



Big Data, Little Data, No Data  
Scholarship in the Networked World

大数据管理丛书

# 大数据、小数据、无数据

## 网络世界的学术

[美] 克莉丝汀 L. 伯格曼 (Christine L. Borgman) 著  
孟小峰 张玮 赵尔平 译



机械工业出版社  
China Machine Press



大/数/据/管/理/丛/书

Big Data, Little Data, No Data  
Scholarship in the Networked World<sup>1</sup>

# 大数据、小数据、无数据

## 网络世界的学术

[美] 克莉丝汀 L. 伯格曼 (Christine L. Borgman) 著  
孟小峰 张祎 赵尔平 译



机械工业出版社  
China Machine Press

## 图书在版编目 (CIP) 数据

大数据、小数据、无数据: 网络世界的学术 / (美) 克莉丝汀 L. 伯格曼 (Christine L. Borgman) 著; 孟小峰, 张祎, 赵尔平译. —北京: 机械工业出版社, 2017.7  
(大数据管理丛书)

书名原文: Big Data, Little Data, No Data: Scholarship in the Networked World

ISBN 978-7-111-57578-8

I. 大… II. ①克… ②孟… ③张… ④赵… III. 数据处理 IV. TP274

中国版本图书馆 CIP 数据核字 (2017) 第 174611 号

### 本书版权登记号: 图字: 01-2017-0493

Christine L. Borgman: Big Data, Little Data, No Data: Scholarship in the Networked World (ISBN 978-0-262-02856-1).

Original English language edition copyright © 2015 Christine L. Borgman.

Simplified Chinese translation copyright © 2017 by China Machine Press.

Simplified Chinese translation rights arranged with MIT Press through Bardon-Chinese Media Agency.

No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system, without permission, in writing, from the publisher.

All rights reserved.

本书中文简体字版由 MIT Press 通过 Bardon-Chinese Media Agency 授权机械工业出版社在中华人民共和国境内 (不包括香港、澳门特别行政区及台湾地区) 独家出版发行。未经出版者书面许可, 不得以任何方式抄袭、复制或节录本书中的任何部分。

本书立足于大数据背景, 深入探讨了学术界的数据现状, 并详细阐释了数据与学术之间的复杂关系。书中首先通过确定调查研究的前提——六项“挑战”——来吸引读者对学术数据的兴趣, 然后就自然科学、社会科学和人文学科三个领域中的数据实践进行案例分析, 最后阐述其研究成果对学术实践和研究政策的推动意义。作者提出, 要实现长期的数据管理和挖掘, 就要对知识基础设施进行大量投资, 这对学术研究的未来至关重要。

本书适合相关学者、研究人员、高校领导、出版商、图书馆和数据档案馆的工作人员以及投资决策者和政策制定者阅读。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 曲熠和静

责任校对: 李秋荣

印刷: 北京文昌阁彩色印刷有限责任公司

版次: 2017 年 8 月第 1 版第 1 次印刷

开本: 170mm × 242mm 1/16

印张: 18

书号: ISBN 978-7-111-57578-8

定价: 99.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzjsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东



当下大数据技术发展变化日新月异，大数据应用已经遍及工业和社会生活的方方面面，原有的数据管理理论体系与大数据产业应用之间的差距日益加大，而工业界对于大数据人才的需求却急剧增加。大数据专业人才的培养是新一轮科技较量的基础，高等院校承担着大数据人才培养的重任。因此大数据相关课程将逐渐成为国内高校计算机相关专业的重要课程。但大数据人才培养课程体系尚不尽如人意，多是已有课程的“冷拼盘”，顶多是加点“调料”，原材料没有新鲜感。现阶段无论多么新、多么好的人才培养计划，都只能在 20 世纪六七十年代编写的计算机知识体系上施教，无法把当下大数据带给我们的新思维、新知识传导给学生。

为此我们意识到，缺少基础性工作和原始积累，就难以培养符合工业界需要的大数据复合型和交叉型人才。因此急需在思维和理念方面进行转变，按大数据应用需求对现有的课程和知识体系进行延展和补充，加入新的可以因材施教的知识模块。我们肩负着大数据时代知识更新的使命，每一位学者都有责任和义务为此“增砖添瓦”。

在此背景下，我们策划和组织了这套大数据管理丛书，希望能够培养数据思维的理念，对原有数据管理知识体系进行完善和补充，面向新的技术热点，提出新的知识体系 / 知识点，拉近教材体系与大数据应用的距离，为应对现代技术带来的大数据领域的新问题和新挑战扫除障碍。我们相信，

假以时日，这些著作汇溪成河，必将对未来大数据人才培养起到“基石”的作用。

**丛书定位：**面向新形势下的大数据技术发展对人才培养提出的挑战，旨在为学术研究和人才培养提供可供参考的“基石”。虽然是一些不起眼的“砖头瓦块”，但可以为大数据人才培养积累可用的新模块（新素材），弥补原有知识体系与应用问题之前的鸿沟，力图为现有的数据管理知识查漏补缺，聚少成多，最终形成适应大数据技术发展和人才培养的知识体系和教材基础。

**丛书特点：**借鉴 Morgan & Claypool Publishers 出版的 Synthesis Lectures on Data Management，特色在于选题新颖，短小精湛。选题新颖即面向技术热点，弥补现有知识体系的漏洞和不足（或延伸或补充），内容涵盖大数据管理的理论、方法、技术等诸多方面。短小精湛则不求系统性和完备性，但每本书要自成知识体系，重在阐述基本问题和方法，并辅以例题说明，便于施教。

**丛书组织：**采用国际学术出版通行的主编负责制，为此特邀中国人民大学孟小峰教授（xfmeng@ruc.edu.cn）担任丛书主编，负责丛书的整体规划和选题。责任编辑为机械工业出版社华章分社姚蕾编辑（yaolei@hzbook.com）。

当今数据洪流席卷全球，而中国正在努力从数据大国走向数据强国，大数据时代的知识更新和人才培养刻不容缓，虽然我们的力量有限，但聚少成多，积小致巨。因此，我们在设计本套丛书封面的时候，特意选择了清代苏州籍宫廷画家徐扬描绘苏州风物的巨幅长卷画作《姑苏繁华图》（原名《盛世滋生图》）作为底图以表达我们的美好愿景，每本书选取这幅巨卷的一部分，一步步见证和记录数据管理领域的学者在学术研究和工程应用中的探索和实践，最终形成适应大数据技术发展和人才培养的知识图谱，共同谱写出我们这个大数据时代的盛世华章。

在此期望有志于大数据人才培养并具有丰富理论和实践经验的学者和专业人员能够加入到这套书的编写工作中来，共同为中国大数据研究和人才培养贡献自己的智慧和力量，共筑属于我们自己的“时代记忆”。欢迎读者对我们的出版工作提出宝贵意见和建议。

## 大数据管理丛书

主编：孟小峰

### 大数据管理概论

孟小峰 编著

2017年5月

### 异构信息网络挖掘：原理和方法

[美] 孙芝洲 (Yizhou Sun) 韩家炜 (Jiawei Han) 著

段磊 朱敏 唐常杰 译

2017年5月

### 大规模元搜索引擎技术

[美] 孟卫一 (Weiyi Meng) 於德 (Clement T. Yu) 著

朱亮 译

2017年5月

### 大数据集成

[美] 董欣 (Xin Luna Dong) 戴夫士·斯里瓦斯塔瓦 (Divesh Srivastava) 著

王秋月 杜治娟 王硕 译

2017年5月

### 短文本数据理解

王仲远 编著

2017年5月

### 个人数据管理

李玉坤 孟小峰 编著

2017年5月

### 位置大数据隐私管理

潘晓 霍峥 孟小峰 编著

2017年5月

### 移动数据挖掘

连德富 张富峥 王英子 袁晶 谢幸 编著

2017年5月

云数据管理：挑战与机遇

[美] 迪卫艾肯特·阿格拉沃尔 (Divyakant Agrawal) 苏迪皮托·达斯 (Sudipto Das) 阿姆鲁·埃尔·阿巴迪 (Amr El Abbadi) 著

马友忠 孟小峰 译

2017年5月

大数据、小数据、无数据：网络世界的的数据学术

[美] 克莉丝汀 L. 伯格曼 (Christine L. Borgman) 著

孟小峰 张祎 赵尔平 译

2017年8月

## || 译者序

已故的图灵奖得主吉姆·格雷（Jim Gray）在其《事务处理》一书中提到：6000年前，苏美尔人（Sumerians）就使用了数据记录的方法，已知最早的数据是写在土块上的，上面记录着皇家税收、土地、谷物、牲畜、奴隶和黄金等情况。随着社会的进步和生产力的提高，类似土块的处理系统演变了数千年，经历了殷墟甲骨文、古埃及莎草纸、羊皮纸等阶段。19世纪后期，打孔卡片出现，并用于1890年的美国人口普查。用卡片取代土块，使得系统可以每秒查找或更新一个“卡片”（土块）。可见，用数据记录社会由来已久，而数据的多少和系统的能力是与当时社会结构的复杂程度和生产力水平密切相关的。

随着人类进入21世纪，尤其是互联网和移动互联网技术的发展，使得人与人之间的联系日益密切，社会结构日趋复杂，生产力水平得到极大提升，人类创造性活力得到充分释放。与之相应的数据规模和处理系统也发生了巨大改变，从而催涌了当下众人热议的大数据局面。

其实数据本无所谓大小，也没有高低贵贱之分，只有应用才是促使我们前行的动力。人们使用“数据”这一术语已经是第五个世纪了，但其至今仍没有一个公认的定义。克莉丝汀L. 伯格曼在此书中认为数据既不是客观事物，也不是带有自身特征的自然对象，数据只是对客观对象中某些可观测到的现象的描述，这些描述会因人而异、因地而异和因时而异。如何把客观事物概念化为数据，就是大数据的本质。

本书立足于大数据背景，深入探讨了学术界的数据现状，并详细阐释了数据与学术之间的复杂关系，最终提出了数据学术在政策与实践中的挑战。全书主要分为三部分。第一部分是数据与学术，主要介绍数据相关的基本概



念。其中，第1章描述了学术界“大数据”“小数据”和“无数据”三者并存的现状，“大数据”的特征，以及“小数据”甚至“无数据”的原因，并总结了大数据时代学术界面临的六大挑战。第2章全面介绍了“何为数据”这一问题，总结了不同的“数据”定义及其分类方式，并随之提出了本书中“数据”一词的含义；即出于研究或学术目的，一些现象需要观察结果、研究对象和其他实体承担论据的角色，而数据就是这些论据的表示方式。第3章在“知识基础设施”背景下，从社会与技术、开放学术、交流融合几个角度描述了数据与学术之间的复杂关系。第4章阐述了学术活动中数据的“多样性”特点，提出了数据大小问题，以及包括距离问题和外部影响因素在内的数据产生问题。该章内容奠定了第二部分中数据学术案例分析的基本框架。一般而言，学术可以分为自然科学、社会科学和人文学科三大领域，每个领域又囊括了不同学科。不同学科之间的数据量存在天壤之别。因此，第二部分从自然科学、社会科学和人文学科三个领域出发，在每个领域中分别选择了“大数据学科”和“小数据甚至无数据学科”的案例，从而进一步论证了数据与学术之间的复杂关系。基于前两部分内容，第三部分主要围绕三个问题展开：数据共享、发布与重用，数据信誉、归属与发现，保存何种数据及其原因。这一部分主要探索了数据政策与实践中的涉及的相关问题和严峻挑战。

“在中国传统文化中，最高的状态是意会的境界。大，意味着多。多，意味着无穷无尽，无穷无尽就是空。既无穷莫测，故实则虚之。实则虚之，是中国人的文化密码，投射到每个人的心中。”（引自《故宫100：至大无外》。）本书以数据为主线，将学术研究的三大学科贯穿起来，形成交叉学科的典范。希望本书能成为交叉学科研究的文化密码，投射到每个人的心中，打破数据的利益壁垒，构筑数据学术的大同世界。其实大数据的归宿是无穷无尽，无穷无尽就是空。

本书的翻译、统稿和审校由孟小峰组织完成。具体翻译分工如下：第1章由赵尔平翻译；第2章由张祎翻译；第3章由张祎、朱敏杰和杜治娟翻译；第4章由杜治娟和郭胜娜翻译；第5章由郭胜娜、李进、王春凯和吴文妹翻译；第6章由吴文妹、叶青青和杨晨翻译；第7章由杨晨、郝泽慧、任玮和郭豫龙翻译；第8章由郭豫龙、秦楷迪和翁祖建翻译；第9章由翁祖建、王硕和孙管阳翻译；第10章由孙管阳和忻日辉翻译。本书于2016年冬译出初稿，之后由孟小峰、张祎和赵尔平逐章进行了修改或重译，这期间三易其稿，最后由孟小峰负责统一定稿。

本书涉及自然科学、社会科学和人文学科三大领域的多个学科，内容丰富，术语量大，翻译工作的难度可想而知。本书术语主要遵从对应学科领域

中学学术论文的习惯用法。译者在翻译过程中备感力不从心，不当之处在所难免，恳请读者批评指正并不吝赐教。如果有任何建议或意见，欢迎发邮件至 [xfmeng@ruc.edu.cn](mailto:xfmeng@ruc.edu.cn)。

译者

2017年5月于北京

在“大数据”风靡的当代，学术研究中的“小数据”依旧不容小觑。随着绝对数据量的增加，学者们进行个体研究的能力却不断退化。因为他们从未距离研究对象如此遥远。新工具和新视角成为学术研究的必需品。但相对于“小数据”而言，“大数据”不一定是最好的选择。因为研究者距离数据来源（即数据采集、处理、提炼和转换方式及其对应的研究假设和目的）越远，就越难理解数据代表的实际意义。因此，学者们更倾向于使用少量数据进行更深入细致的研究。但当数据尚未被发现甚至不可能被发现时，学者们将面临“无数据”的窘境<sup>①</sup>。

相对于一般商品而言，研究类数据（research data）的利用价值或多或少。数据管理计划、数据发布需求以及由资助机构、学术期刊和科研机构提出的积极政策都很难适应数据多样性和跨领域的新实践。除了按例定义之外，很少有政策尝试给出数据的定义。那些反映学术界利益相关者竞争机制和动机的政策更是少之又少。数据的含义因人而异，因时不同；数据可以像资产一样受控、积累、交易、组合投资、挖掘甚至可能被发布；同时，数据也可以像债务一样被管理、保护或者毁约。数据的敏感性和保密性可能很高，因此数据发布往往伴随着高风险。数据价值可能立即显现，也可能很久之后才引人注目。有些数据值得长期监护，但大部分数据都只具有短期价值。某些研究类数据的价值可能在几个小时或几个月之内，随着技术和领域前沿工作的进步而不复存在。

要理解数据在学术研究中的价值，首先必须明确数据绝非一般事物，其不

① 这就是本书所说的“大数据、小数据、无数据”。——译者注

是具有本质内涵的自然对象。恰恰相反，出于研究或学术目的，学者们往往需要证明现象发生的论据，即观察结果、客观对象以及其他实体，而数据正是这些论据的表现形式。而且这种具体表现形式会因学者、情景以及时间的不同而有所差异。学者们虽然在自然科学、社会科学以及人文学科领域内创造、使用、分析和解释数据，但都尚未就数据定义达成共识。把某种事物概念化为数据的过程本身就是一种学术行为。学术活动包括寻求论据、解释现象以及论证观点等，这些活动以数据为媒介，最终往往表现为期刊论文、专著、会议论文以及学术界认可的其他形式。考虑数据重用的研究极为少见。

伽利略使用笔记本记录观察到的现象，19世纪的天文学家把图像刻在玻璃板上，而今天的天文学家则使用数字设备捕获光子。普通相机拍摄的夜空图像之所以能与航空任务采集的图像基本相似，是因为天文学家已经在数据描述和映射的表示体系上达成一致。为集成几个世纪以来的观测数据，天文学家已经在标准、工具和档案方面做了大量工作。即便如此，该领域的知识基础设施还远未建成，更不用说完全自动化了。就数据获取的组织和协调而言，信息专业技术人员在天文学等领域发挥着重要作用。

出版物和数据之间的联系体现在多个方面，这正是研究类数据在学术交流范围内得以充分利用的原因。数据创造可能需要长期的深思熟虑，这一漫长过程积累了许多随时间可增值的资源。但数据创造也可能历时很短，甚至具有偶发性。在后两种情况下，研究者需要尽可能记录现象发生时的所有仪器数据。在天文学、社会学和人种学等领域，不论研究计划制定得有多好，数据采集都具有随机性，因为每个阶段的研究发现都会影响下一阶段的数据选择。因此，任何领域的学者都需要学会评估数据优劣、判定数据信度和效度以及适应实验室研究、实地调查和档案馆的条件。出版物是特定领域中假设读者具有专业知识的条件下，对研究成果进行报告的形式，其为读者理解研究成果的论点、方法及结论提供了必要信息。但因为假定读者熟悉该领域，所以其往往省略重复研究所需的细节。虽然人们对数据发布各执己见，但数据的重复性与再现性只在特定领域具有相关性。而且即使是在这些领域，真正实现重复与再现也很困难。而判断哪个学术研究成果更具有保存价值更是难上加难。

数据管理、发布和共享等相关政策模糊了数据在学术活动中的复杂作用，而且在很大程度上忽略了领域内与领域间政策实践中的差异性。“数据”的概念在自然科学、社会科学和人文学科之间甚至同一学科内都存在很大差异。数据管理在很多领域内可学不可教，这使得现有解决方案缺乏系统性。一般而言，研究人员重用数据都有很大困难，更不用说使这些数据为具有不同需求的其他人所用。由于数据共享实施难度大、激励机制有限，而且在知识基

基础设施工程建设方面需要大量投资，所以其规范仅存在于部分领域。

本书面向包括研究类数据利益相关者在内的广大读者，具体包括学者、研究人员、高校领导、出版商、图书馆和数据档案馆的工作人员以及投资决策者和政策制定者等。第一部分使用四章篇幅界定“数据”与“学术”的概念，详细讨论了数据、学术、知识基础设施以及研究实践的多样性。第二部分包括三章，分别探究了自然科学、社会科学和人文学科领域的数据学术（data scholarship）。这三个案例分析结构平行，便于进行跨领域对比。第三部分涵盖三个章节，主要阐述数据策略和数据实践相关内容，并尝试分析数据学术面临众多棘手问题的原因。这部分具体包括：数据共享、发布与重用，数据信誉、归属与发现，保存何种数据及其原因。

“学术”与“数据”之间有着长远而错综复杂的关系，二者都不是新概念。当前，二者之间的关系又出现了新要素，即从学术活动中提取数据进行挖掘，并应用于其他目的。研究类数据相关的成本、收益、风险和报酬等正在相互竞争的利益相关者之间进行重新分配。本书旨在推动各方进行内容更全面、信息更充分的讨论。这对学术的未来至关重要。

克莉丝汀 L. 伯格曼  
加利福尼亚，洛杉矶  
2014 年 5 月



一本书的完成往往需要群策群力，这本多主题多领域的专著更是如此。此书受到各领域同事的帮助，这些帮助在学术用语中被称为“无形学院”。这些同事对所在领域的数据怀有满腔热情，他们在本书完成过程中花费大量时间不断讨论，并加入我们的专题讨论和研习会，还帮忙检查修订本书部分章节的内容。

本书的创作灵感产生于多年前，这里特别感谢那些曾影响我的思想，甚至在某种程度上直接帮我完成本书的人。这些人中的大部分都可以在参考文献中看到。虽然也有很多人在书中的部分话题讨论中给我留下了深刻印象，但我也可能会不可避免地遗漏某些人。

我对学术数据实践的研究可以追溯到 20 世纪 90 年代后期，具体包括数字图书馆前期建设工作、信息寻求行为、人机交互、信息检索、文献计量学和学术交流。我在数据实践研究过程中有幸结识了许多优秀的合作伙伴，他们极大地帮助我把已有的零散隐性知识整合为本书中的显性知识。我们的合作成果已得到广泛引用。合作伙伴中的许多教师、学生和博士后研究员都参与了很多项目，这里按字母顺序逐一列出。学术数据实践的研究项目包括：亚历山大数字地球原型项目（Alexandria Digital Earth Prototype Project, ADEPT）；传感器网络研究中心（Center for Embedded Networked Sensing, CENS）；网络学习任务小组（Cyberlearning Task Force）；监控、建模和存储（Monitoring, Modeling, and Memory）；数据管理（Data Conservancy, DC）；知识基础设施（Knowledge Infrastructures）以及长尾研究（Long-Tail Research）。

参与这些项目的教师合作伙伴包括：丹尼尔·阿特金斯（Daniel Atkins）、杰弗里·博克尔（Geoffrey Bowker）、赛义德·乔杜里（Sayeed Choudhury）、保罗·戴维斯（Paul Davis）、蒂姆·迪劳罗（Tim DiLauro）、乔治·乔尔戈夫斯基（George Djorgovski）、保罗·爱德华兹（Paul Edwards）、诺埃尔·埃涅迪（Noel Enyedy）、黛博拉·埃斯特林（Deborah Estrin）、托马斯·芬霍尔特（Thomas Finholt）、伊恩·福斯特（Ian Foster）、詹姆斯·弗鲁（James Frew）、乔纳森·费纳（Jonathan Furner）、安妮·吉利兰（Anne Gilliland）、迈克尔·古德柴尔德（Michael Goodchild）、艾莉莎·古德曼（Alyssa Goodman）、马克·汉森（Mark Hansen）、托马斯·哈蒙（Thomas Harmon）、布莱恩·黑岛恩（Bryan Heidorn）、威廉·豪（William Howe）、史蒂文·杰克逊（Steven Jackson）、卡尔·凯塞尔曼（Carl Kesselman）、卡尔·莱格若（Carl Lagoze）、格里高利·利泽（Gregory Leazer）、玛丽·玛丽洛（Mary Marlino）、理查德·迈耶（Richard Mayer）、卡罗尔·帕尔默（Carole Palmer）、罗伊·波蒂（Roy Pea）、格里高利·博迪（Gregory Pottie）、艾伦·若尼（Allen Renear）、戴维·诺贝兹（David Ribes）、威廉·桑多瓦尔（William Sandoval）、特伦斯·史密斯（Terence Smith）、苏珊·莱·斯达（Susan Leigh Star）、亚历克斯·绍洛伊（Alex Szalay）、查尔斯·泰勒（Charles Taylor）和沙伦·特拉维克（Sharon Traweek）。参与这些项目的学生、博士后和研究人员合作伙伴包括：丽贝卡·卡明斯（Rebekah Cummings）、彼得·达奇（Peter Darch）、戴维·费伦（David Fearon）、里奇·加沙（Rich Gazan）、米莱娜·戈尔尚（Milena Golshan）、埃里克·格雷厄姆（Eric Graham）、戴维·格温（David Gwynn）、格雷戈·杰尼（Greg Janee）、伊莱恩·莱维亚（Elaine Levia）、雷切尔·曼德尔（Rachel Mandell）、马太·曼也卡（Matthew Mayernik）、斯达萨·米洛贾维奇（Stasa Milojevic）、阿尔贝托·波珀（Alberto Pepe）、伊丽莎白·罗兰多（Elizabeth Rolando）、阿什利·金沙（Ashley Sands）、卡蒂·希尔顿（Katie Shilton）、吉利安·沃利斯（Jillian Wallis）和劳拉·威郝兹（Laura Wynholds）。

2012~2013年，我在牛津大学休假时，完成了本书的大部分构思和写作工作。牛津大学的同事们是我知识和新思想的源泉，他们饶有兴致地回答了我的问题——“你眼中的数据是什么？”贝利奥尔学院盛情邀请我担任奥利弗·史密斯访问学者和讲师。同时，我也是牛津大学互联网研究所和牛津大学电子化研究中心的访问学者。各种高低层交谈使我开始深入思考数据相关的所有事物，包括佛教、宇宙学、但丁、基因组学、手性、纳米科技、教育学、经济学、古典名著、哲学、数学、医学、语言学、文学和计算学等。牛津大学的体制为一群素未谋面甚至跨领域的人提供了研讨平台。我永远对此次休假的东

道主心怀感激，他们是贝利奥尔学院院长德拉蒙德·博恩（Drummond Bone）先生、资深导师尼古拉·特罗特（Nicola Trott）、牛津大学互联网研究所的威廉·达顿（William Dutton）、牛津大学电子化研究中心的戴维·德洛雷（David de Roure）和牛津大学图书馆管理员莎拉·托马斯（Sarah Thomas）。我在牛津大学也遇到了很多给人灵感的真诚合作伙伴：科菲·艾格乌（Kofi Agawu）、马丁·伯顿（Martin Burton）、乔治（George）、卡梅拉·爱德华兹（Carmella Edwards）、帕纳伊斯·菲利普特普洛斯（Panagis Filippakopoulos）、玛丽娜·伊罗特卡（Marina Jirotko）、威尔·琼斯（Will Jones）、埃琳娜·隆巴尔迪（Elena Lombardi）、埃里克·迈耶（Eric Meyer）、康塞普西翁·纳瓦尔（Concepcion Naval）、彼得（Peter）、雪莉·诺索弗（Shirley Northover）、拉尔夫·施罗德（Ralph Schroeder）、安妮·特雷费森（Anne Trefethen）和斯特法诺·扎基尼（Stefano Zacchetti）。

同样对我有所启发（这些启发甚至比他们本人想象得还要多）的牛津同事还有威廉·巴福德（William Barford）、格兰特·布兰克（Grant Blank）、达姆·林恩·布林德利（Dame Lynne Brindley）、罗杰·卡什莫尔（Roger Cashmore）、艾恩·查默斯（Iain Chalmers）先生、卡罗尔·克拉克（Carol Clark）、道格拉斯·杜普里（Douglas Dupree）、蒂莫西·恩迪科特（Timothy Endicott）、戴维·厄尔多斯（David Erdos）、贝特朗·佛揣克斯（Bertrand Fauchaux）、詹姆士·福德（James Forder）、布莱恩·福斯特（Brian Foster）、约翰保罗·贾布里勒（JohnPaul Ghobrial）、安东尼·格雷厄姆（Anthony Graham）先生、莱斯利·格林（Leslie Green）、丹尼尔·格里姆利（Daniel Grimley）、基思·凯勒巴斯（Keith Hannabus）、克里斯托弗·欣奇克利夫（Christopher Hinchcliffe）、沃尔夫·霍斯特曼（Wolfram Horstmann）、圣姬·吉姆（Sunghee Kim）、唐娜·库尔茨（Donna Kurtz）、威尔·拉尼尔（Will Lanier）、克莉丝·林托特（Chris Lintott）、保罗·勒夫（Paul Luff）、布莱恩·马吉（Bryan Magee）、海伦·马吉茨（Helen Margetts）、菲利普·马歇尔（Philip Marshall）、阿什利·诺德（Ashley Nord）、多米尼克·奥布莱恩（Dominic O'Brien）、德莫特·奥黑尔（Dermot O'Hare）、理查德·奥文登（Richard Ovenden）、丹尼斯·诺布尔（Denis Noble）、谢默斯·佩里（Seamus Perry）、安德鲁·彭岑（Andrew Pontzen）、雷切尔·夸勒尔（Rachel Quarrell）、戴维·罗比（David Robey）、安娜·桑德（Anna Sander）、布鲁克·西蒙斯（Brooke Simmons）、罗布·辛普森（Rob Simpson）、谭金崇（JinChong Tan）、琳内特·泰勒（Linnet Taylor）、罗萨林德·托马斯（Rosalind Thomas）、尼克·特雷费森（Nick Trefethen）、戴维·维恩斯（David Vines）、丽莎·沃克（Lisa Walker）、戴

维·华莱士 (David Wallace)、杰米·华纳 (Jamie Warner)、弗雷德里克·威尔莫特·史密斯 (Frederick Wilmot-Smith) 和蒂莫西·威尔逊 (Timothy Wilson)。

同时,我也特别感谢在第5、6和7章的案例研究中给予重要帮助同事们。第5章的天文学案例在很大程度上依赖于哈佛-史密森天体物理中心 (Harvard-Smithsonian Center for Astrophysics) 的艾莉莎·古德曼 (Alyssa Goodman) 及其合作伙伴,包括阿尔贝托·艾克玛择 (Alberto Accomazzi)、梅尔塞·科洛萨斯 (Merce Crosas)、克莉丝·埃德曼 (Chris Erdmann)、迈克尔·库尔茨 (Michael Kurtz)、格斯·明奇 (Gus Muench) 和阿尔贝托·佩佩 (Alberto Pepe)。这一案例也借鉴了加利福尼亚大学洛杉矶分校科研团队的知识基础设施研究成果。案例的具体内容也受益于古德曼 (Goodman) 教授的多次修改以及阿尔贝托·艾克玛择 (Alberto Accomazzi)、克莉丝·林托特 (Chris Lintott)、迈克尔·库尔茨 (Michael Kurtz)、帕特里克·麦克雷 (Patrick McCray) 和布鲁克·西蒙斯 (Brooke Simmons) 等其他天文学家或天文历史学家的评论。天文学家乔治·乔尔戈夫斯基 (George Djorgovski)、菲尔·马歇尔 (Phil Marshall)、安德鲁·彭岑 (Andrew Pontzen) 和亚历克斯·绍洛伊 (Alex Szalay) 也在明确科学问题方面提供了帮助。第5章的传感器网络科学与技术案例借鉴了传感器网络研究中心之前发表的成果,对应书稿也经过了戴维·卡伦 (David Caron)、埃里克·格雷厄姆 (Eric Graham)、托马斯·哈蒙 (Thomas Harmon)、马太·曼也卡 (Matthew Mayernik) 和吉利安·沃利斯 (Jillian Wallis) 等合作伙伴和传感器网络研究中心科技研究人员的审阅。第6章以社会科学为例进行分析,其中第一个关于互联网的研究案例主要基于对牛津大学互联网研究所格兰特·布兰克 (Grant Blank)、科琳娜·迪詹纳罗 (Corinna di Gennaro)、威廉·达顿 (William Dutton)、埃里克·迈耶 (Eric Meyer) 和拉尔夫·施罗德 (Ralph Schroeder) 等研究员的访谈,同时,他们也很热情地对相应书稿提出修改意见。第二个案例是社会技术研究方面的。正如参考文献所示,此案例引用了我与合作伙伴们之前发表的成果,并由合作伙伴马太·曼也卡 (Matthew Mayernik) 和吉利安·沃利斯 (Jillian Wallis) 审稿。第7章的人文学科案例专门为此书所写。CLAROS 案例主要基于对牛津大学唐娜·库尔茨 (Donna Kurtz) 的访谈及其提供的资料,戴维·罗比 (David Robey) 和戴维·肖顿 (David Shotton) 也提供了进一步帮助。本书对比萨狮鹫 (Pisa Griffin) 的分析借鉴了对牛津大学彼得·诺索弗 (Peter Northover) 的访谈及其提供的资料,同时也参考了英国伦敦大学东方与非洲研究学院安娜·柯涛蒂尼 (Anna Contadini) 提供的资料。最后一个案例是关于佛教学术的,这部分内容的完成要感谢牛津大学从事佛教研究的斯特法诺·扎基尼 (Stefano