

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容和观点?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,即使像是像BAT这样的互联网巨头,它所

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,巴克斯特竟想探植物的“内心秘密”。他把一台测谎仪安装到一株早干枯的、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柑树。他依然将测谎仪绑在石柑树身上,所不同的是这次选用火柴——点燃一根沾上油的火把,凑近石柑树的叶片。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后,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,对象依然是石柑树,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,

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,对人生命有威胁的有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



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——

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。在这一座座孤岛被打破之前,大数据都无法称“全”。这种大而全的数据,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。

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,这些都需要依靠于基因研究。人身体中的细胞数量有10的14次方个,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。

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没人看得见的地方,观察体验店里的行为——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,这些都需要依靠于基因研究。人身体中的细胞数量有10的14次方个,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。

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没人看得见的地方,观察体验店里的行为——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

本报讯 (记者于究尼)日前,“穿阅·中塔·2014 3D 中国行”大型3D体验展自北京开幕以来,将3D打印、3D扫描、3D雕刻等一系列先进的3D技术带给了消费者。据悉,去年以来国内3D打印企业诞生了五六百家,中国有潜力成为世界最大的3D打印市场,这一市场在未来3年可从约10亿元人民币增长到100亿元。

3D打印用金属、塑料或陶瓷等材料层层堆砌,“印”出3D立体成品。美国西雅图发明家霍尔曼预测,随着这项科技改进,大型高速打印机将带动材料的需求。尤其在一些3D打印关键专利今年到期后,制造业将更加广泛使用3D打印机。记者在体验展上看到,一台英国的3D巧克力打印机,使用液态巧克力作为“油墨”,让巧克力从10毫升的“注射



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近的,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,在天文史上其又被称之为“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肠胃有强烈的刺激作用。豆角如果没有炒熟煮透,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。严重时有尿困、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。

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,可将土豆贮藏在干燥阴凉的地方,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。该人群在食用青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。严重时有尿困、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。

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。

● **豆角中毒** 俗称的“毒扁豆”中毒多因

将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,可将土豆贮藏在干燥阴凉的地方,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。该人群在食用青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。严重时有尿困、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。

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。

● **豆角中毒** 俗称的“毒扁豆”中毒多因

将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。

● **豆角中毒** 俗称的“毒扁豆”中毒多因

将可将大部分水溶性