



E-CAS

中国科学院信息化工作领导小组办公室

The Cyberinfrastructure Office, Chinese Academy of Sciences

需求牵引

开放前瞻

整合资源 服务第一

[首页](#) | [组织体系](#) | [规章制度](#) | [院机关动态](#) | [研究所动态](#) | [出版物](#) | [参考资料](#) | [专家视点](#) | [媒体关注](#) | [联系我们](#)

站内搜索

搜

[首页](#)>国内外的动态

李国杰院士谈大数据热：近期力戒忽悠

2013-01-28 【大 中 小】 【打印】 【关闭】

“其实，生物信息学、脑科学、空间科学等基础研究早就采用数以PB级计的大数据，却没能引发大数据浪潮。大数据如今引起各方重视，主要还是因为它蕴藏着巨大的经济价值。”

近日，在北京召开的“大数据背景下的计算机和经济发展高层论坛”上，中国工程院院士李国杰表示，我国开展大数据研究，一定要以企业为主体，充分调动企业和经济界学者的积极性。

大数据在经济领域应用的一个有趣案例，是通过社交网络预测股价涨落。社交媒体监测平台DataSift监测了社交网站Facebook首次公开募股当天Twitter上的情感倾向与Facebook股价波动的关联。在Facebook开盘前，Twitter上的情感逐渐转向负面，25分钟之后，Facebook股价便开始下跌。而当Twitter上的情感转向正面时，Facebook股价在8分钟之后也开始了回弹。最终当股市接近收盘、Twitter上的情感转向负面时，10分钟后Facebook的股价又开始下跌。该机构得出最终的结论是：Twitter上每一次情感倾向的转向都会影响Facebook股价的波动。

“预测股价，是公认的很难做到的事，但大数据分析也能发挥作用。”李国杰说，“其他过去认为做不到的事情，也许大数据分析都能派上用场。因此，我们的当务之急是多实践，实践出知识，实践出理论。只有当每个领域的数据分析做得相当深入时，才能提出数据科学的共性问题。”

虽然大多数单位其实仍处于“小数据”处理阶段，但李国杰认为，企业和政府部门都不必太在意自己正在分析的是不是“大数据”，更不必花精力对各种名词和定义作无谓争论。“只要在纵向上有一定的时间积累，在横向上有较丰富的记录细节，通过多个源头对同一个对象采集的各种数据有机整合，进行认真仔细的数据分析，就可能产生大价值。”

李国杰还提醒，要对我国在大数据方面的技术基础落后有清醒认识。由于我国人口最多，网民也最多，许多人估计我国每年新数据的增量应该在世界上名列前茅。但仅就数据量而言，我国还不是数据大国。根据麦肯锡公司的报告，2010年美国新增数据量为3500PB，欧洲为2000PB，中国仅为250PB。也就是说，我国每年新增数据量仅为美国的7%、欧洲的12%。

“当然，大数据的关键并不在数据量的大小，而在于获取有效信息的能力和数据分析能力的高低。”李国杰说，“但在大数据领域，我国与世界先进国家也不在同一条起跑线上。”

现在涉及大数据的会议很多，但李国杰觉得政府、企业和学术界未必真正理解大数据的重大意义。“许多人对变革性技术的效益估计往往是，近期过于乐观，远期估计不足。”

他自己的判断是，如同云计算、物联网一样，近几年内大数据的市场规模不会很大，全世界大概只有数百亿美元。但是，大数据市场的增速将是传统IT产业的好几倍。“对近期的大数据市场不能估计过高，不能自己忽悠自己。开展大数据研究和应用一定要有长远眼光。”

站点搜索

请输入关键字

搜索

