

# 人工智能 与 价值观

Artificial Intelligence and Values

李仁涵 黄庆桥 等 编著



上海交通大学出版社  
SHANGHAI JIAO TONG UNIVERSITY PRESS

# 人工智能 与价值观

Artificial Intelligence and Values

李仁涵 黄庆桥 等 编著



上海交通大学出版社  
SHANGHAI JIAO TONG UNIVERSITY PRESS

## 内容提要

人工智能正在以润物细无声的态势,进入各个产业领域和社会方面,预示着科技与经济、社会、生活融合的未来景象。深入了解人工智能的应用情况,准确评估其对社会道德价值体系产生的影响,将是我国应对下一个快速发展时期的重要手段,以便明确前进方向,实现精细治理,维护社会价值体系在科技大潮中的稳定和可持续发展。

## 图书在版编目(CIP)数据

人工智能与价值观/李仁涵等编著. —上海:上海交通大学出版社, 2021

ISBN 978-7-313-24957-9

I. ①人… II. ①李… III. ①人工智能 IV. ①TP18

中国版本图书馆 CIP 数据核字(2021)第 092254 号

## 人工智能与价值观

RENGONGZHINENG YU JIAZHIGUAN

编 著: 李仁涵 黄庆桥 等

出版发行: 上海交通大学出版社

邮政编码: 200030

印 制: 上海新艺印刷有限公司

开 本: 880mm×1230mm 1/32

字 数: 139 千字

版 次: 2021 年 6 月第 1 版

书 号: ISBN 978-7-313-24957-9

定 价: 68.00 元

地 址: 上海市番禺路 951 号

电 话: 021-64071208

经 销: 全国新华书店

印 张: 7.125

印 次: 2021 年 6 月第 1 次印刷

版权所有 侵权必究

告读者: 如发现本书有印装质量问题请与印刷厂质量科联系

联系电话: 021-33854186

# 编 委 会

主 任 李仁涵 黄庆桥  
成 员 杨小康 何黎明 沈辛成 叶 璐  
翟云杰 周 正 李芳薇 黄蕾宇

## 章节分工

第一、二章：李芳薇（北京大学）

第三章：沈辛成（上海交通大学）

第四章：叶璐、翟云杰（上海交通大学）

第五章：周正（上海交通大学）

第六章：黄庆桥、黄蕾宇（上海交通大学）

后记：李仁涵、沈辛成（上海交通大学）

# 序

2015—2020年，人工智能呈现爆发性增长，已开始“嵌入”到人们生产和生活之中，美国、欧盟、英国、日本、韩国、东盟、中国等国家和地区，纷纷加强顶层设计，成立专门机构统筹推进人工智能重大科技研发项目。目前看，人工智能已从自由探索的科研模式转型升级为国家战略推动创新发展的模式。各国政府、大学、科研机构、产业、金融界等竞相争夺该领域的前沿阵地，不同层面的布局都指向同一个预判：在未来数十年间，人工智能有着改变全球社会的巨大潜力。

2018年1月，上海交通大学人工智能研究院成立并揭牌，同年2月，获批建设人工智能教育部重点实验室。2019年，上海交通大学成为全国首批人工智能专业建设高校；在研究生教育层面，上海交通大学设立了吴文俊人工智能荣誉博士班，着重培养人工智能领域领军人才。与此同时，上海交通大学坚持“扎根上海、辐射长三角、服务全国”的宗旨，为中国人工智能产业的发展提供坚实的智力支持。上海交通大学牵头共建的上海人工智能研究院有限公司，正实现平台实体化运行，并与上

上海交通大学苏州人工智能研究院和上海交通大学宁波人工智能研究院共同完成“上海—宁波—苏州长三角人工智能协同创新平台”布局，联合探索人工智能校地之间“政产学研用金”的合作。

三年来，上海交通大学不仅在人工智能理论、技术、应用及人才培养等方面谋求发展与进步，同时也高度重视人工智能带来的社会层面问题的研究，《人工智能与价值观》一书就是其中的一项研究成果。该书在编写过程中，以历史眼光和全球视角审视人工智能的发展，资料翔实、文风简朴，秉持理性、客观、严谨的学术研究态度，通过对技术革命中的价值观演变、人工智能发展带来价值观的重塑、人工智能与美国价值观、人工智能与欧洲价值观、人工智能与日本价值观、人工智能与中国价值观等方面的研究，提出了重视人工智能发展的国家需要团结协作，求同存异，寻求最大的价值公约数，是关乎到人类集体共有的未来；提出了全球在人工智能发展的同时，高度共同关注以人为本、和而不同、人类命运共同体等问题。

当前，由于人工智能发展处于初级阶段，该书在编撰中尚存在不足、疏漏和错误之处，希望编委们在今后能够持续跟踪、不断完善与修正，为人工智能持续健康发展做出应有的贡献。

上海交通大学校长

2021年2月28日



# 目 录

001	<b>第 1 章 技术革命中的价值观演变</b>
001	比较视野中价值观的多样性
007	价值观的变与不变
013	技术驱动下价值观的历史嬗变
020	本章小结
022	<b>第 2 章 人工智能发展带来价值观的重塑</b>
022	从技术到产业：人工智能的革命性变革
028	从宏观到微观：人工智能对社会发展的冲击
034	从局部到整体：人工智能与价值观体系的重塑
039	本章小结
042	<b>第 3 章 人工智能与美国价值观</b>
044	被公开的不平等：算法歧视与技术出路
057	安全与归责：人机结合中的道德挑战
072	自由及其代价：从表面选择到真实赋能
088	本章小结

091 **第4章 人工智能与欧洲价值观**

092 差异竞争：确立以人为本的价值观战略

101 增进福祉：保障社会及公民权利

108 价值维系：技术、人才、法律与伦理“三轮驱动”模式

133 本章小结

135 **第5章 人工智能与日本价值观**

135 封闭与独立：独特外部文化接受模式

144 价值选择：新技术的发展与社会困境

154 文化锚点：传统与创新之间的羁绊

163 本章小结

165 **第6章 人工智能与中国价值观**

166 价值演进：从人机道德到数据治理

182 追求发展：技术赋能与生产力重构

196 融会贯通：至真至善、科技为民

208 本章小结

210 **后记 共同关注**

## 第 1 章

# 技术革命中的价值观演变

### 比较视野中价值观的多样性

价值观是多种人文、社会学科共同关注的问题，哲学、经济学、伦理学、教育学、人类学、社会学、社会心理学等学科都在这一领域进行过不同角度的探索。<sup>①</sup> 其中，哲学关注价值观所反映的主体和客体之间的关系；伦理学关注价值观对人的行为的规范性；人类学关注价值观表达的文化特征；教育学关注影响价值观形成和改变的个体社会化过程及其教育干预；经济学关注人类经济行为的深层心理原因和类型；社会学关注社会结构及社会变迁对价值观的影响；社会心理学则关注价值观的心理结构、过程、功能及其测量。<sup>②</sup>

早在 20 世纪 30 年代，美国心理学家奥尔波特（G. W. Allport）和弗农（P. E. Vernon）采用德国心理学家斯普兰格

---

① 李德顺. 价值学大辞典 [M]. 北京：中国人民大学出版社，1995.

② 杨宜音. 社会心理领域的价值观研究述要 [J]. 中国社会科学，1998（02）：3-5.

(E. Spranger) 对人的六种分类（经济的、理论的、社会的、审美的、宗教的、权力的）制订了一份“价值观研究量表”，进行了具有开创意义的价值观研究。<sup>①</sup> 20 世纪 50 年代，由美国人类学家克莱德·克拉克洪（Clyde Kluckhohn）提出的价值观定义确立了支配地位，从操作层面对价值观的各种定义进行了整合。他认为：价值观是一种外显的或内隐的，有关什么是“值得的”看法，它是个人或群体的特征，它影响人们可能会选择什么行为方式、手段和结果来生活。<sup>②</sup> 20 世纪 70 年代，美国社会心理学家米尔顿·罗克奇（Milton Rokeach）把价值观理解为一种信念，他把价值观分为终极价值观和工具性价值观两个方面，开始了从维度而不仅仅是从内容上对价值观的分析和测量，使价值观的研究进一步走向深入，<sup>③</sup> 由他编制的《罗克奇价值观调查表》至今仍是国际上广泛使用的价值观问卷。20 世纪 80 年代以来，以美国社会心理学家谢洛姆·施瓦茨（Shalom H. Schwartz）为代表的研究者开始从需要和动机出发来解释价值观的深层内涵，试图在此基础上构建一个具有普遍文化适应性的价值观的心理结构。施瓦茨认为，价值观是合乎需要的超越情境的目标，它们在重要性上不同，在一个人的生活中或其他

---

①（美）米尔顿·罗克奇（Milton Rokeach），人格理论（第 8 版）[M]，高峰强，等，译，西安：陕西师范大学出版社，2005：152。

②（美）克鲁克洪（Kluckhohn, C.），文化与个人 [M]，高佳，译，杭州：浙江人民出版社，1986：40。

③ Milton Rokeach. The Nature of Human Values [J], American Journal of Sociology, 1973, 89 (2).

社会存在中起着指导原则的作用。<sup>①</sup>

由于视角、方法的不同，研究者对价值观概念的界定难以统一，但我们不难从中看出一些共识。从价值观的主体角度考虑，它既可能是一种个体现象，也可能是一种社会现象，还可能是一种文化现象；从价值观的表现形式看，它既可能是外显的也可能是内隐的；从价值观的功能看，多数研究者认为价值观对行为具有导向作用。

与世界观、人生观一样，价值观具有多样性，涉及社会生活的各个领域，任何一个社会都存在着多种多样的价值观，它们反映了社会多种多样的文化传统，人们多种多样的生存条件、活动方式和利益等。在传统社会，由于社会生活分化的不充分以及社会关系的狭隘，价值观总体上具有一定的单调性和封闭性。现代社会的复杂多样及其冲突，打破了传统社会单调、封闭、僵化的状态。特别是由社会生产力从低到高的发展以及科技进步对全球空间版图的重构而推动的经济全球化，极大拓展了人类的生存空间，促进了全球治理体系的革新。与此同时，经济全球化也加剧了不同发展道路、文化、宗教信仰之间的激荡碰撞，价值观的多样性成为一个显著的事实。

社会层面的价值观研究在西方发达国家是较为普遍的。长期以来，许多学者致力于在社会、文化以及政治不断变迁的情

---

<sup>①</sup> Shalom H. Schwartz. Are There Universal Aspects in the Structure and Contents of Human Values? [J]. Journal of Social Issues, 1994, 50 (4).

境下，探讨全球不同国家和地区居民的价值与行为，并通过普查的方式探寻个人的微观价值观变化与社会变迁之间的联系。其中，由美国政治学家罗纳德·英格尔哈特（Ronald Inglehart）主持的世界价值观调查项目（World Values Survey, WVS）是涉及范围较广、产生影响较大的典型代表。该项目由瑞典的非营利组织“世界价值观研究协会”（World Values Survey Association, WVSA）主持进行，是一项描述世界社会文化和政治变迁等问题的调查研究。WVS的调查对象是全球范围内具有代表性国家或地区的普通民众，目前已发展到六大洲 97 个社会群体中，覆盖了 88% 的世界人口。具体说来，WVSA 从 20 世纪 80 年代起以 4 至 5 年为一个周期开展价值观普查，通过委派组织成员进行调查督导，委托当地研究机构进行以随机抽样访谈的方式来完成数据采集的工作。

基于 WVS 的数据，英格尔哈特和德国政治学家克里斯蒂·韦尔泽（Christian Welzel）设计出了一套四维世界价值观地图（见图 1-1）来展现全球范围内的跨文化差异（Cross-Cultural Variation）。在四维地图中，人类价值观被定义为“自我表达”“世俗理性”“传统价值观”“生存价值观”。其中，“世俗理性”与“传统价值观”反映了社会群体之间两个相反的过程，靠近“传统价值观”一侧更加强调家庭教育、传统礼仪和对权威的遵从，以及反对自杀、堕胎和讲究伦理等传统家庭观念。相反的一侧，即“世俗理性”则更加倾向于个人自身

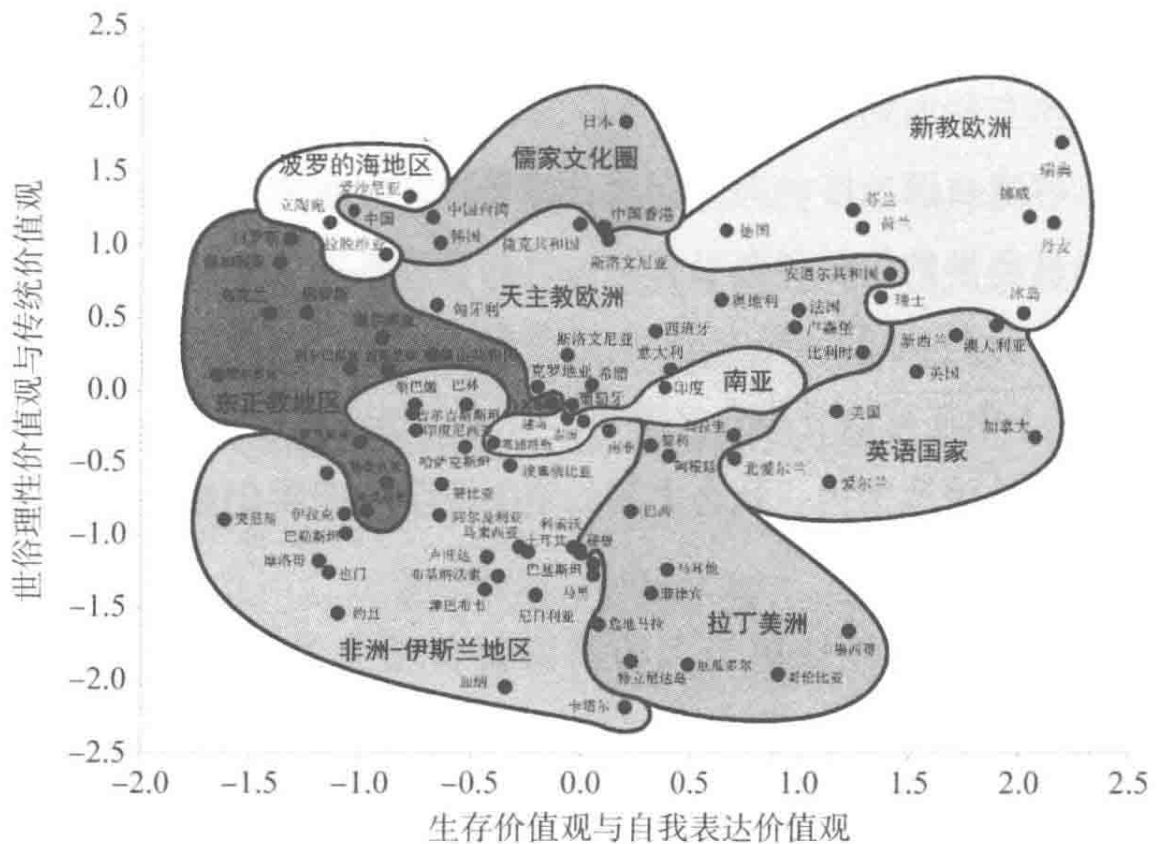


图 1-1 英格尔哈特—韦尔泽世界文化地图  
(Inglehart-Welzel cultural map of the world)

的实际需要。<sup>①</sup>

在图 1-1 上，纵坐标向上移动反映了从“传统”到“世俗理性”的转变，传统价值观强调宗教、亲子关系、尊重权威和传统家庭的重要性，信奉这些价值观的人拒绝离婚、堕胎、安乐死和自杀，同时一般都有高度的民族自豪感和民族主义观。世俗理性价值观与传统价值观有相反的偏好。在世俗理性价值观占优的社会宗教、传统家庭和权威不会得到特别的重

① Inglehart R., Haerper C., Moreno A., et al. (eds.). 2014. World Values Survey: Round Six-Country-Pooled Datafile Version: <https://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp>. Madrid: JD Systems Institute.

视，离婚、堕胎、安乐死和自杀被认为是相对可以接受的。横坐标向右移动则反映了从“生存”到“自我表达”的转变。生存价值观强调经济和人身安全，这与相对种族中心主义的观点以及低水平的信任和宽容有关。自我表达价值观高度重视环境保护，对外国人、男女同性恋者和不同性别更为平等。从图 1-1 中还可以看出，经济的全球化并未掩盖价值观的多样性，在全球范围内，社会之间的基本价值观差异仍然比社会内部大得多。

价值观多样性源于人的需要是价值关系形成的主体依据。<sup>①</sup>人们正是在意识到需要的基础上对各种价值关系进行判断、反思和整合，才形成了价值观。不同的人有不同的需要和自我意识，从而形成不同的价值观。人的需要的多层次性，决定了价值观的多层次性；人的需要的社会性，决定了价值观的社会性；人的需要的历史性，决定了价值观的历史性。

价值观的多样性必然引发价值观的冲突。现代社会价值观的冲突具有广泛性和复杂性，它表现为个人与个人之间，个人与群体、社会之间，以及群体与群体之间的价值观冲突。<sup>②</sup>在效率与公平、自由与平等、利益与道义、环境价值与经济价值等一系列重要问题上，不同主体常常得出不同的乃至截然相反的看法；同一个主体在不同领域、不同方面的价值取向也往往呈现出多变性与矛盾性。这些矛盾和冲突实质上源于不同形态的

---

① 季明. 核心价值观概论 [M]. 北京: 人民日报出版社, 2013: 9.

② 吴向东. 论价值观的形成与选择 [J]. 哲学研究, 2008 (5): 22-28.

价值观，如传统价值观与现代价值观、本土价值观与外来价值观、主导价值观与非主导价值观、宗教价值观和世俗价值观、精英价值观与大众价值观等之间的一系列的矛盾和冲突。不存在一个抽象的、永恒不变的、适应于任何时代、任何民族、任何阶级的价值观。不同的价值观，体现着不同的民族、阶级、社会集团对价值关系应然状态的期盼与展示。

## 价值观的变与不变

作为社会意识系统的有机组成部分，价值观在形成后将渗入人们的一切价值活动之中，是人们进行价值评价、选择、创造的导向和依据。只要人们始终站在现实历史的基础上，不是从观念出发来解释实践，而是从物质实践出发来解释观念的形成，<sup>①</sup>那么价值观就必然是建构在一定的社会经济基础之上的，是一定时代人们的社会存在、社会实践、生活经历的产物和表现，是一定时代文化传统、生活方式、风俗习惯、社会心理等因素潜移默化地濡染和熏陶的结果。

作为文化系统的深层结构，价值观具有相对的持久性或延续性，在相当长的时期内会自觉或不自觉地发挥作用，影响和支配人们的思想和行为。在特定的时间、地点、条件下，人们的价值观总是相对稳定的。例如，对某种事物的好坏总有

---

<sup>①</sup> 《马克思恩格斯全集》第30卷 [M]. 北京：人民出版社，1995：92，291.

一个看法和评价，在条件不变的情况下，这种看法不会骤然改变。

但是，人们的观念、观点和概念，随着人们的生活条件、人们的社会关系、人们的社会存在的改变而改变。<sup>①</sup> 价值观的不变是相对的，它必然会随着时代发展而发展、变革而变革，并随时接受人们的社会生活实践的检验、修正和完善。在社会发展平稳时期，价值观寓于日常生活之中，人们只是自觉或不自觉地听从它的导引。而在社会变革时期，新旧价值观常常产生碰撞、发生冲突，对价值观的比较、反思、批判、变革就会随之提上日程。社会变革的初期，价值观出现多元、紊乱和失序的情况，许多人感到茫然失措和无所适从。接下来，维护旧秩序的价值观和反映历史进步要求的价值观发生激烈的冲突，旧价值观的缺陷逐渐显现，而某些新的价值观则开始展示强大的生命力，得到大众的认同和拥护。最后，经过理性的论证，特别是实践的检验，人们选择和接受新的价值观，从而完成价值观的新旧交替。这种情况不只出现在历史上社会形态发生重大变革之时，即使在同一社会形态中，当社会生活发生巨大变革的时候，也可能会出现。

全球范围内，已有多项跨代际的长期调查试图研究价值观的变迁问题。1953年至1988年间，日本统计数理研究所国民性调查委员会每隔五年就进行一次关于日本国民性的调查。日本

---

① 《马克思恩格斯全集》第30卷 [M]. 北京：人民出版社，1995：291.