

法律人 读懂 人工智能

华宇元典法律人工智能研究院 / 编著

A LEGAL PROFESSIONAL'S
GUIDE TO THE
ARTIFICIAL INTELLIGENCE

看得懂、读得进、用得上

智知：突破法律人的“认知障碍”

践行：AI与法律行业融合的实践展现

趋势：未来“人工智能+法律”的发展方向



法律出版社 | LAW PRESS

让法律人 读懂 人工智能

A LEGAL PROFESSIONAL'S
GUIDE TO THE
ARTIFICIAL INTELLIGENCE

华宇元典法律人工智能研究院 / 编著



法律出版社 | LAW PRESS

图书在版编目(CIP)数据

让法律人读懂人工智能 / 华宇元典法律人工智能研究院编著. -- 北京: 法律出版社, 2018

ISBN 978-7-5197-2917-2

I. ①让… II. ①华… III. ①人工智能—应用—法律—工作—研究—中国 IV. ①D92-39

中国版本图书馆 CIP 数据核字(2018)第 261535 号

让法律人读懂人工智能

RANG FALÜREN DUDONG RENGONG ZHINENG

华宇元典法律
人工智能研究院 编著

策划编辑 杨大康

责任编辑 杨大康

装帧设计 汪奇峰 鲁娟

出版 法律出版社

总发行 中国法律图书有限公司

经销 新华书店

印刷 北京通州皇家印刷厂

责任校对 王晓萍

责任印制 吕亚莉

编辑统筹 法律应用·大众读物出版第二分社

开本 A5

印张 12.5

字数 278 千

版本 2019 年 1 月第 1 版

印次 2019 年 1 月第 1 次印刷

法律出版社/北京市丰台区莲花池西里 7 号(100073)

网址/www.lawpress.com.cn

投稿邮箱/info@lawpress.com.cn

举报维权邮箱/jbwq@lawpress.com.cn

销售热线/010-83938336

咨询电话/010-63939796

中国法律图书有限公司/北京市丰台区莲花池西里 7 号(100073)

全国各地中法图分、子公司销售电话:

统一销售客服/400-660-6393

第一法律书店/010-83938334/8335 西安分公司/029-85330678 重庆分公司/023-67453036

上海分公司/021-62071639/1636 深圳分公司/0755-83072995

书号:ISBN 978-7-5197-2917-2

定价:79.00 元

(如有缺页或倒装,中国法律图书有限公司负责退换)

序

我曾在 2013 年对学生的毕业致词里引用辛波丝卡的话：

“We're extremely fortunate not to know precisely the kind of world we live in.”（我们何其幸运，无法确知，自己生活在什么样的世界）。

今天，这句话或许可略做修改：我们何其幸运，无法确知，自己将生活在什么样的世界。

无法确知，无法预知，是因高速发展的科技正在重塑我们身边的点点滴滴，且它的速度没有减缓的趋势；幸运，是因这意味着未来充满更多新鲜的可能性，更多此刻我们无法想象的便利和惊喜，当然，也会有更多挑战。挑战对于法律人来说是值得期许的，随着新的事物而来的总是新的规则，若世上永远是一成不变的纠纷，法律人恐怕会渐无立足之地，新，恰是法律人的饭碗，以及存在意义。

当下，人工智能便是这“新”的代表。人工智能会不会取

代法律人？如果不会，它又能为我们做些什么？想起围棋界被AlphaGo支配的恐惧，法律人也会惴惴不安，然而低头望着累累卷宗，又会叹息人工智能为何还不能拍马赶来积极分忧解难。事实上，这个神秘存在正在日渐耳聪目明，而法律人对它尚知之甚少。

你的问题，这本书可能并不能允诺给你确切无疑的答案，但它汇集了法律人工智能这条道路上早早出发的探索者，以及对此感兴趣的思索者。他们为你展开的是人工智能触达法律界时，站在边界上的人们的思考图景和实践路径。他们叙述的不是终点的灿烂华美，而是站在起点上，张望着无法确知的未来世界，如何绘制抵达，或者准确地说，出发的路线。

人类并不是第一次遇到来自技术的挑战。读完这本书，你大概也会赞同这样一个说法——每一个法律人不应当去考虑技术是否会替代我，而应当去考虑当我拿起这个武器，如何去捍卫我所守持的正义的天平。

林 维*

* 中国社会科学院大学副校长。

序 二

2017年伊始，大家可以明显地察知人工智能成为了法律人热议的话题，但对于普遍文科背景的法律人，“读懂”与技术关系密切的人工智能并非易事。市面上与人工智能相关的书籍、演讲、文章多是通用领域或技术视角，并非为法律人量身定制。但是，人工智能要在法律领域实实在在地落地，法律人就需要有了解它的靠谱渠道，需要有法律人“看得懂、读得进、用得上”的人工智能书籍。

法律圈对人工智能的热炒是一种既期望又焦虑的情绪发酵：期望人工智能带来真正的变化，快速进入萨斯金所预言的明日世界，同时也担心人工智能会和法律人抢饭碗，甚至进一步取代法律人。其实，只有了解法律人工智能的“真实当下”，才能预判和迎接它的“可能未来”。

这本书梳理出与法律行业密切相关的专项技术，从法律人的角度出发，对这些技术要素重新做解读，形成一套法律人可

感、可知的人工智能技术底层知识架构，也展现过去几年智能技术与法律行业的融合应用实践，构建一幅多维度的探索图景；既集合和尝试回答法律人对人工智能争议最盛的问题，也邀请对这个领域的探索家与关注者分享他们的洞见及展望。希望让法律人在这里读懂人工智能，为法律人进入这个陌生领域尽力提供一些实在的帮助，这样，在讨论之前，我们便能有更多了解；在预言之前，便能有更多深思。

我们希望这个领域风生水起突破不断，而又能迎来更加新鲜、健康、扎实的讨论与研究氛围，那样，法律人必然能以冷静与睿智，拥抱人工智能将为这个行业带来的一切。

余贵清 *

* 北京互联网法院副院长、北京法院信息技术专家。

1

Chapter

1

第一章

智知：突破法律人的“认知障碍”

看上去很美的“智能”，可能不是人工智能！ | 002

与法律人相关的人工智能是什么？ | 010

法律知识图谱的概念与建构 | 019

机器学习的“黑盒”是什么？ | 032

人工智能进击法律的三条路径 | 039

人工智能的双面历程 | 045

Alpha Zero 横空出世，距离横扫法律界还有多久 | 053

人工智能系统的法理学思考 | 062

AI 时代数据之争，我们需要什么样的“数据权”？ | 085

数据法视角下“个人信息”的概念剖析 | 092

2

第二章

践行：AI 与法律行业融合的实践展现

法律人？工程师？法律 + 人工智能靠谁连接？ | 110

法律智能产品经理之我见 | 119

大数据如何助力司法改革科学决策 | 127

大数据时代下，如何进行司法行政信息化建设 | 135

智慧法检：法律大数据为法检行业升级赋能 | 152

民商事案件智能辅助办案系统的研发路径 | 163

律师端决策辅助的应用 | 174

大数据下法律实证研究的可能创新与传统依赖 | 182

生活在算法之下：AI 算法的歧视与反歧视 | 191

人工智能在公开的裁判文书中能够挖掘出什么？ | 201

数据时代的数据合规最佳实践 | 207

人工智能的“黑箱”真的“黑”吗？ | 222

延伸阅读：吉林电子法院迈向智慧法院的探索实践 | 229

延伸阅读：四川智慧法院建设实践 | 240

延伸阅读：深度解读《检察大数据行动指南

（2017~2020 年）》 | 262

3

第三章

趋势：未来“人工智能 + 法律”的发展方向

法律人为什么要拥抱人工智能？ | 278

与 AI 共处，法律人职业能力向何处迁徙？ | 284

在 AI 时代，法律人需要学编程吗？ | 290

人工智能，你可能还不“懂”法律人 | 298

数据竞争与司法裁判 | 304

用问答系统代替人工做法律咨询，距离我们还有
多远？ | 315

机器人能代替法官吗？ | 325

人工智能的法官职业之路 | 330

人工智能全面进军法律界？先做好这三件事！ | 342

纽约律师眼中的人工智能：到底砸谁的饭碗？ | 350

智能时代的数字正义理论 | 368

打破人工智能算法黑箱 | 377

是技术驯化了我们，还是我们驯服了技术？ | 384

后记 | 390

第一章

智知：突破法律人的 “认知障碍”

法律人了解人工智能，最大的屏障在于“认知障碍”。对于普遍文科背景的法律人来说，人工智能显得有些隔膜。本章从法律人的视角出发，逐一解读人工智能的通识知识。这里既汇集了法律人对这个行业前瞻性的早期思考，又汇集了科技进步下最新的行业思想，充分展现了法律人思想的流变，以此为基础帮助我们更好地建立起对人工智能的认知，深刻理解何为“法律 + 智能”。

看上去很美的“智能”，可能不是人工智能！

涂存超*

自 AlphaGo 战胜李世石之后，想必大家对人工智能这个概念耳熟能详，那么到底什么是人工智能呢？

首先，我们来看几张图片。



图 1



图 2

* 幂律智能 CEO。

大家认为哪些更应该是人工智能，哪些不太能称得上人工智能呢？对于图 1，大家可能在日常生活中都习以为常了，而图 2 索菲亚机器人和“高大上”的大数据统计界面是不是更符合大家对于人工智能的预期？答案恰恰相反，其实图 1 是人工智能的典型任务和应用；图 2 的索菲亚机器人，目前不少研究者认为其更像一个炒作，单纯的数据统计几乎没有利用到人工智能技术，即使它有一个非常酷炫的界面。

那么人工智能到底包含哪些内容呢？严格来说，人工智能是计算机学科的一个分支，主要研究如何让机器来模拟人的智能，来处理一些特定的场景和应用问题。具体来说，人工智能研究一般包括：语音识别、计算机视觉、自然语言处理、信息检索、机器学习理论、智能控制机器人、无人机无人车等。

从一个人工智能研究者的角度来说，目前我们对于人工智能主要关注点在于内部智能。比如说，对于一个具体任务，我们更关注解决方式和应用潜力上的智能，而不是这个任务目前的呈现效果看起来够不够智能。这里，关注解决方式和应用潜力的智能我们称为内部智能，关注外部呈现效果的智能我们称之为外部智能。像百度、Google 等搜索引擎、今日头条等新闻的推荐、搜狗拼音输入法、中英翻译系统等。这些人工智能技术，也许看起来不是那么有未来感、奇妙感，但却是我们这些人工智能研究者一直在研究的方向，也是一直有所进展的领域。

当然，在基础技术的支持下，人工智能技术同样能够实现非常多有趣的应用。例如，笔者所在的实验室在孙茂松老师的带领下，推出了一个自动作诗的 AI 系统——“九歌”，在央视《机

智过人》节目上获得了嘉宾的赞誉。九歌还根据这个节目的名字《机智过人》作出一首藏头诗。“机心造化本无私，智略工夫巧笑时。过客不须频眺望，人间天上有新诗。”我其实不太懂古诗，站在一个普通人的角度，我觉得九歌作出的诗还是非常有可取之处的。

下面这张图（图3）是九歌与三位专业选手比拼的集句诗，以“心有灵犀一点通”开头。



图3

不知道大家能否分辨出哪首是九歌作的诗？正确答案应该是第三首，“心有灵犀一点通，小楼昨夜又东风。无情不似多情苦，镜里空嗟两鬓蓬。”通过现场投票，这首诗骗过了大多数人，被大家认为是最像人作的一首诗。可以看到，九歌作出来的部分诗句几乎让人分不清真伪。这些应用是我们真正在内部智能的基础上所实现的有趣的应用。

与人工智能研究者更关注于内部智能相比，对于普通人来

说，往往更会关注外部智能，也就是说一个任务或者应用呈现出来的外部信息看起来智不智能。而这个判断会受很多主观因素的影响，人们往往不会关注背后技术的实现难度、智能程度。例如，展示界面做得够不够好，富不富有科技感、现代感、神秘感等。

在索菲亚机器人用一些极其流利且巧妙的回答让大家惊叹人工智能技术的同时，也引发了更多关于人工智能威胁论的担忧。微博上或者新闻里，每当有索菲亚机器人相关的新闻或者消息，总会有很多人发表对于她的智能水平深信不疑的言论，也会发表一些人工智能威胁的言论，以及发表对于人工智能技术的错误判断和过分期望的言论。而实际上，目前的人工智能技术，还远没有达到在开放领域进行这种流利的自然语言对话的程度，更别提很多对人来说都十分巧妙的回答。这个机器人只能回答好有准备的问题，或者已经记录下来问题。其实，这些对于人工智能技术的大肆不实宣传对于人工智能学科、产业界的发展有百害而无一利。在人工智能发展历史上也有惨痛的教训，比如 20 世纪 80 年代，以专家系统为代表的人工智能受到了狂热追捧，很多国家对专家系统的盲目期望，造成了大量政府、企业科研经费的投入用来开发专家系统；但是随着一些代表性工程的最终“流产”，人工智能的“泡沫”被戳破，进入了长时间的低谷。

人工智能的价值往往体现在它能够具体解决哪些问题上。但是大家一般会对问题的解决难度和思路存在非常大的偏差。这种情况在出现跨专业学科或问题的时候尤其突出。有一段时

间，网络上流传一个新闻，产品经理提出了让手机 APP 根据手机壳的颜色变换主题的需求，开发人员对于需求过于绝望，双方大打出手。这个问题概括起来就是，需求者有时候不清楚技术能够解决哪些问题，缺少对于技术能力和实现逻辑的理解。

我们拿 NLP 领域的机器翻译任务来举例。机器翻译和问答系统、对话系统，包括目前的自动作诗等都一样，都是 NLP 领域最上层也是最难的任务。为什么说它难呢，是因为对于待翻译的一句话，我们需要从一个庞大的可能的空间进行搜索，假设我们翻译出的句子长度为 10，总共有 10 万个词，那么可能的组合方式就有 10 万的 10 次方。

那么，机器翻译应该怎么做才合理，才对呢？

最简单的就是基于规则、模板的方法，利用一个双语的词表，进行词的翻译，然后根据语法规则来对翻译之后的内容进行语法调整。机器翻译最开始出现的时候，人们对于这种方法盲目乐观，认为机器翻译的问题会很快被解决，人类翻译员的工作很快会被取代。

但是后来发现，语言本身是非常复杂的，语言的多样性、歧义性、递归性，靠人类的知识定义翻译规则，规则是永远写不完的。不真正去解决这个问题，往往很难看清问题的难度和陷阱在哪里。

基于规则的专家系统行不通之后，统计机器翻译的方法被提了出来。这种方法是利用大量的双语对齐的训练数据，来学习两种语言之间的翻译规律；而不是让人类来定义这种翻译规