



用大数据丰富和完善官方统计数据

国家统计局统计科学研究所

孙学光



主要内容：

1. 大数据为官方统计提供了新的机遇
2. 大数据在中国官方统计中的应用
3. 为大数据应用创造良好环境



1

1.1

大数据时代，数据资源急剧增加，数据量巨大，极大丰富了官方统计数据来源。



1.2

大数据来源

来源分类	数据信息类型	数据归属
行政记录数据	个人信息记录数据	公安、卫生、教育、人力资源和社会保障等部门
	单位信息记录数据	质检、工商、税务、民政和编办等管理部门
	自然和资源记录数据	国土资源、环境保护、气象、地震、海洋、测绘等部门
	其他管理记录数据	知识产权、海关、出入境管理及资质评定等部门
商业记录数据	电子商务交易数据	各网上商城、网店
	生产经营管理数据	国民经济各个行业的企业
	信息咨询报告数据	专业数据库公司，咨询机构
互联网(包括搜索引擎)数据	社交网数据	国内各社交网站注册的博客、微博、微信等，国外专门社交网如Facebook、Twitter等
	媒体数据	新闻媒体、广播电台、电视台和出版社等
	搜索引擎数据	百度和谷歌等



1.3

大数据类型众多，有音频数据、视频数据、文本数据，等等。更加直观、简化，便于决策者和广泛的利益相关人士理解和使用数据。



1.4

很多大数据可以实现实时提供，满足即时性需要，有助于提高官方统计的时效性，方便用户。



1.5

大数据CPI与官方CPI比较

指数名称	每日网上价格指数	淘宝网ISPI	百度搜索数据CPI	官方CPI
指数用途	物价变动趋势和变化程度	网络零售价格变动趋势和程度	预测物价变动趋势	居民消费价格变动趋势和变化程度
发布频率	日	日	日	月
预警能力	领先官方CPI	领先官方CPI	领先官方CPI	一般
切合实际	较高	较高	较高	一般



1.6

大数据的利用，为传统统计数据提供更多验证和补充，有利于增强官方统计数据真实性，进一步完善数据质量控制体系。



2

2.1

2013、2014年，中国国家统计局分两批与百度、阿里巴巴、腾讯、浪潮等17家大数据企业签署了大数据战略合作框架协议，打造大数据合作平台。



2.2

在许多官方统计领域开展大数据应用，例如：



2.3

在贸经统计中：

2015年起，按月发布全国网上零售额，如最新的2016年1-7月份，全国网上零售额26268亿元，同比增长27.5%。

同期社会消费品零售总额182966亿元，同比增长10.3%。



2.4

在农业统计中：

结合使用卫星影像、空间高分辨图像和地面定位系统、传感器、物联网等技术，获取有关农作物种植面积、农作物种类等方面的海量数据。



2.5

在居民消费价格（CPI）统计中：

2014年起，试点手工采集和网络抓取部分商品的网络价格并测算价格指数，2016年，在全国进一步推广网络采价。



2.6

在行政记录使用方面：

正在积极推进官方统计与电子化行政记录的共享，加大电子化行政记录的应用力度。



2.7

在经济分析预测方面：

与百度公司合作开展“经济雷达”项目，探索利用搜索引擎数据建立模型，预测宏观经济走势。



2.8

在数据发布方面，按照公开为常态、不公开为例外原则，通过互联网为用户提供600多万笔数据和丰富多样的可视化产品。开发“中国统计”、“数据中国”等移动设备客户端，建立政府统计微信、微博平台，及时推送官方统计数据。



3

3.1 积极推动制定大数据相关标准

3.2 积极推动数据开放共享

3.3 积极推动大数据应用法律法规的制定



Thanks !

sunxueguang@hotmail.com