,日照市电动公交车的数量将达到290台,超过全市公交车辆总数的一半。据悉,这些电动公交车年可节约燃料油近200万升,大大降低了污染物的排放,运营成本也降低了约60%。

新华社记者 范长国摄

鄱阳湖两重点流域 江豚数量不足百头

据新华社电 (记者袁慧晶 程迺)记者近日从江西省山江湖办获悉,据观测情况显示,鄱阳湖重点流域孑子水域和龙口水域观测到的江豚种群数量不足百头。

参与此次科考的南昌大学生命科学研究院教授吴小平分析说,此次公布的是在两个江豚主要栖息地在一段时间内一定范围内观测到的江豚数量,并不是整个鄱阳湖流域的江豚数量,所以不能作为鄱阳湖重点流域江豚数量锐减的证据。记者了解到,由于长江流域水体污染加剧,人类肆意采挖江砂,非法使用渔具等原因,给长江江豚的生存和繁育带来一定影响,近20年来种群数量快速衰减。鄱阳湖是长江江豚重要的栖息地,经过多年保护,江豚的数量比较稳定。江西省渔政局监测数据显示,近几年,鄱阳湖流域的江豚数量稳定在450头左右。

长江江豚是全球唯一的江豚淡水亚种,在地球上生存超过2500万年,被誉为长江生态的“活化石”和“水中大熊猫”,属中国二级保护动物,数量不足千头,主要分布在长江中下游干流以及洞庭湖和鄱阳湖等水域。

国家种业科技成果 产权交易中心启动

本报讯 (记者黄哲雯)在全国现代种业发展工作会议召开期间,依托中国农科院建设的国家种业科技成果产权交易中心正式启动运行。

该交易中心的建设是我国深化种业改革进程中具有标志性意义的一件大事,对于推动民族种业发展、做大做强现代种子产业将发挥重要作用。据了解,该交易中心目前已建成近1000平方米办公用房,自主研发了国家种业科技成果电子交易平台、种业科技成果价值评估平台,建立了全球种业知识产权数据库,制定了《种业科技成果转化产权交易管理暂行办法》、《种业科技成果展示和产权交易规程》等一系列办法和技术规程,组建了一支包括产权交易、评估、法律、生物技术、经济管理、信息技术等30余人组成的专业团队。



宁夏回族自治区盐池县位于毛乌素沙漠边缘,常年干旱少雨,风大沙多,生态环境十分脆弱。近年来,盐池县在全自治区率先实施封山禁牧工作,坚持草原禁牧与舍饲养殖、封沙育林与退牧还草、防沙治沙与沙产业开发相结合。目前,全县人工造林以每年10万亩的速度推进,林木覆盖率和植被覆盖率分别达到31%和68%,与十年前相比分别提高了17%和11%,逐步形成了跨区域综合治理、“乔灌草”相得益彰的“盐池生态模式”。

新华社记者 彭昭之 摄

地铁盾构施工 成功穿越南宁南湖

本报讯 近日,继6月23日右线盾构成功穿越南湖之后,由中铁四局承建的南宁地铁1号线13标经理部左线地铁盾构也成功穿越南湖。至此,南宁市首条地铁左右线盾构施工均成功穿越东西宽约260米的南湖,为项目2015年底全线贯通奠定了坚实基础。

为确保盾构穿湖施工安全,该局南宁地铁1号线13标经理部选取具有盾构施工经验的管理、技术人员和生产骨干组成盾构施工架子队,盾构始发前顺利、严格地完成盾构组装、调试、验收;编写了《南湖加固方案》等针对性强的盾构施工方案20余个,为盾构施工奠定技术基础;盾构始发前将埋深4.2m~6m浅埋段进行膜袋围堰,围堰后对隧底至隧顶范围内进行注浆加固处理;对盾构机操作员、管片拼装人员和注浆人员定期进行技术培训,提高盾构施工人员的专业技能;综合各方意见不断改进盾构施工工艺,确保了施工安全

和质量。

(张召王征日)