

农业大数据的获取与利用

谢润梅 (中国科学技术信息研究所, 北京 100038)

摘要 大数据的发展日新月异,已经深入到包括农业在内的各个行业。因此,农业大数据作为农业信息化的发展趋势,是新一代信息技术的集中反映,是一个具有无穷潜力的新兴科技产业发展方向。该研究结合农业大数据的基本特征与产业链构成,介绍了农业大数据的基本概念,基于相关农业信息平台、农业数据库等渠道,阐释了农业大数据的获取方式,并从良种培育、精准种植、农业生态环境监测、天气预测等维度对农业大数据的利用进行探讨,以正确面对农业大数据面临的机遇与挑战,为农业的发展、创新奠定基础。

关键词 农业大数据;数据获取;数据利用;智慧农业

中图分类号 S126;TP311.13 **文献标识码** A **文章编号** 0517-6611(2015)30-383-03

Acquisition and Utilization of Agricultural Big Data

XIE Run-mei (Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract The development of big data with each passing day, has already gone deep into the various industries including agriculture. Therefore, the agricultural big data as the development trend of the agricultural information, is a new generation of information technology and a concentrated reflection of the new technology industry. Based on the basic features and industrial chain composition, this paper expounds the basic concepts of agricultural big data. Based on the related agricultural information platform and agricultural database, this paper elaborates the acquisition ways of agricultural big data, and discusses the utilization of agricultural big data from aspects of seed breeding, precision planting, agricultural ecological environment monitoring, weather forecast, in order to face the opportunities and challenges and lay the foundation for the development and innovation of agriculture.

Key words Agricultural big data; Data acquisition; Data utilization; Wisdom agriculture

农业作为基础产业,为紧跟大数据时代步伐,传统的农业生产方式应向数据驱动的智慧化生产方式转变。而云计算、大数据、互联网等科学技术将是这场变革的主要推动力。基于此,理论界对农业大数据进行大量的探索,张浩然等^[1]介绍了大数据的概念与应用技术,温孚江^[2]阐释了农业大数据研究的战略意义与协同机制,孙忠富等^[3]论述了农业大数据管理整个生命周期;然而在农业大数据领域研究最多的还是农业大数据的应用,柳平增^[4]以渤海粮仓科技示范工程大数据平台为例,介绍了农业大数据平台在智慧农业中的应用,许世卫^[5]介绍了农业大数据对农产品监测预警的重大作用,郭承坤等^[6]就发展农业大数据的主要问题和主要任务进行了探讨。可见,在农业大数据研究方兴未艾的今天,探讨其获取与利用有着非常重要的意义。

1 农业大数据基本概念

农业大数据是大数据的一个分支,是大数据理论、技术和方法在农业领域的应用实践。农业大数据的特性满足大数据的5个特性,一是数据量大(Volume);二是处理速度快(Velocity);三是数据类型多(Variety);四是价值大(Value);五是精确性高(Veracity)。农业大数据涉及的领域广,环节多,是跨行业、跨专业的数据集合。因此,从某种程度讲,农业大数据还具有复杂性(Complexity)及分散性(Dispersion)的特征。

农业大数据作为一个新生事物,其内涵还没有准确的界定。目前我国已形成农业大数据“农业数据源—农业数据管理—农业数据分析—农业数据应用”的产业链,而IT基础为产业链中的各个环节提供基础支撑。

农业数据源主要从数据获取方式及数据的类型进行描述。数据获取方式主要包括Web、移动端、流媒体、传感器等;而数据类型主要包括结构化数据、半结构化数据、非结构化数据。农业数据源是农业大数据产业生态的基础,为农业数据管理、农业数据分析及农业数据应用做铺垫。

农业数据管理作为大数据产业技术重要的组成部分,承载着农业大数据的采集、传输、存储与处理等功能。农业数据采集与传输是基于计算机硬件、软件、互联网等工具与技术对数据进行生产集聚;农业数据存储从简单的DAS、NAS、SAN等存储方式,逐步扩展到数据聚合平台,最终形成云服务;而农业数据处理是以海量数据分布式处理、非结构化数据处理、实时数据处理为主要方式,基于Hadoop、Spark等核心技术,实现数字、文本、图像、视频、音频及多媒体等数据的处理功能。

农业数据分析作为农业大数据产业链的核心环节,肩负着从海量数据中挖掘其行业价值的使命。即运用大数据挖掘工具及技术,对农业数据进行定制化分析,实现数据服务的个性化、精准化和智能化。

农业数据应用作为农业大数据产业的落地点,呈现分析挖掘的数据价值,还原大数据结论反映的行业问题,即将农业大数据应用在粮食安全、土壤治理、病虫害防治、动植物育种、农业结构调整、农产品价格、农副产品消费等领域,解决农业生产过程遇到的诸多问题。大数据技术与农业生产的紧密结合,为从事农业领域工作的师生、社会公众、政府部门及涉农企业提供了新方法、新思路。

2 农业大数据的发展机遇

目前,不论是国际、还是国内,中央、还是地方,都对大数据发展尤为重视。农业大数据作为大数据的重要分支,面临着前所未有的发展新机遇^[7]。

作者简介 谢润梅(1983-),女,河北涿鹿人,在读硕士,从事竞争情报研究。

收稿日期 2015-09-28

联合国在2012年发布了大数据政务白皮书,指出大数据对于联合国和各国政府来说是一个历史性的机遇,人们如今可以使用极为丰富的数据资源,来对社会经济进行前所未有的实时分析,帮助政府更好地响应社会和经济运行。欧盟支持大数据建设的第一步是进行监管和推动公开,欧委会于2012年1月提交了“运用数据保护条例”等规定,旨在以较低的费用和简捷的重复使用条件,更加便捷地使用和重新使用公共数据。2012年3月,美国宣布投资2亿美元拉动大数据相关产业发展,将“大数据战略”上升为国家意志,并表示一个国家拥有数据的规模、活性及解释运用的能力将成为综合国力的重要组成部分。

李克强总理在2014年3月5日第十二届全国人民代表大会第二次会议做政府工作报告时提出,要设立新兴产业创新创业平台,在新一代移动通信、集成电路、大数据、先进制造、新能源、新材料等方面赶超先进,引领未来产业发展。

大数据产业发展,受到国内外的高度重视。然而,“十二五”期间,全国现代农业发展规划(2011-2015年)指出,加快发展现代农业,既是转变经济发展方式、全面建设小康社会的重要内容,也是提高农业综合生产能力、增加农民收入、建设社会主义新农村的必然要求。因此,大数据作为新一代信息技术,在农业领域的跨界应用任重道远。农业大数据应在良好的大环境中,抓住机遇,紧跟时代发展潮流。

3 农业大数据的获取

考虑到农业大数据具有多类型、复杂性及分散性的特点,其获取的难度可想而知。农业大数据分为结构化农业数据和非结构化农业数据。结构化农业数据是专业化、系统化的农业领域数据,即行数据,存储在数据库里,可以用二维表结构来逻辑表达实现的数据,如专业农业数据库、农业信息平台等。因此,结构化农业数据的获取性更好。

随着互联网及物联网的发展,我国农业网络数据资源越来越多。农业信息网站主办单位主要包括各级政府部门、农业科研教育机构及涉农企业单位。农业数据类型也呈现出多样化趋势,如农业图书、期刊杂志等正式资源,政府农业简报、农业相关会议信息等非正式资源,还有一些诸如农产品市场信息、农业科技成果等动态信息资源。总之,从涉农国家部委到科研院校、企业单位都在完善自己的农业数据资源信息库。

中国农业科技信息网涵盖了农业科技领域比较全面的信息资源,如世界农业科技动态数据库、中国农业科学院研究生学位论文数据库、农业科技信息数据库、农业科技成果数据库、国土资源文摘数据库、农业科技项目数据库、国外科技资料目录数据库、科技期刊目录数据库、图书目录数据库、农业科技报纸文摘数据库、农业科技期刊全文数据库等不同类型数据库。

当然,我国还有一些其他涉农数据库,如全国农业经济统计资料数据库、中国农作物种质资源数据库、中国畜禽品种资源数据库、中国农林文献数据库、农业合作经济数据库、农牧渔业科技成果数据库、农产品集市贸易价格行情数据库

等,都为我国农业信息化作出突出贡献。

此外,我国很多综合性数据库也涵盖了大范围的农业领域数据,如万方数据、中国知网、维普等数据平台,通过关键词、中图分类号、学科码、行业码等检索途径,都可以获得很多论文、成果、专利、标准等有关农业领域的大量信息。

除了国内这些结构性涉农数据库,国际知名的农业数据库系统有CABI、AGRIS、AGRICOLA,其质量高、数量多并具共享性,为农业大数据获取提供了便利条件。

然而,虽说涉农信息网站和农业数据库资源数量大,种类多,数据质量却参差不齐。许多规模小的涉农网站在数据质量上存在严重问题;部分企业构建的涉农数据库存在数据结构混乱、建库重复性现象严重、数字化程度偏低、资源共享难等问题。同时,随着信息技术与现代农业发展,非结构化数据急速增长,与结构化数据相比,非结构化农业数据具有数据存储分散、数据格式多样化及数据类别混乱等特点,增加了农业数据收集、利用的难度。因此建议国家有关部门建立标准化数据存储及发布平台,提高数据利用率,为农业大数据在农业领域的进一步应用奠定基础。

4 农业大数据的利用

农业大数据应用,依托部署在农业生产现场的各种传感节点(环境温湿度、土壤水分、二氧化碳、图像等)和无线网络,完成农业大数据采集、传输、存储、处理等环节的数据管理,结合大数据分析挖掘技术,最终实现农业生产环境的智能感知、智能预警、智能决策、智能分析、专家在线指导,为农业生产提供精准化种植、可视化管理、智能化决策。如今,国内外农业大数据快速发展,其典型应用体现在以下几个方面。

4.1 培育良种 对人类营养状况数据、生物群体的基因组等数据进行分析,通过对农作物的基因组进行测序,培育一些营养价值较高的作物品种,有助于提升人们的健康水平。

4.2 精准种植 精准种植是基于3S技术(遥感技术、地理信息系统、全球定位系统),实施一整套现代化农事操作技术与管理的系统,主要用于土壤肥力的精准化监测、农田边界图智能管理、病虫害精准定位和防治、精准施肥和灌溉等。

4.3 农业生态环境监测 生态环境监测主要是对农作物生长相关的土壤、水质、气候、气象和灾害等情况进行全面监测,并对它们之间复杂的关系进行分析,可以判断不同生态环境对农作物生长的影响。

4.4 天气预测 通过分析历史天气变化规律,建立天气识别模型,结合当前的气候特征和近期的天气情况,对某地未来一定时期的天气状况进行预测分析,对农业生产和日常生活具有重要的指导作用。

4.5 农产品与食品安全监测 通过对农产品与食品的产地环境、产业链管理、产前产中产后、仓储加工、物流等数据进行监测,并通过对影响农产品与食品安全的关键性指标设置警兆、分析数据、发布预警、寻找警源、消除警情等一系列操作,实现对农产品与食品的安全监控。

4.6 农产品物流 农产品物流涉及农产品的收购、储存、加

工、包装、运输、卸载搬运、配送等环节,通过整合、分析各个环节的数据,不仅可以连接农业主体和消费需求主体,还能实现农产品保值增值,甚至可以为整个物流管理提供有力的决策支持,如物流中心选址、最优化配送路线、合理管理库存等。

4.7 农产品市场追踪 通过对农产品销售价格、销售量、销售需求、消费者购买行为数据进行分析,可以判断农产品的供需、价格变动以及消费者的购买习惯等。我国的农业大数据应用虽已取得阶段性成果,但在整个农业产业链推广过程中仍存在许多问题。而农业大数据作为农业信息化的发展趋势,是新一代信息技术的集中反映,是一个具有无穷潜力的新兴科技产业方向。

5 农业大数据面临的挑战

大数据与农业的融合,作为一新兴的技术领域,存在着一些困难与挑战。重点体现在数据采集、数据存储、数据处理、数据分析与数据应用等方面。农业数据采集,需运用互联网、物联网技术,尽可能收集异源及异构数据,多角度体验农业数据的全面性和可信性;农业数据存储,需结合分布式、云计算技术,达到低成本、低能耗及高可靠性的目标;农业数据处理,需将数据、文本、高维图像等类型数据降维后度量与处理,导出可理解的内容;农业数据分析,需结合降雨、气温、土壤状况与历年农作物产量等农业相关数据,运用大数据挖掘技术,预测农业未来产量,服务精准种植;农业数据应用,

结合农业产前、农业生产、农产品加工、流通及消费的特点,将大数据广泛应用到农业产业链的各个环节。

此外,农业大数据在政策法律、数据开放、数据共享及人才培养等方面都面临着严峻挑战,希望国家、企业及个人共同努力,为大数据在农业中的应用带来跨时代的变革。

6 结语

农业大数据首要任务是推动国家农业的发展,通过把握数据获取整理、存贮管理、计算处理、数据分析、知识展现与决策支撑等环节,努力走出一条生产技术先进、经营规模适度、市场竞争力强、生态环境可持续的中国特色新型农业现代化道路,在新一轮农业现代化建设中,紧抓机遇,夯实智慧农业的基石,让大数据创造出真正的智慧,支撑智慧农业的稳健发展。

参考文献

- [1] 张浩然,李中良,邹腾飞,等. 农业大数据综述[J]. 计算机科学,2014(S2):387-392.
- [2] 温平江. 农业大数据研究的战略意义与协同机制[J]. 高等农业教育,2013(11):3-6.
- [3] 孙忠富,褚金翔,杜克明,等. 农业大数据管理整个生命周期[J]. 高科技与产业化,2015(5):58-61.
- [4] 柳平增. 农业大数据平台在智慧农业中的应用:以渤海粮仓科技示范工程大数据平台为例[J]. 高科技与产业化,2015(5):68-71.
- [5] 许世卫. 农业大数据与农产品监测预警[J]. 中国农业科技导报,2014(5):14-20.
- [6] 郭承坤,刘延忠,陈英义,等. 发展农业大数据的主要问题及主要任务[J]. 安徽农业科学,2014(27):9642-9645.
- [7] 蔡书凯. 大数据与农业:现实挑战与对策[J]. 电子商务,2014(1):3-4.
- [8] 陈四龙,李玉荣,程增书,等. 用 GGE 双标图分析种植密度对高油花生生长和产量的影响[J]. 作物学报,2009,35(7):1328-1335.
- [9] 赵长星,邵长亮,王月福,等. 单粒精播模式下种植密度对花生群体生态特征及产量的影响[J]. 农学学报,2013,3(2):1-5.
- [10] 李安东,任卫国,王才斌,等. 花生单粒精播高产栽培培育特点及配套技术研究[J]. 花生学报,2004,33(2):17-22.
- [11] 郑亚萍,许婷婷,郑永美,等. 不同种植模式的花生单粒精播密度研究[J]. 亚热带农业研究,2012,8(2):82-84.
- [12] 冯焯,李宝龙,郭峰,等. 单粒精播对花生活性氧代谢、干物质积累和产量的影响[J]. 山东农业科学,2013,45(8):42-46.
- [13] 郭峰,万书波,王才斌,等. 不同类型花生单粒精播生长发育、光合性质的比较研究[J]. 花生学报,2008,37(4):18-21.
- [14] 王亮,王桥江,李艳,等. 天山北坡经济带机收花生膜下滴灌高产栽培技术[J]. 农业科技通讯,2015(6):238-240.
- [15] 万书波. 花生优质安全增效栽培理论与技术[M]. 北京:中国农业科学技术出版社,2009.
- [16] 甄志高,王晓林,段莹,等. 不同种植密度对花生产量的影响[J]. 中国农学通报,2004,20(2):90-91.
- [17] 高飞,翟志席,王铭伦. 密度对夏直播花生光合特性及产量的影响[J]. 中国农学通报,2011,27(9):320-323.
- [18] 于振文,岳寿松,沈成国,等. 不同密度对冬小麦开花后叶片衰老和粒重的影响[J]. 作物学报,1995,21(4):412-418.
- [19] HOFMANN W C, KITTOCK D L, ALEMAYEHU M. Planting seed density in relation to cotton emergence and yield[J]. Agron J, 1988, 80:834-836.
- [20] CAMPBELL T A, WHITE G A. Population density and planting date effects on Kenaf performance[J]. Agron J, 1982, 74:74-77.
- [21] 葛再伟,杨丽英. 不同种植密度对花生生育及产量的影响[J]. 花生学报,2002,31(3):33-35.

(上接第 382 页)