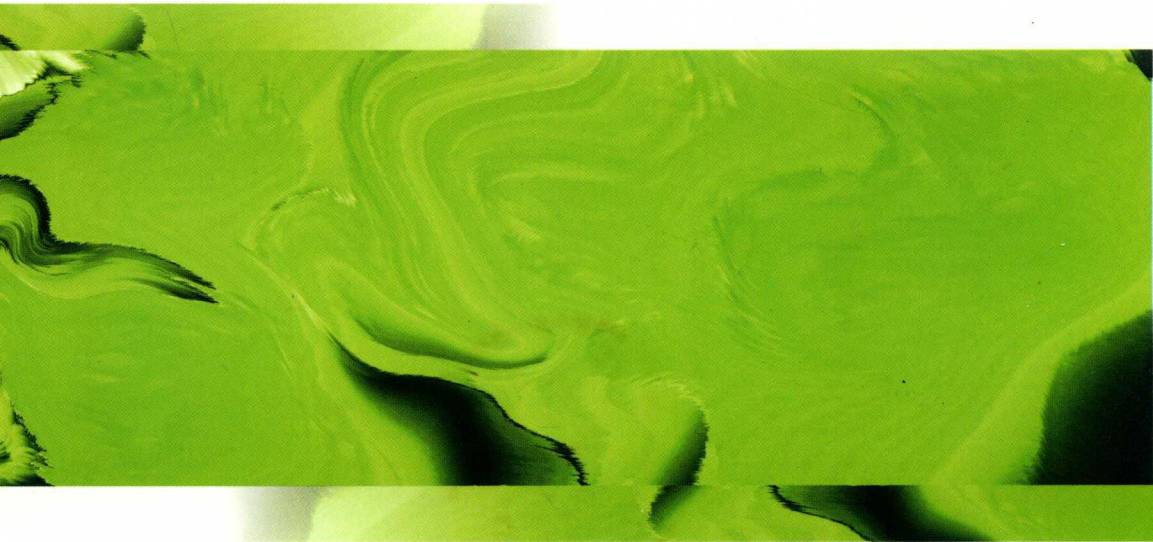


基于教育大数据的
教师专业成长丛书

王陆◎丛书主编

基于教育大数据的 知识发现方法与技术

JIYU JIAOYU DASHUJU DE
ZHISHI FAXIAN FANGFA YU JISHU



王陆 马如霞◎编著



北京师范大学出版集团
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP
北京师范大学出版社

图书在版编目(CIP)数据

基于教育大数据的知识发现方法与技术 / 王陆, 马如霞编著.
—北京: 北京师范大学出版社, 2019. 7
(基于教育大数据的教师专业成长丛书)
ISBN 978-7-303-23933-7

I. ①课… II. ①王… ②马… III. ①互连网络—应用—
教育方法—研究 IV. ①G4-39

中国版本图书馆 CIP 数据核字(2018)第 160345 号

出版发行: 北京师范大学出版社 www.bnup.com.cn

北京新街口外大街 19 号

邮政编码: 100875

印 刷: 天津旭非印刷有限公司

经 销: 全国新华书店

开 本: 730 mm×980 mm 1/16

印 张: 13.25

字 数: 180 千字

版 次: 2019 年 7 月第 1 版

印 次: 2019 年 7 月第 1 次印刷

定 价: 30.00 元

策划编辑: 林 子

责任编辑: 王星星

美术编辑: 焦 丽

装帧设计: 焦 丽

责任校对: 李云虎

责任印制: 陈 涛

版权所有 侵权必究

反盗版、侵权举报电话: 010-58800697

北京读者服务部电话: 010-58808104

外埠邮购电话: 010-58808083

本书如有印装质量问题, 请与印制管理部联系调换。

印制管理部电话: 010-58808284

北京出版集团
北京出版人
北京出版人

内容介绍

知识发现 (Knowledge Discovery) 源于人工智能和机器学习, 是机器学习、人工智能、数据库和知识库等众多研究领域相互融合而形成的具有交叉学科特色的新的研究领域。本书作为一本系统介绍基于教育大数据的知识发现的学术专著, 其阅读对象定位在从事教育技术领域研究的专业人员、教育技术学硕士研究生或博士研究生, 以及相关教育信息化企业的研究者。

本书作为教育部人文社会科学研究“教师在线实践社区中的知识管理与知识创新研究(13YJA880077)”项目资助的研究成果, 以教师在线实践社区(The Teacher's Online Communities Of Practice, 简称 TOCOP)为基本的研究情境, 采用科学研究第四范式(the Fourth Paradigm), 即数据密集型发现(Data-Intensive Scientific Discovery)范式, 以课堂教学行为大数据和实践性知识两类基础性教育大数据为例, 结合具体的研究案例, 为读者介绍了典型的基于教育大数据的知识发现方法与技术。

全书共分为六个章节。第一章是全书的背景知识, 主要介绍了以教师在线实践社区为背景的教育大数据及基于大数据的教育研究。第二章是大数据的知识发现方法与技术的基本框架, 包括相关的基本概念、过程与方法等内容。第三章至第五章是本书的核心部分, 以具体案例为依托, 详细介绍了课堂教学行为大数据的知识发现方法与技术, 优秀教师特质的知识发现方法与技术, 以及教师在线实践社区中的知识转移发现的方法与技术。第六章结合当前大数据知识发展的趋势和面临的挑战进行了总结与展望。

本书也可以作为教育技术学硕士研究生及博士研究生相关课程的教学参考用书。

丛书总序

2012年北京师范大学出版社出版了由我任总主编的远程校本研修丛书，包括《课堂观察方法与技术》《教学反思方法与技术》和《教师网络研修活动设计方法与技术》三本著作。这三本著作以2009年以来由我任课题负责人所做的教师在线实践社区项目（简称“靠谱 COP 项目”）和由我任首席专家的“国培计划（2011）中小学骨干教师研修项目”中的高中数学集中培训与远程校本研修一体化试点项目（简称“一体化项目”）的研究实践及研究案例为基础，勾画出了以课堂教学行为大数据为基础的教师混合式研修的新方法与新技术，出版7年来，该套丛书受到了广大读者的高度好评和喜爱。

时光荏苒，岁月如梭，转瞬已经过去了7年。随着科学技术的迅猛发展，信息时代教育治理的新模式，以及“互联网+”的人才培养模式正在成为研究的焦点，从应用切入深度融合的教育信息化2.0时代已经来临。在这7年中，我们以课堂为研究田野，我所领导的科研团队也通过在全国17个省、近300所学校，面向近1万名教师的深入课堂和深入学校的研究中，逐渐完善起基于课堂教学行为大数据的研究方法论，在原有远程校本研修丛书的基础上，进行了继承性创新，补充了大量的研究实例和研究数据，提出了更多的研究方法，形成了这套基于教育大数据的教师专业成长丛书，包括《基于课堂教学行为大数据的课堂观察方法与技术》《基于课堂教学行为大数据的教学反思方法与技术》《教师网络研修活动设计方法与技术》和《基于教育大数据的知识发现方法与技术》四本著作。

在这7年中，伴随我们在中小学开展的课堂教学行为大数据的深入研究，我们无时无刻不感受着古德和布罗菲所说的课堂之美：课堂是质朴的、守成的、思辨的、分析的、批判的、创新的、激昂的；课堂是思想生命的火花的碰撞与展现，是情不自禁从灵魂深处流露出的不断滋润精神之园的甘泉的发源地(Good & Brophy, 2017)。我们享受着研究带来的深层快乐，我们也深刻地感悟着课堂的复杂多变性：共时性、不可预料性、错综复杂性、非间接性和公开性，迎接着一个接一个研究难题的挑战。

虽然，大数据目前还没有一个全球公认的定义，作为这套丛书的总主编，我认为，大数据即海量的资料，指的是专业领域中所创造的大量非结构化和半结构化数据。大数据具有4个特点：大容量(Volume)、多样性(Variety)、高速度(Velocity)和多维价值(Value)(简称“4V”)。正如大数据的定义一样，课堂教学行为大数据目前也不存在一个公认的定义。但是，我想可以借用我们对大数据的认识，对课堂教学行为大数据做一个定义，课堂教学行为大数据是指，在课堂情境中，伴随教与学过程而产生的大规模、多样性，蕴含了丰富的教与学含义的非结构化与半结构化的特殊数据集。经过19年的课堂教学行为大数据的研究，我们发现，目前中小学课堂中的教学行为大数据共有4种类型：模式数据、关系数据、结构数据和行为数据。模式数据是指反映教学模式要素及要素之间关系的数据。关系数据是指反映课堂中行动者之间的相互关系结构的数据。结构数据是指反映为完成一定的教学目标，构成教学的诸因素在时间、空间方面所呈现的比较稳定的倾向及流程的数据。行为数据是指反映教与学行为主体特征的数据。

自从2000年我决定率领团队开始进行课堂教学行为大数据研究至今，已经过了整整19个年头。19年来，我们在深入中小学课堂，与中小学教师组成密切研究共同体的过程中，越来越清晰地感受到，只有当教师具体的教学行为在课堂教学行为大数据中清晰可见且被条分缕析时，只有当教师拥有了对课堂行为的描述和表达能力并建立起概念系统时，教师才会增

强其对课堂中所发生的所有事情的认识，也才真的会发生行动中反思和行动后反思。

课堂教学行为大数据是一种无形资产，是教师和学校专业发展的重要资源。2016年10月13日《光明日报》整版报道了我们的研究，并且在“编者按”中指出：大数据时代，来自课堂教学行为的大数据，不仅可以帮助我们清晰地认识不同教育发展水平的地区教师课堂教学行为的差异与特点，从而助力中等和薄弱地区的学校与教师通过改进课堂教学行为实现课堂教学质量的提高；同时，课堂教学行为大数据还可以促进优质教育区域更快地总结概括出课堂教学的优秀经验和优秀教师的实践性知识，从而实现教育优质资源在知识层面的共享与传播，助力教育均衡化发展。

正如《大数据时代》一书的作者所说，互联网世界的变化速度与日俱增，但万变之中有一点不曾变过，就是通过互联网，大数据将改变一切，可能超越我们所有人的想象。课堂教学行为大数据透视出的不仅仅是本套丛书中提到的各种现实问题，也为撬动教育供给侧的改革提供了思路。运用大数据分析方法与技术，寻找到教育教学中的真正短板，开发出面向广大中小学教师的专业发展公共服务，有效推进教师培训与研修的结构调整，矫正相关要素的配置扭曲，提高教师专业学习与培训研修的供给结构，有的放矢地实施供给侧改革是我们“靠谱 COP”团队的责任与愿景。我们坚信，课堂教学行为大数据将在教师教育供给侧改革中扮演越来越重要的角色。

本套丛书由首都师范大学博士生导师、“靠谱 COP”联盟首席专家王陆教授进行总体设计。同时，王陆教授与张敏霞副教授共同担任了《基于课堂教学行为大数据的课堂观察方法与技术》及《基于课堂教学行为大数据的教学反思方法与技术》两本书的作者，并负责《基于课堂教学行为大数据的教学反思方法与技术》一书的统稿工作；王陆教授与马如霞副教授一起担任了《基于教育大数据的知识发现方法与技术》一书的作者。张敏霞副教授负责《基于课堂教学行为大数据的课堂观察方法与技术》的统稿工作。首都师范

大学杨卉教授与冯涛副教授担任了《教师网络研修活动设计方法与技术》一书的作者。首都师范大学的硕士研究生房彬、刘霜和罗一萍同学参与了《基于课堂教学行为大数据的课堂观察方法与技术》一书的编写工作；中央电化教育馆的张静然副研究员及首都师范大学的硕士研究生张薇、刘文彬、马晔和林子同学参与了《基于课堂教学行为大数据的教学反思方法与技术》一书的编写工作，其中张静然同志还参与了《基于课堂教学行为大数据的教学反思方法与技术》一书的统稿工作；首都师范大学的硕士研究生张莉、耿雪和李爽同学参与了《教师网络研修活动设计方法与技术》一书的编写工作，北京优学社教育咨询服务有限公司的数据分析工程师彭功老师、首都师范大学的张敏霞副教授、首都师范大学的博士研究生张薇及三位硕士研究生李瑶、李旭和任艺参与了《基于教育大数据的知识发现方法与技术》部分撰写工作。

感谢在首都师范大学现代教育技术重点实验室做国内访问学者、来自内蒙古农业大学外语学院的陈金凤副教授，陈金凤副教授参加了本次丛书再版的修订工作，为本书的再版工作付出了智慧和辛勤的劳动。感谢北京优学社教育咨询服务有限公司对本套丛书修订工作的支持。

感谢首都师范大学教育技术系董乐老师、司治国老师对本套丛书撰写所提供技术上的各种支持与精神上的热情鼓励，感谢北京优学社教育咨询服务有限公司的工程师王鹏，有你们的无私陪伴与幕后奉献，才使得我们这个团队能够不断向前。感谢参与“国培计划(2011)”中小学骨干教师研修项目中的高中数学集中培训与远程校本研修一体化试点项目的全体老师，感谢参与“靠谱 COP”项目的全体老师，是你们的智慧贡献和全力投入才使得我们能拥有今天的成果与成绩，你们的课堂绩效改进和学生的进步一直是我们的最大心愿。也要衷心感谢参与“一体化项目”和“靠谱 COP”项目的全体助学者同学们，是你们的创造性工作才使得我们这个共同体生机勃勃，不断焕发出年轻生命的动人活力。

本书之所以能够顺利完成，还应该感谢北京师范大学出版社集团北京

京师普教文化传媒有限公司栾学东董事长的亲自指导和大力帮助，以及对本套丛书作者的关怀与理解。本套丛书的责任编辑林子，作为曾经担任过3年“靠谱 COP 项目”的助学者，满怀热情地投入本套丛书的策划与实施工作中，认真、细致、严谨地完成了书稿的各项编辑工作，使得本套丛书能够顺利出版。

本套丛书参考与引用了国内外大量的资料，其中的主要来源已在参考文献目录中列出，如有遗漏，恳请原谅。由于作者经验与学识所限，加上时间紧迫，书中谬误之处在所难免，欢迎读者指正。

王陆于北京

前 言

知识到底是什么，目前仍然存有争议。然而，著名管理学者彼得·德鲁克(Peter Drucker)早在 1965 年就预言：“知识将取代土地、劳动、资本与机器设备，成为最重要的生产因素。”目前，哲学视角的知识观认为“所谓知识，是客观事物的属性与联系的反映，是客观世界在人脑中的主观映像，可以分为感性知识和理性知识”。认识论视角的知识观认为，知识具有显性知识和隐性知识两种不同的知识种类。存在论视角的知识观认为，知识具有 4 个层次：个体知识、团队知识、组织知识和跨组织(interorganization)知识，其中个体知识是最根本的知识，是其他各个层次知识的基础。

教师的专业发展要以实现教师的实践性知识发展和课堂教学实践行为的改进为目标。为此，教师网络研修活动必须为发展教师的实践性知识和课堂教学实践行为的改进提供必要的支持与助力。由于教师的实践性知识以隐性知识为主，不能以正规的形式加以传递，而是需要通过非正式途径进行知识共享和知识创新而获得，因此，教师在线实践社区(The Teacher's Online Communities of Practice, 简称靠谱 COP)作为一般传统教师网络研修的补充和发展便应运而生了。

知识具有波粒二象性。当我们在对知识进行分类、组织甚至测度时，知识具有实体的性质，并且具有“粒子”的属性；而在对知识进行创造、提高及应用的持续过程中，知识又具有了过程的性质，即“波”的属性。知识的波粒二象性并不是相互对立的，而是相互依存和统一的。作为实体的知

识粒子是作为过程知识波的内容、载体和对象，因此，没有作为粒子的知识实体存在，就不会有可以传播流动的知识波，也就不会发生知识的传播、转移、共享与创造。同样，如果没有作为过程的知识波存在，作为实体的知识也就无法发挥作用，同样也不会发生知识的传播、转移、共享与创造。

一般的教师网络研修以教师的学科取向的显性知识发展为主，注重知识作为“粒子”的实体属性，即注重将庞大的教师知识体系逐一分解为最小的不可再分的知识粒子，也就是知识点，并由此利用知识的粒子特征及其可继承性的特点，设计一系列以识记、理解、应用、分析、评估为目标的网络学习活动。一般作为教师网络研修的培训专家或在线主讲教师及在线辅导教师，在教师网络研修中所做的大量工作是针对已经脱离教师思维空间，并以其他形式运动的知识，即静态知识，进行知识内容的管理工作。教师在线实践社区与一般的教师网络研修不同，它更注重知识的“波”属性，是以教师的实践性知识这种隐性知识为核心的。因而，教师在线实践社区不仅为知识波的流动提供了“知识收集→知识编码→知识转移与扩散→知识共享与交流→知识创新”等由一系列互不相同但又相互关联的研修活动所构成的知识价值链，还提供了基于知识价值链的、连接教师不同群体与外部组织知识的，如大学的教师教育专家等外部组织的知识、经验和技能等，最终形成了复杂的教师知识网络。故教师在线实践社区不仅可以汇聚大量的教师课堂教学行为大数据和教师实践性知识大数据，还可以促使教师的实践性知识迅速增长，从而涌现出众多的知识创新成果。

伴随着教师在线实践社区中教师课堂教学行为大数据和实践性知识大数据迅猛增长，传统的网络研修平台的信息服务体制和服务方式已经完全不能满足教师的知识管理需求，迫切需要以数据挖掘和数据分析为基础，通过现代技术手段将资源整合、知识发现、信息推送融为一体的知识发现服务。为此，本书所涉及的众多研究案例就一一诞生了。

本书由首都师范大学博士生导师、靠谱 COP 联盟首席专家王陆教授进行总体设计。王陆教授同时还与马如霞博士共同担任主编，负责全书的统

稿工作，同时完成了第一章、第三章、第四章的书稿撰写。马如霞博士与彭功老师共同撰写了本书的第二章和第六章的内容。张敏霞副教授撰写了本书第五章的内容。张薇博士研究生及李瑶、李旭和任艺三位硕士研究生分别参与了第一章和第四章部分内容的撰写工作。

感谢多年来担任教师在线实践社区项目指导教师的杨卉教授、冯红副教授、冯涛副教授、王彩霞副教授、董乐老师及司治国老师等，以及靠谱COP联盟的陈莉理事长、王鹏老师和赵炜老师等对本书撰写所提供的技术上的各种支持与精神上的热情鼓励，因为有你们的无私陪伴与奉献，靠谱COP团队才能够不断向前。感谢来自全国15个省近300多所参与教师在线实践社区项目的中小学的研修教师，是你们多年来的知识与智慧的贡献，才使得我们能够拥有今天的研究成果，你们的实践性知识增长和课堂教学实践行为的改进是我们最大的期盼与心愿。也要衷心感谢参与教师在线实践社区项目的全体同学，是你们创造性的工作，才使得我们这个学习共同体生机勃勃，不断焕发出年轻生命的动人活力。

本书之所以能够顺利完成，还应该感谢北京师范大学出版社的栾学东副总编辑的亲自指导和大力帮助。本书的策划编辑林子，作为曾经担任过三年教师在线实践社区项目的助学者，满怀热情地投入本书的策划中，认真、细致、严谨地完成了书稿的各项编辑工作，使得本书能够顺利出版。

本书参考并引用了国内外大量的资料，其中的主要来源已在参考文献目录中列出，如有遗漏，恳请原谅。由于作者经验与学识所限，加上时间紧迫，书中谬误之处在所难免，欢迎读者指正。

王陆

于北京

2017年11月22日

目 录

第 1 章 教师在线实践社区中的大数据

1.1 教师在线实践社区 / 1

1.1.1 教师在线实践社区的内涵 / 1

1.1.2 教师在线实践社区的特征 / 3

1.2 教育大数据 / 9

1.2.1 教育大数据的内涵 / 9

1.2.2 大数据时代教育的变革 / 12

1.2.3 教育大数据的价值与应用领域 / 14

1.2.4 教师在线实践社区中的大数据 / 18

1.3 基于大数据的第四范式引领的教育研究 / 21

1.3.1 第四范式引发科学研究方法的革命 / 21

1.3.2 教育研究迈向第四范式的必然趋势 / 25

【本章参考文献】 / 28

第 2 章 基于大数据的知识发现

2.1 基于大数据的知识发现的过程与方法 / 36

2.1.1 基于大数据的知识发现：连接“数据”和“知识”的桥梁 / 36

2.1.2 DIKW(数据—信息—知识—智慧)模型 / 40

2.2 基于大数据的知识发现数据处理 / 43

2.2.1 基于大数据的知识发现数据采集 / 43

2.2.2 基于大数据的知识发现与数据挖掘方法 / 47

2.2.3 基于大数据的知识发现知识重组过程 / 54

2.3 教师在线实践社区中的大数据知识发现服务 / 58

2.3.1 教师在线实践社区中知识发现服务的过程模型 / 59

2.3.2 教师在线实践社区知识发现服务类型 / 61

【本章参考文献】 / 67

第3章 课堂教学行为大数据的知识发现

3.1 课堂教学行为分析与研究 / 73

3.2 提问行为倾向的知识发现 / 76

3.2.1 研究方法 / 77

3.2.2 研究案例 / 79

3.3 教学现象与教学行为的知识发现 / 83

3.3.1 研究方法 / 83

3.3.2 研究案例 / 84

【本章参考文献】 / 95

第4章 优秀教师的特质发现

4.1 优秀教师的大数据研究 / 99

4.1.1 优秀教师及优秀教师特质 / 99

4.1.2 基于课堂教学行为大数据的知识发现 / 101

4.2 优秀教师的实践性知识特征发现 / 104

4.2.1 研究方法 / 105

4.2.2 研究案例 / 107

4.3 优秀教师课堂教学行为特征发现 / 120

4.3.1 研究方法 / 121

4.3.2 研究案例 / 122

4.4 课堂教学行为与实践性知识相关性的发现 / 130

4.4.1 研究方法 / 130

4.4.2 研究案例 / 131

【本章参考文献】 / 136

第5章 教师在线实践社区中的知识转移发现

5.1 知识转移研究 / 139

5.1.1 知识转移的内涵 / 139

5.1.2 知识转移的要素 / 141

5.1.3 知识转移过程模型 / 143

5.1.4 教师在线实践社区中知识转移的分类 / 145

5.1.5 教师在线实践社区中的非正式隐性知识转移 / 148

5.2 教师在线实践社区中的知识转移网络发现 / 152

5.2.1 研究方法 / 152

5.2.2 研究案例 / 154

5.3 教师在线实践社区中的知识转移影响因素研究 / 166

5.3.1 研究方法 / 167

5.3.2 研究案例 / 170

【本章参考文献】 / 177

第6章 大数据的知识发现展望

6.1 基于大数据的知识发现面临多重挑战 / 182

6.1.1 大数据的定义与存储规范缺乏统一标准 / 182

6.1.2 非结构化数据处理技术的复杂性 / 183

6.1.3 大数据存储与数据冗余问题 / 184

6.1.4 数据质量问题 / 185

- 6.1.5 数据挖掘技术面临的挑战 / 185
- 6.1.6 知识发现中组织与机构的协作问题 / 186
- 6.1.7 数据安全与隐私保护的问题 / 187

6.2 基于大数据的知识发展的发展趋势 / 188

- 6.2.1 基于大数据的知识发现相关技术的发展趋势 / 188
- 6.2.2 大数据知识发现促进知识生产的变革 / 190
- 6.2.3 基于大数据的知识发现服务应用领域的细分 / 191
- 6.2.4 基于大数据的知识发现中的信息安全问题将成为关键 / 191

【本章参考文献】 / 192

第4章 优秀教师的教育大数据研究

4.1 优秀教师的教育大数据研究 / 99

4.1.1 优秀教师及优秀教师特质 / 99

4.1.2 基于课堂教学行为大数据的教师研究 / 101

4.2 优秀教师的实践性知识研究 / 105

4.2.1 研究方法 / 105

4.2.2 问题提出 / 105

4.3 优秀教师的实践性知识研究 / 105

4.3.1 研究背景 / 105

4.3.2 研究案例 / 105

第 1 章 教师在线实践社区中的大数据

目前世界范围内对教师专业学习的研究领域仅仅处于起步阶段，但在最近大约 20 年的时间里，已经取得了巨大的进展，并且已有证据显示教师的专业学习与专业化发展能够有助于提高教师的教学实践和学生的学习。教师在线实践社区作为一种教师专业学习的理想学习环境，日益蓬勃发展，并引起众多研究者的高度关注。在从事教师在线实践社区的理论与实践的 18 年中，作者团队一直认为：教师在线实践社区是由教师、专家及助学者所组成的一种正式学习与非正式学习相混合的学习系统，是一种由专业人员提供基于大数据的专业知识发现服务的学习型组织，也是一种聚焦教师知识的新的知识管理与知识创生的新途径。

1.1 教师在线实践社区

1.1.1 教师在线实践社区的内涵

教师在线实践社区的最初概念是解决发生在面对面情况下的学习问题。老师在线实践社区是一种支持成人通过日常的社会实践学习而不是聚焦有意设计的课程来支持学习的学习环境，是一种正式学习与非正式学习相混合的系统(Wenger, 1998; Wenger, 2001)。在过去的 10 年中，由于实践社区在教师专业发展中具有独特的不可替代性，所以许多国家都在教育改