

·科技纵横捭阖·

大数据与小数据以及西医与中医

现在大家都热炒大数据,其实小数据也很值得注意。昨天我有点胃痛,于是想:生活中的什么变化引起了胃痛呢?想了一会发现虽然我每天喝1两酒,但是前天喝的酒换了一个牌子。可能就是这个小牌子的酒引起我胃痛。这个小数据成为我分析健康情况的依据。

康奈尔大学的计算机科学教授 Deborah Estrin^[1]说,她从父亲2013年去世前几个月的数字社会脉动中看到了信号的异常——父亲停止了发电子邮件,不去超市买食物,到附近散步的距离也越来越短。而这些逐渐衰弱的状态却从心脏病科医师那里看不出来,因为心电图并没有显示异常。所以,日常生活中的小数据能够给我们更多警示,它可以看作是一种新的医学证据。这启发 Estrin 开始研究小数据。有人说,小数据是从大数据中拿出来的,研究大数据自然就包括研究小数据。其实不然。首先,小数据不是从一个超大数据库里挑出来的,而是在小数据产生过程中,用APP(移动应用程序)搜集起来的;其次,大数据分析要从大量的样本中统计分析出规律性的东西,而小数据却是非常个性化的数据,针对某一个人、某一件事;再次,个人数据资料室是高度隐私的信息,除了本人可以存取以外,其他任何人甚至系统管理员都不能存取。如果把一个数据元素允许存取人数的倒数称为隐私度的话,个人数据资料室的隐私度为1,为最高值。而电子邮件的隐私度 ≤ 0.5 ,因为至少发送者和接收者可以存取。一般的公用信息,由于大家都可以存取,其隐私度为0。要做到隐私度为1,技术实现起来是很难的。

研究个人数据可以了解人们某方面的情况。譬如在医学领域,Estrin认为:mHealth是将移动装置用于医疗和公共卫生。mHealth开发一种体系结构,集成个人使用特定应用程序的数据,以帮助医疗服务人员提出推荐性的建议。她对小数据的定义是:your row of their data。如果你把使用信息的移动用法、有线用法、公用程序的用法、电子商务行为、搜索行为、社交媒体、电子邮件、智能汽



车使用数据、游戏、音乐、视频等等组合起来,可以看出什么信息?如果分析年老的父母的这些集成数据,又可以看到关于他们健康和幸福感方面的什么信息?是否可以通过这些集成的数据,比较不同的医学治疗方案?Estrin的目的是开发一个应用的生态系统,个人可以运行他/她自己的数据流之集合,这称为个人数据资料室。她希望能够允许个人数据跟踪。个人数据应用程序接口可以产生实时的个人数据APP。小数据描绘个人随时间变化的健康情况,由一个新的云端APP产生,它可以连续地、隐私地追踪分析个人的工作、购物、睡眠、吃饭、锻炼和通信等情况。

从日常生活到社会网络、搜索引擎、在线游戏和电子商务,我们每天存取和使用它们留下的“足迹”,可以形象地称为个性化的社会脉动。一个脉动体现出个人幸福感的微妙而有意义的变化,补充自己认购的APP。该APP可以根据个人的数字轨迹自动分析位置和词汇数据,来建造个人日常活动的比较图,或者从运行移动跟踪的APP看到药物对个人的治疗效果。尤其是许多慢性病的变化,不仅与许多复杂因素有关,而且还因人而异,病情的变化都是缓慢的。所以,个人很难靠自己的记忆追踪新治疗方法的效果。但是,这些数据需要从杂乱的数据源中提取真实的内容。因此,数据解放是首要任务,即以丰富的APP市场和服务,从移动和网络服务中获取数据。

西医是根据化验结果或其他科学仪器测试得到的结果进行的诊断,它是从大数据里得出结论,再应用到个体。而中医不同,它是根据用个性化的方法从个人数据资料室里拿出来分析的小数据,对个人进行医疗和保健的一种手段。中医用望、闻、问、切的方法,推断应该用什么

药来缓解病情;再根据服药以后的病情发展情况,随时调整药方。研究自动化的人都知道,自动化系统的自修复能力是系统最重要而又不易达到的性能,而人有很强的自修复能力。中医就是充分利用个人不同的自修复能力治病,而不是像西医那样主要靠从外部给以药物或手术治疗。

再说西药,它根据化学试验,得出药物的化学名称、结构式、分子式,再分析其可能的副作用,进行实验,因此对大多数有相似病症的患者有疗效,而副作用较小。这是从大数据里得来的结论。中药则不同。先是神农尝百草,打下基础,再根据长期的经验积累,得出某些草药搭配可能治愈某种病。病人服用后,病情好转则继续;反之则再调整药方。因此,即使相同病症的2个人,所用的中药品种或者剂量可能不同。如果要求中医列出药物的结构式、分子式,并给出可能的副作用,那真是勉为其难了。当然,中医诊断过程中隐含的推断过程,也有错误的可能。推断不正确,病症可能就治不好。但是,中国几千年的医疗经验表明,中医这种个性化服务,治好病的概率还是挺高的。对于某些疾病,有时已不得不采取中医疗法。

有人反对中医,说它不科学。其实,大数据是科学,小数据也是科学。中医根据个人的情况用药,反复试验,也很科学。我认为,老人尤其适合中医。因为西医动不动就做手术,而老人可能经不起手术的折腾,术后的恢复能力也差,用中药调理也许更合适。本人以为,中西医结合不是治疗疾病更好的方法,中医应该走自己的路。

参考文献

- [1] Deborah Estrin. Small data, where $n = me$ [J]. Communications of the ACM, 57 (4): 32-34.

文阅应骋

作者简介 中国科学院计算技术研究所,研究员,IEEE终身Fellow。

栏目主持人 关增建,电子信箱:guanzz@sjtu.edu.cn。

(编辑 王丽娜)