

大数据

农业

温孚江◎编著

BIG DATA
AGRICULTURE

农业大数据即是利用大数据的思维、方法以及技术来解决农业以及涉农行业的具体和客观问题的技术体系。本书以大数据思维为前提，从大数据概念、技术原理以及在农业中应用的案例或潜力入手，对农业大数据的应用进行了较为全面的阐述，借此来梳理农业大数据的广泛用途以及在农业现代化中的重要作用和重大价值，使更多人认识农业大数据，并投入到农业大数据的研究中来，使我国农业现代化插上“大数据的翅膀”，促进农业产生颠覆性的变革，与互联网农业相结合，促进我国早日实现农业现代化。

 中国农业出版社

大数据农业

温孚江 编著



中国农业出版社

图书在版编目 (CIP) 数据

大数据农业 / 温孚江编著. —北京: 中国农业出版社, 2015. 9 (2016. 3 重印)

ISBN 978-7-109-20910-7

I. ①大… II. ①温… III. ①农业现代化—研究—中国 IV. ①F320.1

中国版本图书馆 CIP 数据核字 (2015) 第 216536 号

中国农业出版社出版

(北京市朝阳区麦子店街 18 号楼)

(邮政编码 100125)

责任编辑 闫保荣 张 欣

北京中兴印刷有限公司印刷 新华书店北京发行所发行
2015 年 9 月第 1 版 2016 年 3 月北京第 2 次印刷

开本: 700mm×1000mm 1/16 印张: 21.5

字数: 380 千字

定价: 50.00 元

(凡本版图书出现印刷、装订错误, 请向出版社发行部调换)

自序

写作本书的动议来自 2014 年 7 月 18—20 日在山东省龙口市召开的“农业大数据论坛”。会议期间，来自复旦大学的数据科学家朱扬勇教授建议，既然大数据对农业如此重要，并且在中国已经掀起了农业大数据研究热潮，我们有责任写一本农业大数据的专著，研究农业大数据技术和应用模式，探寻农业大数据促进农业现代化的方式方法，同时介绍我们在农业大数据方面的成果和经验，以期推动农业大数据的研究与发展。几位与会专家一拍即合，接着组建了由温孚江、朱扬勇、姜春玲、米庆华、王云诚、陈盛伟、谢楚鹏为成员的编委会，并召开第一次编委会会议。此乃本书孕育、诞生的来由。

何为大数据？这个概念好像还没有一个各方都能接受的定义，但不影响大数据的应用。全面系统地掌握数据，使之具有“立体感”，让复杂的数据说话，提升生产生活方式，为人类带来福祉，这是大数据的魅力。大数据是全球生产力提升的前沿领域，蕴含巨大的应用价值和广泛的市场机会；大数据技术将创造新一轮产业格局，促进信息消费和“互联网+”经济发展，带动经济社会的深刻变革。互联网的巨大功能正在改变着人们的工作和生活，互联网和任何事物结合，往往产生 $1+1>2$ 的效果，这就是“互联网+”。“互联网+”实际上是互联网连接，再加以数据驱动。“互联网+农业”，可以称之为“互联网农业”，它大大充实了“信息农业”的概念，使信息农业的概念提升到一个更高的水平，内涵发生了巨大变化，如智慧农业，农产品电子商务等。大数据的概念在更深层次、更高水平上丰富了信息农业的内涵，对农业的影响将大大超过互联网农业。笔者认为，下一个更高级的产业变革模式可

能是“大数据+”模式。大数据和农业的深度融合可表述为“大数据+农业”，这是农业现代化的加速器，大数据改变了人们的思维模式和生活方式，不管你认同不认同，也不管你是否感觉得到，大数据几乎无处不在，你不知不觉中已经生活在大数据中。

大数据时代带给我们的是一种全新的思维方式，这就是大数据思维方式。大数据思维方式包含以下三层意思：简单地说，就是更多，更乱，关联。具体一点，一是分析的数据是全面的和多维度的（数据粒度密，数据种类多）；二是重视数据的复杂性，弱化精确性；三是关注数据的相关性，而非因果关系。全面系统地掌握数据，使之具有“立体感”，让复杂的数据说话。尽管以前没有“大数据”词句，但笔者认为，中国或许是最早应用大数据的国家。如中医，尽管“郎中”不知道每一种草药的有效成分是什么（每一种草药都含有繁多的物质和生物化学成分，即数据高“粒度”），但将特定的几种草药放在一起（“数据多维”），经过一定条件下的熬制（“数据分析和处理”），便成为一剂能治病的良药（“结果”）。其中任何一种草药单独都达不到预期的结果，只有“相关”的几种草药综合熬制才能成功，这是一种由量变到质变的过程。这完全符合上述所述的“大数据思维方式”的内涵。

大数据农业的应用研究还刚刚起步，尚无很多的成功案例，重要的是我们开始用大数据的思维来审视农业以及涉农问题，这意味着大数据农业时代已经到来。在大数据农业研究与应用中，数据的采集和挖掘是当前第一要务。上述提到的互联网农业的过程，也是一个数据积累和数据采集的过程。数据是资源，有价值。在互联网时代，数据是核心资源，但数据只有流动起来才能产生更大的价值，这也许是数据价值的属性。政府掌握着大量的农业大数据，但“数据孤岛”问题尚未解决。尽管有关数据的法律问题值得商榷和研究，但数据的流动是一种趋势，将来甚至会产生规范的“数据市场”，实现数据有序交易或共享。农业大数据需要各

种数据有序地流动起来，来推动农业现代化的实现，让数据产生不可估量的价值，推动生产力的发展，提高农业生产效益。

我国正在处于由传统农业向现代农业转变的转型期，农业将是互联网和大数据的下一个“风口”。“四化同步”中，农业现代化是其中一化。笔者认为，现代农业至少包括五个特征。一是规模化，这是首要。没有适当的规模化经营，就不可能走出小农经济的阴影，就谈不上效益和质量，更谈不上现代化。从国内外的形势看，目前是实现规模化的最佳时期，千载难逢，万万不可错过。二是机械化。机械化是规模化的前提，是提高效益，保证质量以及提高生产力的需要。这里所谈的机械化是比较高的，一般指高度智能化的机械化。三是标准化。标准化是保障农产品和涉农产品质量、生态保护、食品安全和品牌化的需要。四是信息化。一方面，农产品及涉农产品的供需双方在各个环节具备尽可能对称的市场信息；另一方面，高度信息化和数字化，可促进互联网农业的发展，这不但可提高农业生产率，还可将大数据应用到每一个领域。五是市场化。按照市场的性质，市场大致可分为“场内市场”和“场外市场”。场内市场是一种相对规范的期货市场平台，主要指农产品和涉农商品的期货商品交易。这种交易因具有平台（交易所），具有“规避风险”和“寻找价格”的功能而备受发达国家的青睐和高效使用。显然，我国在这方面还有很长路要走。我国现行的农产品和涉农产品的交易大量的场外交易，这种交易存在行为不规范和交易风险的弊端，缺少权威的，并具有信用约束功能的供需信息平台，因而农产品价格的波动时有发生，对国家的粮食安全和食品安全带来不利影响，亟待加强市场的监管。目前发展起来的农产品（以及农资）电商，是场外市场的一种新的表现形式。

分析可发现，现代化农业的这五个特征都离不开大数据的参与。大数据农业可助推农业现代化的实现。大数据农业即是利用大数据的思维、方法以及技术来解决农业以及涉农行业的具体和

宏观问题。农业大数据如何在农业现代化进程中起作用呢？农业大数据可服务农业的每一个环节，每一个领域。例如粮食安全问题、食品安全问题、现代育种及良种推广、土壤信息的变化与改良、生态农业的设计与发展、涉农产品市场监控及价格的预警、病虫害的预测、畜牧养殖业疫情的监控及预警预测，等等。大数据的主要功能是预测。归纳起来，大数据农业主要在三方面会发挥战略作用。一是指导生产和经营，为农业（广义农业）生产的产前、产中、产后提供全程服务，指导农业生产或涉农企业的经营。二是服务决策，为政府决策提供参考、咨询和指导服务，为企业生产、转型、市场营销提供咨询、指导。三是预测未来，如未来人口变化趋势和粮食供给能力，未来气候变化趋势对粮食生产影响，市场对某产品需求变化趋势等。农业数据的来源可以从互联网中获取，但主要的数据还是在农业生产和经营以及互联网农业中原创产生，这是大数据农业的一大特点。每一个数据都来之不易，值得倍加珍惜。农业大数据的另一个特点是其“准公益性”。因为农业是国家基础性战略产业，面对的是弱势的“三农”问题，这就决定了其具有公益性的一面。

本书以大数据思维为前提，从大数据概念、技术原理，以及在农业中应用的案例或潜力入手，对农业大数据的应用进行了较为全面的阐述，借此来梳理大数据的广泛用途以及在农业现代化中的重要作用和重大价值，使我国农业现代化插上“大数据的翅膀”，促进农业产生颠覆性的变革，促进我国早日实现农业现代化。

由于大数据农业尚处于探索阶段，书中难免留有很多遗憾和不足，希望读者们提出宝贵意见，以便我们不断充实和完善。

温孚江

2015年4月18日于东岳泰山

目 录

自序

第一章 绪论	1
第一节 大数据的概念与现状	1
一、数据及数据资源	1
二、大数据及其发展	2
三、大数据大变革	6
第二节 农业大数据	8
一、农业大数据的基本内涵	8
二、农业数据资源	9
三、农业大数据的开发应用	11
第三节 从传统农业到现代农业	12
一、现代农业的内涵	12
二、现代农业与传统农业	15
三、实现农业现代化的必要条件	16
第四节 农业大数据在现代农业中的作用	17
一、发展农业大数据的战略意义	17
二、大数据的国家战略	19
三、农业大数据应用	19
本章小结	20
本章参考文献	21
第二章 现代农业与大数据	23
第一节 农业规模化	23
一、农业规模化的内涵	23

二、我国的农业规模化	25
三、大数据在农业规模化发展中的应用	27
第二节 农业机械化	36
一、世界农业机械化发展概况	36
二、我国农业机械化发展历程与现状	37
三、我国农业机械化发展的新趋势与大数据应用	40
四、基于大数据技术的农业机械智能化	41
第三节 农业标准化	43
一、农业标准与农业标准化的内涵	43
二、国外农业标准化的基本特点与发展动态	44
三、我国农业标准化的发展历程与存在的问题	46
四、大数据推动我国农业标准化发展的思路	50
第四节 农业信息化	52
一、农业信息化内涵及其特征	52
二、我国农业信息化发展的进程与现状	53
三、大数据进一步推动农业信息化	56
第五节 农业市场化	59
一、我国农业市场化的发展进程	59
二、农业市场化面临的挑战	60
三、大数据推动农产品市场交易模式的创新	62
本章小结	65
本章参考文献	66
第三章 农业大数据的应用与开发	71
第一节 农业大数据资源的建设	71
一、农业大数据资源的内涵	71
二、农业本体与农业大数据资源建设	72
三、农业数据的获取途径与方法	75
四、农业数据的质量与清洗	78
五、农业大数据资源平台的建设	80
第二节 农业大数据在现代农业中的应用	82
一、农产品生产与大数据应用	83

二、农产品需求的大数据分析	87
三、农产品流通与大数据应用	90
四、农产品期货市场与大数据分析技术	94
第三节 农业大数据产业	97
一、大数据产业的兴起	97
二、国外农业大数据产业现状	98
三、我国农业大数据产业发展	99
本章小结	101
本章参考文献	102
第四章 农业领域数据资源	104
第一节 种植业数据资源	104
一、种植业概述	106
二、种植业数据类型与特点	106
三、种植业数据资源现状	109
四、种植业数据面临的问题	110
第二节 林业数据资源	111
一、林业概述	111
二、林业数据类型与特点	112
三、林业数据的获取方法	113
四、林业数据的管理	114
五、林业数据面临的问题	115
第三节 畜牧业数据资源	116
一、畜牧业概述	116
二、畜牧业数据类型与特点	116
三、畜牧业数据资源国内现状	117
四、畜牧业数据面临的问题	118
第四节 渔业数据资源	119
一、渔业与渔业数据信息化	119
二、水产养殖数据类型与特点	120
三、捕捞渔业数据类型与特点	120
四、渔业数据资源国内现状	121

五、渔业数据获取条件	122
第五节 农业水利资源数据资源	122
一、农业水资源与水利信息化	123
二、农业水利资源数据及信息化建设内容	123
三、农业水利资源数据类型与特点	124
四、农业水资源数据的获取	125
第六节 农产品加工数据资源	125
一、农产品加工概述	126
二、农产品加工数据类型与特点	126
三、农产品加工数据资源国内现状	127
四、农产品加工数据的获取	128
本章小结	129
本章参考文献	130
第五章 农业相关领域数据资源	131
第一节 生物基因数据资源	131
一、生物基因信息数据的重要性	132
二、生物基因数据资源	132
三、生物信息在农业生产中的作用	133
第二节 气候气象数据资源	134
一、气候气象数据	135
二、农业气象数据及其应用	135
第三节 地理数据资源	136
一、地理数据及获取	136
二、地理空间数据的获取	136
三、地理空间数据的统计	137
四、地理信息系统	137
五、地理信息在农业中的应用	138
第四节 农业生产资料数据资源	139
一、农业生产资料数据	139
二、农业生产资料数据的获取	139
三、农业生产资料数据的统计与应用	139

第五节 农产品物流与市场数据资源	140
一、农产品物流与市场数据采集	141
二、农产品物流与市场数据应用	143
第六节 国际农业数据资源	145
本章小结	148
本章参考文献	148
第六章 农业大数据应用技术体系	150
第一节 大数据采集预处理与存储管理	150
一、大数据采集与预处理方法技术	150
二、数据采集与预处理系统	152
三、从原始数据到高质量信息的采集和预处理技术	153
四、大数据存储与管理系统	155
五、发展趋势	159
第二节 大数据计算与挖掘分析	162
一、问题与挑战	162
二、主要进展	163
三、发展趋势	166
四、挖掘分析技术方法	168
五、问题与挑战	169
六、主要进展	170
七、发展趋势	172
第三节 大数据可视化分析	172
一、问题与挑战	172
二、主要进展	173
三、发展趋势与技术难题	175
第四节 大数据隐私与安全	177
一、问题与挑战	177
二、主要进展	179
三、发展趋势	180
本章小结	181
本章参考文献	182

第七章 国家粮食安全的监控与预测	184
第一节 粮食安全的概念与目标	184
一、粮食安全的有关概念	184
二、粮食安全的目标	186
第二节 世界粮食供需现状与挑战	188
一、世界粮食供应现状	188
二、世界粮食供应面临的挑战	190
第三节 国内粮食供需现状与问题	192
一、国内粮食供给现状	192
二、国内粮食需求现状	193
三、粮食生产的制约因素	196
第四节 粮食安全监控指标体系	201
一、食物的监控种类	201
二、食物供给监控指标	202
三、食物需求监控指标	204
第五节 粮食安全监控平台	206
一、食物供给数据采集平台	206
二、食物供给数据分析	206
三、粮食需求数据采集平台	210
四、粮食需求数据分析	210
五、粮食安全预警	215
本章小结	215
本章参考文献	216
第八章 农产品质量安全保障	218
第一节 农产品质量安全概述	218
一、农产品质量安全的基本概念	218
二、农产品质量安全保障存在的主要问题	219
第二节 农业生产条件安全控制	223
一、农产品生产条件安全控制的难点	223
二、农业生产中的质量安全数据采集与传递	226

三、大数据背景下的农业生产条件安全控制行为	229
第三节 农产品规范生产与加工控制	232
一、先进的生产与加工安全控制体系	232
二、农产品产业链中的数据采集与传递	233
三、大数据背景下的农产品产业链质量安全控制	235
第四节 农产品流通控制及产品追溯	236
一、农产品安全的追溯机制	236
二、农产品流通中的数据采集与传递	239
三、农产品质量安全的精准追溯	243
四、我国农产品安全追溯平台的发展现状	244
本章小结	248
本章参考文献	249
第九章 农业环境与生态安全	256
第一节 农业环境生态化	256
一、农业环境质量	256
二、农业环境数据平台	257
三、农业生态化	258
四、农业生产的环境污染问题及对策	262
第二节 农业灾害绿色防控	266
一、农业灾害	267
二、农业气象灾害	269
三、农业生物灾害	271
四、农业灾害防控的对策措施	273
第三节 精准农业与可持续发展	278
一、数字农业与精准农业	278
二、精准农业提高农业资源利用	282
本章小结	283
本章参考文献	284
第十章 有关大数据政策法规建设	289
第一节 大数据相关政策法规概览	289

一、大数据相关政策法规的界定	289
二、国际组织的相关立法	292
三、我国现行相关政策法规	295
四、发达国家逐渐推出大数据政策	297
第二节 数据权属政策法规	298
一、数据主权政策法规	298
二、数据权利及其政策法规	300
三、农业基础数据的产权	305
第三节 数据出版政策法规	308
一、数据出版	308
二、科学数据出版的现状	309
三、数据出版政策法规	311
第四节 数据流通与数据产业政策法规	315
一、数据流通	316
二、数据资产	319
三、数据产业	321
四、保护数据产业政策法规	324
本章小结	327
本章参考文献	328
后记	330

第一章 绪 论

大数据是数据科学的一个分支，是将数据科学原理与计算机科学、软件科学以及互联网科学相交叉融合的新兴学科。人类对客观世界的认识总是从感性认识开始，随着科学技术的发展，逐渐上升到理性认识，进而升华为科学。本章简要介绍大数据的基本概念和功能，并对大数据与农业现代化的关系进行简要论述。

第一节 大数据的概念与现状

我们经常遇到这种情况，天空浓云密布，各种因素似乎都符合降雨的条件，但雨就是下不来，干旱持续。在过去我们只能望天兴叹，但现在通过对影响降雨的因素分析我们知道，这可能是云中缺乏“凝结核”，这是水汽变成雨滴不可或缺的因素。据此，我们用飞机、火箭向云中播撒干冰、碘化银等作为凝结核，催化降雨的形成，这就是我们常见的人工降雨。其实，这就是通过调控改变事物发展轨迹的案例，是典型的大数据干预的结果。雨云的形成并形成降雨是由各种相关因素互相作用的结果，如太阳辐射强度、温度、大气湿度、风向、风速、气流对流强度、凝结核的形成等等，每一个因素的变化信息可理解为数据，这些因素（数据集）组合互相作用可理解为“大数据”。气象工作者根据这些因素的互相作用情况进行天气预测预报。

一、数据及数据资源

“数据”的含义既指 1062、2751 之类的阿拉伯数字，又指“dataology”（数据学），如“小舟扬帆出海”等的文字表述、“15/06/02”等符号、字符、日期的表达形式。总而言之，数据是指能够输入到网络空间的数字、文字报告、字符、声音、图表、照片等元素。

这里所指的网络空间（Cyber Space）主要是指互联网、物联网、广播电视网络、终端通信网络、卫星网络等所有人工智能网络和设备构成的空

间,这个空间真实存在。信息化的本质是将现实世界中的事物转化成数据并存储到网络空间中,即信息化是一个数据生产加工的过程,网络空间中的所有数据构成数据界(Datanature)。

从另一角度讲,数据是指网络空间中唯一存在,即网络空间中的任何东西,是可度量的、可处理的、可观测的并占有空间的。数据可以分为两类,一类是表示现实事物的数据,称为现实数据,另一类则不表示现实事物,只在网络空间中存在,称为非现实数据或虚拟数据。

经过 20 多年的国民经济与社会信息化发展战略的实施,信息技术(IT)已经为大家所熟悉。当今一部智能手机的储存和处理数据能力相当于 20 世纪 80 年的一台台式计算机或更强。今天的工作、学习、生活无不依赖于信息技术。信息化为我们带来便利,我们已经不能退回到信息化之前的时代了,这是信息化的成就。

信息化带来的变化显而易见。信息化将我们过去手工做的事情转换成计算机来做,并且表现得更准确、方便、高效;将现实事物通过照相机或录像机等的摄像头、录音机装置、传感器等采集到计算机中,进行信息的储存和处理,已成为常规知识,这些信息统称为“数据”。信息化的结果是在计算机系统中不断形成数据,这就是为什么我们要不断地买存储系统、买硬盘、买光盘、买优盘,不断地做备份,不断地保安全,目的是保存好信息化的成果,保存好我们的工作成果,保存好我们值得纪念的东西,等等。因此,从网络空间的视角来看,信息化的本质是生产数据的过程。数据作为一种资源已经获得广泛认识。数据资源被定义为有含义的数据集结到一定规模后形成数据资源。

二、大数据及其发展

(一) 什么是大数据

什么是大数据呢?目前,大数据的概念还没有一个统一的定义。下面是几个典型的大数据定义:

(1) Michael Cox 和 David Ellsworth 定义。1997 年,Michael Cox 和 David Ellsworth 在提出“大数据”术语时指出数据大到内存、本地磁盘、甚至远程磁盘都不能处理,这类数据可视化的问题称为大数据。

(2) 香山科学会议定义。2013 年,在北京香山召开的第 462 次香山科学会议上,对数据科学与大数据的科学原理与发展前景给出了两个定义,即