

摘要

随着刑事案件越来越复杂,大数据技术在刑事司法领域的作用也越来越明显,大数据证据也随之产生。由于大数据证据在法律上没有明确的定位,导致司法实践中对于大数据证据的适用较为混乱,没有形成统一的适用标准,阻碍了大数据证据的发展,因此,本文以刑事司法实践以及学术理论为基础,对大数据证据能否作为独立证据种类进行了论证。

大数据证据在适用过程中也产生了很多问题,在合法性层面上,由于缺乏对基础数据的审查,导致大数据证据的合法性存在缺失;在真实性层面上,算法黑箱问题致使人们对大数据证据的真实性存在质疑;在关联性层面上,大数据证据所呈现的是一种相关关系,从而增加了认定证据关联性的难度。同时,由于大数据证据自身的先进性,导致其不可避免地引发了证据偏在问题,影响了刑事诉讼的公平。

为更充分地适用大数据证据,本文从理论到实践分为五个部分,逐步推进对大数据证据的研究。

第一部分是大数据证据的界定。本章在分析大数据证据的内涵后,对大数据证据与其他证据种类进行了对比,明确了其具备成为独立证据种类的资格,并对其特征和在司法实践中的功能进行了分析。

第二部分对大数据证据的证据属性进行了分析。在明确大数据证据能够成为独立的证据种类后,从证据的三个要素出发并结合大数据证据三位一体的证据结构,对大数据证据的合法性、关联性、真实性进行了分析。合法性审查主要是考查一项证据是否具有作为证据的能力,例如,大数据证据在收集、运用过程中不符合法律规定或损害公民权利时,那么大数据证据就不具有证据能力,不能被作为证据使用;大数据证据的关联性需要结合其三位一体的特点进行综合认定。大数据证据的真实性也需要从基础数据的真实性、大数据技术的真实性以及大数据分析报告真实性出发进行研究。

第三部分主要研究大数据证据在刑事司法中的适用问题。针对大数据证据的证据属性分别在合法性、真实性、关联性三个层面出发指出大数据证据在当前刑事司法实践中存在的问题。

第四部分为大数据证据在刑事司法中的适用完善建议。主要对前文提出的问题设置相关的制度规则。对于大数据证据的法律地位问题，提出分阶段逐步解决的方案；对于合法性难题，主要通过加强对基础数据的合法性审查以及加强对隐私权的保护的方式解决；对于真实性难题，主要通过加强大数据算法透明度的方式以及培养具备综合能力的司法人才的方式进行解决；对于关联性难题，通过建立基础数据综合审查模式以及发挥法官在关联性认定中的作用的的方式解决；对于证据偏在问题，通过利用证据开示制度以及建立相应的数据弱势群体援助制度进行解决。

关键词： 大数据证据 大数据技术 刑事司法 证据规则

目 录

中文摘要	I
英文摘要	III
引言	1
(一) 选题背景与研究意义	1
1. 选题背景	1
2. 研究意义	1
(二) 国内外研究现状综述	1
1. 国内研究现状	2
2. 国外研究现状	3
(三) 研究思路与研究方法	3
1. 研究思路	4
2. 研究方法	4
(四) 重点难点与创新点	4
1. 重点难点	4
2. 创新点	5
一、大数据证据的概述	6
(一) 大数据证据的内涵	6
(二) 大数据证据的性质	7
1. 大数据证据的证据种类	8
2. 大数据证据的独立性证成	9
(三) 大数据证据的主要特征	11
1. 证据的传来性	11
2. 数据的海量性	11
3. 对技术和算法的依赖性	12
(四) 大数据证据在刑事司法中的功能	12
1. 侦查阶段的分析刻画功能	12
2. 审判阶段的证明功能	13

3. 办案质量和司法效率提升功能	14
二、大数据证据的证据属性分析	15
(一) 大数据证据的合法性	15
(二) 大数据证据的真实性	15
(三) 大数据证据的关联性	16
三、大数据证据在刑事司法中的适用问题	18
(一) 大数据证据的合法性难题	18
1. 缺乏对基础数据的合法性审查	18
2. 大数据证据取证过程存在瑕疵	19
(二) 大数据证据的真实性难题	20
1. 算法黑箱导致真实性缺失	20
2. 对大数据技术的陌生性导致真实性缺失	21
(三) 大数据证据的关联性难题	21
1. 海量基础数据关联性认定形式单一	21
2. 大数据证据的相关关系导致关联性认定困难	22
(四) 大数据证据的偏差性难题	23
1. 控辩双方取证能力差异	23
2. 控辩双方举证质证能力差异	24
四、构建大数据证据在刑事司法中的适用规则	26
(一) 明确大数据证据的法律地位	26
(二) 构建大数据证据合法性规则	27
1. 加强对基础数据的合法性审查	27
2. 加强对隐私权的保护	28
(三) 构建大数据证据真实性规则	29
1. 加强大数据算法透明度	29
2. 培养具备综合能力的司法人才	30
(四) 构建大数据证据关联性规则	31
1. 建立基础数据综合审查模式	31
2. 发挥法官在关联性认定中的作用	32

(五) 平衡控辩双方的数据差异	33
1. 试行证据开示制度	33
2. 建立数据弱势群体援助制度	33
结 论	35
参 考 文 献	36
后 记	38

引言

（一）选题背景与研究意义

1. 选题背景

随着时代的发展，大数据技术水平的提高，大数据证据在司法实践中的作用逐渐显露。大数据证据的优势主要有：第一，对于信息数据庞杂的案件，大数据证据能够依靠其自身的技术性，迅速提取出与待证事实有关的信息，提高了办案效率；第二，大数据证据以海量电子数据为基础产生，数据的海量性保证了大数据证据的准确性。

当前司法实践中，新型的犯罪类型逐渐增多，案件涉及信息量庞杂，涉案人员众多，传统的办案形式已经难以满足当前实践中的新需要，但以大数据技术为基础产生的大数据证据，正好能够解决上述问题，因此，在司法实践中引入大数据证据是大势所趋。

2. 研究意义

其一，理论意义：推动大数据证据的理论发展。大数据证据具有自身独特的特点，与传统证据种类有明显区别，这导致与传统证据种类所配套的适用规则，不能满足大数据证据的需要。为顺应时代发展，我们应探索出符合大数据证据自身特点的适用规则，为大数据证据成为新的证据种类寻找出路，进而完善我国的证据适用规则，推动刑事司法的发展。

其二，实践意义：促进大数据证据在实践中的运用。研究大数据证据的适用问题，能够有效解决实践中，对大数据证据适用混乱的问题，从而提高司法效率和司法质量。同时，对大数据证据的研究，顺应了时代发展的要求，有利于促进我国法治建设的进一步发展。

（二）国内外研究现状综述

1. 国内研究现状

目前，我国对大数据证据的研究已经越发深入，目前，主要围绕以下问题展开：一是新兴技术对诉讼法带来的冲击，左卫民认为，随着数字技术的发展，形成了数字社会这样一种新的社会样态，同时，受到数字技术的影响，传统诉讼法

必然会发生重要变革，数字诉讼法将应运而生^①。李训虎指出，虽然我国刑事司法人工智能走在世界前列，但也存在很多问题。要想充分发挥人工智能在司法领域的作用，就需要改变传统思维、创新监管模式，对刑事司法人工智能进行包容性规制，以实现发展与规制的协调。^②

二是关于大数据侦查的研究。大数据侦查，是大数据证据在刑事司法领域的一个重要研究方向。卞建林、钱程指出，大数据侦查通过数据搜索查询、对比碰撞、挖掘分析等方法进行犯罪回溯侦查与预测防控，其高度技术性、适用隐蔽性、运行多维性使得侦查权内部张力增强，侦查权呈现扩张趋势。传统侦查程序规范对大数据侦查缺乏有效的法律控制，泛用大数据侦查存在冲击刑事正当程序、侵犯公民基本权利等风险。^③同时胡铭、张传玺认为，大数据背景下侦查权正在呈现出显性扩张和隐性扩张两类新趋势，大数据在侦查实务中扩张了侦查权限，对公民的基本权益带来了相当程度的侵害风险，并应当予以规制。^④刘玫、陈雨楠认为，在网络犯罪增多，传统犯罪逐渐向网络化发展的新环境下，侦查人员应合理利用数据技术协助办案，同时开辟出大数据侦查在实践运用中的新路径。^⑤

三是关于大数据证据在法律上的定位问题研究。大数据证据在法律上的定位问题，是当前学界讨论最激烈的一个问题。林喜芬指出，随着时代发展，新型证据不断涌现，而由于我国封闭的证据种类规定，导致了实践中司法人员难以充分利用新型证据。同时他还认为，如果将研究重点从法定证据种类转变为法定的证据方法上，就能在一定程度上解决大数据证据的证据种类定位问题。^⑥张吉喜、孔德伦认为，大数据证据的证据种类应当根据具体情形作具体分析，并在归纳了大数据证据具有衍生性、科技性、间接性和与镶嵌论的契合性之后，提出应从特点入手把握大数据证据。^⑦杜鸣晓认为，大数据证据属于电子数据证据的一种，当符合证据属性的大数据被作为证据使用时，应认定为“电子数据”，同时大数据也可以作为电子数据的载体形式，成为电子数据的一种，从而作为合法的证据

^① 参见左卫民：《迈向数字诉讼法：一种新趋势？》，《法律科学(西北政法大学学报)》2023年第41期。

^② 参见李训虎：《刑事司法人工智能的包容性规制》，《中国社会科学》2021年第2期。

^③ 参见卞建林，钱程：《大数据侦查的适用限度与程序规制》，《贵州社会科学》2022年第3期。

^④ 参见胡铭，张传玺：《大数据时代侦查权的扩张与规制》，《法学论坛》2021年第36期。

^⑤ 参见刘玫，陈雨楠：《数字时代侦查措施的体系建构及程序控制》，《河北学刊》2023年第43期。

^⑥ 参见林喜芬：《大数据证据在刑事司法中的运用初探》，《法学论坛》2021年第36期。

^⑦ 张吉喜，孔德伦：《论刑事诉讼中的大数据证据》，《贵州大学学报(社会科学版)》2020年第38期，第85页。

进入诉讼过程。^①刘璐则持不同看法，并认为电子数据与大数据证据存在明显区别：电子数据能够简单、直接地证明案件事实而大数据证据则更为复杂；她指出电子数据是案发时就形成的，而大数据证据则需要后天的技术算法生成。^②通过当前学界对这一问题的研究不难看出，大数据证据的法律定位依然不明确。

当前，有关大数据证据的研究，大多聚焦于法律定位问题上，缺乏了对其适用制度的研究，不够全面。本文在研究大数据证据的法律定位问题的基础上，还着重分析了当前司法实践中，适用大数据证据时存在的问题，并针对实践中存在的问题提出了新的解决方案。

2. 国外研究现状

大数据证据的法律定位以及证明力问题在外国学术界也存在激烈的讨论，不同学者之间的观点也不尽相同。

一是有关大数据证据性质的研究。在与证人证言的关系问题上，国外学者存在不同观点。美国学者加纳（Bryan A. Garner）教授认为，大数据证据不同于普通的证人证言。他指出，大数据证据与普通的证人证言有明显区别。普通证人的意见，主要是以其自身的逻辑判断、社会经验和掌握的知识为依据，对证言中所涉及的事实进行分析后所得出的认知或论断。^③

美国学者洛斯（Andrea Roth）教授则提出“机器证言”这一观点。她指出，机器算法在解决法律争议问题时的作用越来越明显，根据机器算法得出的信息可以被视为“机器证言”。她还解释到，机器算法中存在的“算法黑箱”很可能会削弱“机器证言”的可信性。^④

二是有关算法技术是否公开的研究。一种观念认为，如果将基于算法技术产生的大数据证据提交到法庭当作证据使用，那么开发者就必须公开算法方案，不公开将影响大数据证据的证明力；同时，如果“开发者能够公开机器算法，那么该算法就更容易被法庭接受，进而增加大数据证据的说服力”^⑤。但美国学者洛斯（Andrea Roth）教授认为，“对于不公开算法的谷歌地图数据分析报告，往往

^① 参见杜鸣晓：《大数据作为诉讼证据的可行性分析—以互联网租车行业为例》，《南华大学学报(社会科学版)》2017年第18期。

^② 参见刘璐：《大数据证据治理中存在的主要问题及对策》，《网络空间安全》2022年第13期。

^③ Garner, Bryan A. Black's law dictionary. Thomson Reuters publish ,2004.

^④ Roth, A. Machine testimony. Yale LJ,1972.

^⑤ Simmons R. Quantifying criminal procedure: how to unlock the potential of big data in our criminal justice system. Mich. St. L. Rev,2016.

因其高度的准确性而被认定具有很强的证明力。持反对意见的当事人，不能仅因其没有公开算法，而要求法官认定其无效”^①。

（三）研究思路与研究方法

1. 研究思路

首先，确定大数据证据在我国证据体系中的定位问题，构建大数据证据的证据规则，重构我国的证据体系，将大数据证据作为新兴独立证据，使其更好地运用于刑事司法实践中。其次，从证据的三性出发，解决大数据证据的证据能力和证明力问题，对大数据证据的独立性进行论证，寻找将其作为独立证据的理论基础。最后，提出并解决大数据证据在司法实践中存在的问题，对大数据证据在刑事司法中的运用提出自己的见解。

2. 研究方法

其一，文献分析法：通过阅读相关期刊论文、学位论文、著作建立理论基础。通过阅读文献了解现存理论观点，为本文写作奠定理论基础。同时，在分析文献的过程中，找出现存理论研究中的不足之处，为本文的写作寻找出发点。

其二，实证分析法：通过查询有关大数据证据的案例，归纳出大数据证据在刑事司法实践中的适用领域、适用方式，并总结大数据证据在适用中存在的问题。同时，基于对相关案例和数据的分析，为大数据证据的运用困境寻找出路，并为理论创新打下基础。

（四）重点难点与创新点

1. 重点难点

本文的难点在于，大数据证据作为时代发展的产物，其自身所具备的科学性、专业性，使得在研究这一问题时，需要先学习有关机器学习、算法黑箱、大数据可视化等与大数据技术有关的知识，了解运用哈希值检测大数据一致性的方法。同时，还需要理解大数据证据产生的过程和内涵，并对大数据证据三位一体的特点有充分的理解。

本文的重点是明确大数据证据未来的法律地位，主要是在大数据证据的合法性、真实性、关联性进行论证的基础上，提出大数据证据具有独立的法律地位，同时发现大数据证据在实践中的适用难题并提出解决方案，以此保障大数据证据

^① Roth, A. Machine testimony. Yale LJ,1972.

在实践中的充分运用。

2. 创新点

对于大数据证据在实践中存在的问题，本文采取逐步推进的解决方式，根据不同的社会发展阶段提出不同的解决办法，循序渐进地解决大数据证据的使用问题。同时，针对实践中存在的证据偏在问题，本文从现有制度出发，参考法律援助制度提出数据弱势群体援助制度，为解决大数据证据的证据偏在问题提出了新的解决方案。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/598012016006007005>