网络空间安全学科系列教材

数字取证

网络空间安全学科系列教材

教育部高等学校网络空间安全专业教学指导委员会
中国计算机学会教育专业委员会
共同指导

丛书顾问委员会主任：沈昌祥 | 丛书编委会主任：封化民

数字取证

陈晶 张俊 何琨 郭永健 朱勇宇 编著

根据教育部高等学校网络空间安全专业教学指导委员会编制的
《高等学校信息安全专业指导性专业规范（第2版）》组织编写

清華大學出版社

内容简介

数字取证技术作为网络空间安全的重要组成部分，在获取犯罪证据、打击网络犯罪中起到不可替代的作用。随着“电子数据”时代的到来，电子数据也已经在新的刑事诉讼法、民事诉讼法、行政诉讼法中被明确为证据类型之一，是新的“证据之王”。本书覆盖数字取证技术的主要知识点，是一本用户数字取证技术教学的入门教材，适合国内网络空间安全方向本科生、公安院校网络安全与执法方向学生使用。全书共设计了十二个章节。第一章至第四章节覆盖电子数据的基础知识、数字取证的发展和应用讲起，围绕电子证据的获取、数据过滤和搜索和初步解析、主流文件系统和数据恢复基础等取证知识和技能，涵盖哈希计算、文件过滤、关键字搜索、元数据提取、数据恢复。第五章至第十章讲解操作系统和痕迹，基于国际主流的Windows、Linux 和 macOS 桌面操作系统和 Android 和 iOS 移动终端操作系统的基本原理、安全架构和痕迹特点进行学习，涵盖了回收站和废纸篓、最近执行的程序，最近打开的文件，注册表分析、日志分析和内存分析等核心的取证知识点。最后两章就未来的取证发展和挑战进行了探讨，对区块链、数字货币、侧信道取证、汽车取证、无人机取证、物联网设备取证等方面知识进行学习，并对数据加解密、隐藏和反取证技术进行的简要介绍。本书使用中，可配合为虚拟仿真教学实验环境而设计的全套教学课件、实验案例、实验指导手册和习题，以达到最佳的教学效果。

前言

随着大数据、互联网、人工智能时代的到来，以数字化、网络化、智能化为主要特征的创新浪潮一浪高过一浪，现代科技正在推动社会变革和进步，同时也深刻影响着司法办案的模式、证据类型和证据规则。受高科技的冲击，利用信息技术进行的犯罪急剧增长，为了适应打击新型犯罪的需要，提升司法人员的电子数据取证能力势在必行。网络空间安全专业，致力于培养“互联网+”时代能够支撑国家网络空间安全领域的具有较强的工程实践能力，系统掌握网络空间安全的基本理论和关键技术，能够在网络空间安全产业以及其他国民经济部门，从事各类网络空间相关的软硬件开发、系统设计与分析、网络空间安全规划管理等工作，具有强烈的社会责任感和使命感、宽广的国际视野、勇于探索的创新精神和实践能力的拔尖创新人才和行业高级工程人才。

武汉大学国家网络安全学院于 2018 年开设了“数字取证” 课程，一直在研究并探索如何培养高水平的网络安全人才。随着课程的不断完善，我们发现数字取证作为一门实用技术，能够有效帮助网络安全专业学生开阔眼界，补充必要的专业知识。但在实际教学中我们也发现，国内适用的数字取证教学书籍较少，适合高校使用的教材则更少。为此，我们联合了国内知名专家和学者联合编写了本书。《数字取证》一书汇集了数字取证的基本概念、基本理论、工作原理和基本方法；系统全面地介绍了不同文件系统的取证分析方法；介绍了内存、日志、移动终端、恶意代码的取证分析方法。作者们在本书中融合了近年国内外电子数据取证理论和实践的最新成果，全面介绍了国内外主流的电子数据取证技术。《数字取证》紧密结合电子数据证据取证工作实际，既考虑了电子数据取证的知识体系，又尊重了学习和取证实践的规律，按照理论和技术并重的编写思路，将“实践”与“理论”完美结合，深入浅出地讲解数字取证的基本原理，结合电子数据取证实践性强的特点，通过实际的案例分析和实践练习将理论与实际和现实世界联系起来，帮助学生更好地学习和掌握取证知识。

本书由武汉大学陈晶、郭永健、何琨三位教授，以及湖北警官学院张俊教授、香港大学朱勇宇共五位作者共同编著完成。此外，本书编写过程中得到了皮浩、谢明聪、白虹软件董事长胡力和、深圳安证计算机司法鉴定所副所长魏智煌的大力支持，他们为本书第二章、第三章、第十一章、第十二章提供了内容和素材。

《数字取证》是一本理论实践相结合的课程，仅有理论内容无法培养出具有实战能力的取证人才。武汉大学国家网络安全学院经过五年的教学实践，采用校企联合方式开发了数字取证实验课程虚拟教学环境，利用丰富的教学案例和实验工具配合数字取证理论课程。为配合本书的理论教学，本书作者还有针对性地设计了具有 8 学时、16 学时、32 学时和 48 学时数字实验课程，并在虚拟教学实验环境中配套有数字取证实验指导手册，每一章的教学都对应了相应的虚拟化教学实验题目和实验指导手册。数字取证实验教学中，由于所有学生都有相同的练习环境，授课教师可以很好地进行学习进度控制和实验分数汇总。相信通过配合丰富的数字取证实验内容，学生收获的不仅仅是单纯的数字取证理论知识，也会在实践和动手能力上有显著提升。

本书配套的教学课件、部分实验案例、实验指导手册仅对希望开设数字取证选修或必修课程的大学提供，有需要的老师请请联系 cdf@cflab.net。

目录

第 1 章 数字取证概述

随着网络的普及和信息技术的飞速发展，越来越多人在享受着网络带来的便利。根据 2022 年 8 月中国互联网络信息中心发布的第 50 次《中国互联网络发展状况统计报告》，截至 2022 年 6 月，中国的网民规模已经达到 10.51 亿，互联网普及率达到 74.4%，人们在工作、社交、科学、艺术、商业交易、社会管理等生产生活的各方面，越来越依赖于网络空间提供的信息手段和方式方法。

网络空间已经成为人类新的生存空间，随之而来的是网络空间的安全管理和风险防范等问题，这些问题将网络空间与现实社会关联起来，导致各类新型的安全威胁甚至犯罪层出不穷。数字取证是伴随着计算机的发展，以调查这些安全威胁事件和涉网犯罪行为为目的而逐步形成的。它是涉及计算机科学与技术、法学、信息安全等新兴、交叉和前沿的领域，是网络空间安全学科的一个年轻而又充满活力的分支。

数字取证是一种主动的安全防御措施，不同于在技术层面的对抗以及在管理层面的被动防御，它将防线退到最后----通过取证获得犯罪或违规的证据，然后诉诸法庭。经过近 40 年的发展和演变，与网络空间安全其他领域不断融合发展，数字取证已经逐步形成完整的标准规范、证据规则、方法流程、软件和工具的验证、产业应用体系等，并形成了鲜明的特色，在维护网络空间安全、推进社会法治建设、法庭科学应用、打击治理涉网新型犯罪等方面发挥着巨大的作用。

本章主要介绍数字取证与电子数据证据的基本概念、数字取证的实施、数字取证的技术标准与规范以及数字取证的现状与发展等，帮助读者在正式开始学习数字取证的方法和技术之前，对该领域有系统和全面的理解。

1.1 数字取证与电子数据证据

1.1.1 数字取证的概念

二十世纪八十年代末，计算机应用开始逐步普及。人们将从计算机中发现和恢复与特定案件相关联数据的方法学定义为计算机取证（Computer Forensics），它研究关于计算机犯罪证据的获取、保存、分析和出示的法律规范与科学技术。后来随着计算机网络以及电子设备的广泛使用，数据不再局限于计算机中，电子取证（Electronic Forensics, eForensics）或电子发现（eDiscovery，Electronic Discovery）的概念逐渐被接受。进入二十一世纪后，数据的存在形式和表现方式更加多样，数字取证（Digital Forensics）的概念得到业界更普遍的认同。当前，这些概念通常可以互换使用，但还是有所区别，都从不同侧面反映着研究范畴和发展轨迹。计算机取证的研究对象是存在于计算机系统以及各种电子设备中的数据信息，电子取证或电子发现的研究对象是以电子信号方式存储的数据信息，而数字取证的研究对象泛指数字化存储的数据信息。在国内，数字取证也被称为电子数据取证。

目前，还没有权威组织对数字取证给出一个统一的定义，不同专家有不同看法。其中一种定义是“使用科学的、被证明的方法，对来自于数字设备的数字证据进行保护、收集、确认、辨识、分析、解释、归档和呈供，以协助犯罪事件的重建、预测违规操作的非授权行为”。还有定义是“使用特定的技术，对电子数据进行恢复、验证、分析，以便在法庭上提交证据”。国际电子商务顾问局（EC-Council）认为“计

算机取证指一套方法程序和技术，帮助识别、收集、保存、提取、解释、记录和呈现来自计算机设备的证据，以便任何发现的证据都可以在刑事或民事及行政程序中被接受”。维基百科认为是“法庭科学的分支，包括从能够存储数字数据的数字设备发现的材料中，进行恢复和调查的过程”。

本书认为数字取证是研究如何对数字化存储的数据信息进行获取、保存、分析和出示的法律规范与科学技术。从信息论的角度看，这是一个从媒介（例如存储设备、网络等）到证据的转换过程，即遵循特定的标准方法或作业指导，通过辨识和收集，获得存储的媒介；然后通过对媒介进行恢复和检查，获得与案件相关联的电子数据；通过对数据进行解释和分析，发现数据中的信息；最后通过报告和展示等方式，形成信息中包含的证据，以证明或证伪某种假设。这个通过收集、检查、分析到报告的过程，工作对象从媒介、数据、信息到证据，就是数字取证的总体过程，见图 1-1。

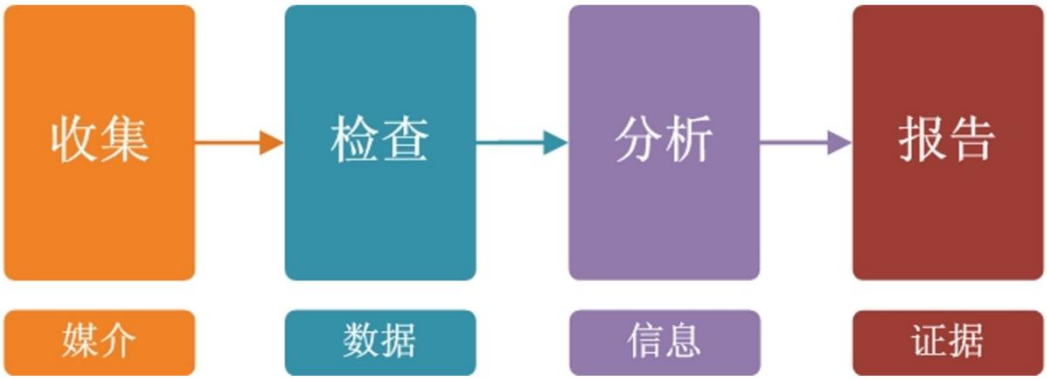

图 1-1 从媒介到证据

1.1.2 电子数据作为证据

1.电子数据及其特点

数字取证的目的是获得电子数据（或称为数字证据）。在数字取证的语境中，电子数据是以数字化形式存储、处理、传输的，能够证明某个事实的数据。其特点在于：

第一，客观存在性。这种客观物质性是决定其证据效力的基础。电子设备在正常运行状态下进行的运算、存储、发送、接收等操作，不管是用户有意或系统自动完成，都会留下痕迹，反映着用户操作或程序运行的意图，所以与传统的证据形成的形式一样，都具有客观存在性。

第二，多样多态性。电子证据以光电磁等物质，通过电子信号或编码形式存在于各种媒介。例如光电磁的信号被以复杂的方式编码和解码，表现为二进制，即以数字 0 或 1 编码的方式存储和传输，不像传统证据，例如书证和物证那样能被人直接感受接触和理解。多样性体现在这些数据通过解析，表现为声音图像、图形动画、文本消息、程序算法、协议信令等，或者其组合。多态性也体现在同样的数字数据上，经过不同的解析方法，可能具有完全不同的含义，而这些解析方法，还都符合某些标准和规范。

第三，脆弱易变性。电子证据的形成、存储和传输依赖于特定的设备和方法，非常容易被改变和破坏。例如网络传输的数据转瞬即逝，删除一个文件只需按下删除键，损坏一个磁盘的盘片会导致其中的数据完全不能读取，修改电子表格的数据导致恢复困难。因此，才需要制定专门的电子证据规则。

第四，抽象间接性。抽象性体现在必须以完整、客观、直观的方法展现电子证据的证明力，以帮助非

专业人士对证据的理解。间接性体现在电子证据往往不能直接证明某个事实，还必须借助客观世界的事件与之关联，所以电子证据的证明力体现在其真实性、完整性和关联性。对其进行的检查、分析、处理、展现，包括现场重建（Scene Reconstruction）、溯源追踪等必须借助最现代化的技术和手段，依赖特定的装备和环境。

2.电子数据的法定地位

数字取证的目的是为了获得电子证据。电子数据能否作为证据？目前，各个国家都持肯定态度。我国有关法律规定："证明案件真实情况的一切事实，都是证据"，目前我国的各项法律规范均认可电子数据的证据地位。例如，我国三大诉讼法都相继明确地赋予电子数据与书证物证相同的法律地位。《中华人民共和国民事诉讼法》第六十六条、《中华人民共和国刑事诉讼法》第五十条、《中华人民共和国行政诉讼法》第三十三条等，都在证据类型中列举了"电子数据"。

在最高人民法院、最高人民检察院、公安部发文《关于办理刑事案件收集提取和审查判断电子数据若干问题的规定》（2016 年法发〔2016〕22 号）中，第一条采取"概括+例举+排除"的方式，对电子数据作了明确界定。其中第一条规定，电子数据是案件发生过程中形成的，以数字化形式存储、处理、传输的，能够证明案件事实的数据。电子数据包括但不限于下列信息、电子文件：

（1）网页、博客、微博客、朋友圈、贴吧、网盘等网络平台发布的信息；

（2）手机短信、电子邮件、即时通信、通信群组等网络应用服务的通信信息；

（3）用户注册信息、身份认证信息、电子交易记录、通信记录、登录日志等信息；

（4）文档、图片、音视频、数字证书、计算机程序等电子文件。

同时规定，以数字化形式记载的证人证言、被害人陈述以及犯罪嫌疑人、被告人供述和辩解等证据，不属于电子数据。

1.1.3 电子数据的证明力

1.电子数据的证据效力

电子数据的证据效力源自证据的客观性。电子数据必须经过关联性、合法性与真实性的检验，才能作为定案的根据。为保证电子证据的证据能力，首先要将电子证据能够转换为法定证据形式。在实际司法实践中，通常可以将电子证据呈现为司法鉴定意见、勘验报告和视听材料等。

2.电子数据的证据力

电子数据的证据力是指证据的真实性、合法性。要求证据的扣押、提取、分析、保管等过程依照标准和规定程序执行。具体应当从以下几个方面加以保证：

（1）证据来源的原始性。可通过自认、证人具结、推定方式、鉴定方式等实现。例如在扣押现场的拍照录像、见证人签字、材料交接过程的保管链（Chain of Custody）等。

（2）可以验证的完整性。是指电子数据本身的完整性以及电子数据的处理所依赖的环境和系统的完整性。例如要验证在现场采集获得的数据没有被篡改，通过对数据计算完整性校验码或者对原始证物封存

加以保护。

（3）技术手段的科学性。是指在实施电子数据勘验、检查、分析等的操作中，使用的软硬件、相应的勘验和检查流程、获取和分析方法、结论时的逻辑推理等，符合一般科学原理，并确保在遇到异常情况时，具有作出合理解释的能力。

（4）操作过程可再现性。保证数字证据真实性和完整性的方法之一，就是要求勘验、检查人员对数字化证据实施的操作应当是可再现的。例如对提取、处理、存储、运输过程有详细的可追踪记录，从而确保对数字证据的操作过程可还原。

3.电子数据的证明力

证明力就是指证据与待证事实之间的关联性，也就是证据与结论之间是否符合逻辑。

（1）原理科学性和推理逻辑性。数字取证所依赖的软硬件、技术原理符合科学原理，能够经受事实的考验。从证据到结论这一过程符合逻辑，这是推断证据运用是否科学的重要依据。

（2）分析结果可再现性。指根据相同的原理、程序和方法，在相同数据集合上进行分析，任何人都能够得到相同的结果，这种稳定不变的特性就是数字证据成为科学证据的首要原则。

4.数字证据的质证

质证是指在庭审过程中，当事人就法庭上所出示的证据，采取辨认、质疑、说明、辩论等形式进行对质核实，以确认其证明力的诉讼活动。由于电子数据的抽象间接性特点，现代司法实践越来越强调电子数据法庭质证的重要性。英美法系创设了专家证人（Expert Witness）制度，司法鉴定人作为专家证人在出庭时接受质询，帮助法庭理解复杂的专业性问题。我国的《刑事诉讼法》规定了专家辅助人制度，当事人可以聘请有专门知识的人出庭对司法鉴定意见提出异议，以增强控辩双方在诉讼活动中的对抗性。

1.1.4 电子数据司法鉴定

数字取证最初是为满足执法部门的案件调查而兴起的，目前执法部门仍然是数字取证最重要的应用领域。近年来，数字取证应用到越来越广泛的领域或行业，例如网络空间安全事件响应和威胁捕获、商业调查和流程监管（包括尽职调查/内部调查等）、军事行动、审计事务、知识产权调查等。现今，数字取证在刑事、民事和行政案件调查中得到广泛应用，成为了法庭科学（Forensic Science）新的组成部分，如警务部门的电子数据勘验和案件调查，司法部门的电子数据司法鉴定等。

司法鉴定是司法诉讼活动中的一个重要环节。当事人可以就某些专门问题提出司法鉴定请求，委托司法鉴定机构，由司法鉴定人运用专业知识和技术，依照法定程序作出鉴定和判断，并出具司法鉴定意见书。随着电子数据的广泛存在，需要进行电子数据司法鉴定的案件越来越多，审判机关对电子数据司法鉴定活动也越来越重视。

根据司法部司规（2020）3 号《声像资料司法鉴定执业分类规定》文件，电子数据鉴定是声像资料司法鉴定执业分类规定中的专业之一，电子数据鉴定是指鉴定人运用信息科学与技术和专门知识，对电子数据的存在性、真实性、功能性、相似性等专门性问题进行检验、分析、鉴别和判断并提供鉴定意见的活动。

包括电子数据存在性鉴定、电子数据真实性鉴定、电子数据功能性鉴定、电子数据相似性鉴定。四种鉴定类型的详细说明如下：

电子数据存在性鉴定：包括电子数据的提取、固定与恢复及电子数据的形成与关联分析。其中电子数据的提取、固定与恢复包括对存储介质（硬盘、光盘、优盘、磁带、存储卡、存储芯片等）和电子设备（手机、平板电脑、可穿戴设备、考勤机、车载系统等）中电子数据的提取、固定与恢复，以及对公开发布的或经所有权人授权的网络数据的提取和固定；电子数据的形成与关联分析包括对计算机信息系统的数据生成、用户操作、内容关联等进行分析。

电子数据真实性鉴定：包括对特定形式的电子数据，如电子邮件、即时通信、电子文档、数据库数据等的真实性或修改情况进行鉴定；依据相应验证算法对特定形式的电子签章，如电子签名、电子印章等进行验证。

电子数据功能性鉴定：包括对软件、电子设备、计算机信息系统和破坏性程序的功能进行鉴定。

电子数据相似性鉴定：包括对软件（含代码）、数据库、电子文档等的相似程度进行鉴定；对集成电路布图设计的相似程度进行鉴定。

1.2 数字取证的实施

1.2.1 数字取证的原则

按照 ISO/IEC 27037:2012《信息技术—安全技术—数字证据的识别、收集、获取和保存指南》，数字取证的原则可以归纳为四条，即合法性（Legality）、关联性（Relevance）、可靠性（Reliability）与充分性（Sufficiency）。即使电子数据不进入司法程序，这四条对于所有的调查活动也都至关重要。

1.合法性

合法性是指数字取证的全部过程都符合法律的要求。合法性是证据可采信的前提，包括主体合法、对象合法、手段合法、过程合法等内容。由于各个国家法律体系的不同导致具体法律要求出现极大变化，所以数字取证在满足一般法律要求前提下，也必须遵照特定法律体系的要求进行。

2.关联性

关联性是指数字取证的目的一定是为证明或证伪某个特定被调查事件的某个方面而展开，例如时间、地点、人物等。首先要能够说明获取的材料与调查相关，其中包含了帮助调查特定事件的有价值的信息，才有理由进行扣押或获取。然后通过审查和判断，取证调查人员才能形成接下来的取证过程，并解释采取各种取证方法的理由，包括：第一，能够说明使用某个方法和技术的决策，是得到所有潜在证据的最佳选择；第二，能够说明复现或验证所采取的行动或方法。

3.可靠性

可靠性是指被调查的电子数据就是它原来的状态或样子。电子数据具有脆弱易变性，有意或无意的不

恰当操作行为，都有可能造成数据发生改变，导致跟原始数据不一致，即被污染。可靠性要求电子数据处理的所有步骤都是可审查和可重复的，并且应用这些步骤的结果都是可复现的。

可审查性（Auditability）是指独立的评审员或权威方可以评估取证调查人员进行的取证活动。只有准确记录所有采取的步骤或操作，才有可能做到这一点。取证调查人员应该能够给出选择一个给定的操作路线的决策理由，这样独立的评估才能决定取证调查人员执行的步骤是否遵循恰当的科学方法、技术或流程。可重复性（Repeatablity）是指在下列条件下具有相同的取证结果：使用相同的测量方法和步骤；在相同条件下使用相同的仪器设备；在初始操作之后，还可以重复该操作任意次。可复现性（Reproducibility）是指在下列条件下具有相同的取证结果：使用相同的测量方法；在不同条件下使用不同的仪器设备；在初始操作之后，还可以重复该操作任意次。

数字取证依赖于科学理论和技术方法。在判断和审查某个科学理论或技术方法时要注意：

（1）该科学理论和技术方法是否经过测试；

（2）该科学理论和技术方法是否经受了同行评审和结果公布；

（3）该科学理论和技术方法的错误率是否已知并报告；

（4）该科学理论和技术方法是否有标准以规范其应用；

（5）该科学理论和技术方法是否被业界广泛接受。

4.充分性

充分性是指在数字取证中必须收集足够的潜在电子数据，以便事件的相关因素都得到考虑和调查。充分性不表示取证调查人员一定要收集到所有的数据，或生成原始材料的完全副本。实际情况中，往往也做不到收集所有的数据或生成原始材料的完全副本，例如部分数据被删除，或部分通信流量数据已经消失等。

1.2.2 数字取证的模型

二十世纪末期，由于数字取证得到广泛应用，人们逐渐意识到，由于数字取证基本理论形式化研究、基本过程模型等缺乏一致性所带来的种种弊端，导致数字取证的可操作性较差，且极易在法庭上遭到质疑，因此专家们开始对一些基本问题进行探索，并从不同侧重点提出了几十种取证过程模型或框架。模型是过程的集合，为数字取证技术的发展提供简洁、一致、抽象、交互、易理解的基础。目前，数字取证的模型大致可以分为五类，分别是基本过程模型（Basic Process Model）、事件响应过程模型（Incident Response Process Model）、执法部门过程模型（Law Enforcement Process Model）、抽象过程模型（Abstract Process Model）和其他过程模型等。从结构上看这些模型包括线性式、层次式、瀑布式或循环式等，都用以描述数字取证的一般过程。

1.基本过程模型（Basic Process Model）

Dan Farmer 和 Wietse Venema 是最早系统地开展电子取证基本理论和方法体系研究的专家，他们在 1999 年提出了一个基本过程模型，主要内容包括：“保护和隔离（secure and isolate），记录现场（record the scene），系统查找证据（conduct a systematic search for evidence），证据的收集和包装（collect and package

evidence），维护保管链（maintain chain of custody）”。虽然此模型的过程粒度较粗糙，例如没有把取证准备作为其中一个阶段，但仍然为数字取证的发展起到了奠基作用。

延伸：Dan Farmer 和 Wietse Venema 所开发的取证工具集合 TCT（The Coroner's Toolkit，TCT）主要针对类 Unix 系统，在业界具有很大影响力。后继者 Brian Carrier 基于 TCT 开发了开源的命令行方式的 TSK（The Sleuth Kit），并进一步推出了 GUI（图形用户界面，Graphical User Interface）方式的电子取证平台 Autopsy（www.sleuthkit.org），到 2023 年 1 月的最新版本是 4.19.3，它在很多研究和训练中被采用。Brian Carrier 还开发了 Sleuth Kit Hadoop Framework 项目，是一个利用云计算帮助分析海量硬盘数据的项目。

2.事件响应过程模型（Incident Response Process Model）

具有代表性的是在 2001 年由 Chris Prosise 和 Kevin Mandia 提出的过程模型。在《应急响应：调查计算机犯罪（Incident Response: Investigating Computer Crime）》一书中，作者将数字取证作为安全紧急事件响应的重要环节，将其分为以下几个阶段：事前准备（Pre-incident Preparation）、事件侦测阶段（Detection of the Incident）、初始响应阶段（Initial Response）、响应策略制定（Response Strategy Formulation）、备份（Duplication）、调查（Investigation）、安全策略实施（Secure Measure Implementation）、网络监控（Network Monitoring）、恢复 （Recovery）、报告（Reporting）、后续追踪（Follow-up）等。这一模型明确提出了“事前准备（pre-incident preparation）”的概念，并将其作为取证流程的一个基本步骤，同时使得数字取证成为区别于其他调查方法，更加凸显专业性质的一项工作。

3.执法部门过程模型（Law Enforcement Process Model）

具有代表性的是在 2001 年由美国司法部（The U.S. Department of Justice，DOJ）提出的数字取证调查模型。在《电子证据取证检验：执法人员指南（Forensic Examination of Digital Evidence: A Guide for Law Enforcement）》”中，此模型将数字取证分为评估（Assessment）、获取（Acquisition）、检验（Examination）、记录（Documenting）、报告（Reporting）共 5 个阶段，其中的检验阶段又分为提取和分析，记录阶段要求所有的操作和观察都要规范地记录在案。这个模型面向的对象是执法人员。

4.过程抽象模型（Abstract Process Model）

包括统一过程模型（Harmonized Process Model）、抽象过程模型（ADFPM，Abstract Digital Forensics Process Model）、扩展的抽象模型等。ADFPM 由 Reith M, Carr C 和 Gunsch G 在 2002 年提出。这一模型包括识别（Identification）、准备（Preparation）、策略制定（Approach Strategy）、保存（Preservation）、收集（Collection）、检验（Examination）、分析（Analysis）、提交（Presentation）等 8 个阶段。这一模型具有较大影响力，为数字取证的基本方法和原理的进一步研究奠定了良好的基础。

5.其他过程模型（Ad hoc Model）

代表性的成果包括网络取证模型、IOT 取证模型、数据融合模型、增强的过程模型等。

1.2.3 数字取证的实施

数字取证作为调查犯罪、解决纠纷、事件响应等的重要活动之一，必须在法律规范下，利用科学验证的方式发现、固定、提取、分析证据材料，以获取符合证据可采标准的电子数据。上述的数字取证的模型从各种角度各种场景提出了不同的步骤和流程，我们可以从中归纳出数字取证在实施中的最一般的内容，包括以下 5 个方面。这些工作并不是一种流程，而是完成一个数字取证任务所必须考虑的各方面的因素。

1.准备与受理

准备工作要完成与取证相关的软件硬件设备与环境、人员职责分工、作业指导书、清单与文书，以及了解人员的技术能力等。受理工作包括获得必要的授权，对任务目标的了解，对取证性质的判断。在这一方面，需要考虑的因素还有：

（1）在完成取证任务的前提下，最小化对原始数字设备或潜在数据来源的需求；

（2）充分考虑取证操作可能带来的不利影响，并作出必要说明；

（3）与当前适用的法律规定的合规，例如涉及国家秘密、商业秘密、个人隐私的，如何保密的问题；

（4）权衡是否有能力与资源完成取证工作。

2.收集与获取

收集的任务是识别并加以妥善保管，即判断哪些载体包含潜在的电子数据，是在本地还是远程，在移动存储还是服务器存储等。收集的过程应该是系统化、清单式和标记性的，防止忽视、遗漏、混淆。收集的目标是硬件设备、存储媒介等。

获取的对象是数据，任务是产生电子数据的副本（整个磁盘、分区或选择的文件等），并记录方法和采取的操作。获取需要利用成熟可靠的方法和过程，提取电子化存储的信息（ESI, Electronically Stored Information），并在可行的条件下创建副本。获取的经验法则之一是决不在原始证据或证据源上执行调查分析操作，而要在获得的数据副本上进行。最好生成 2 个副本，一个用来分析调查，另外一个用作资料存档或质量控制。获取电子数据的工作要考虑场地、时间、类型等问题。场地是指在现场进行还是在移送到实验室进行，这取决于多个因素，例如案件类型、紧急程度、现场环境、数据种类与数量等。时间是指数据有效的长短期限，由数据的易失性顺序决定。类型是指如何获取及保存方式等。

按照获得数据的方法分类，获取的方法类型包括逻辑获取（Logical Acquisition）、稀疏获取（Sparse Acquisition）、比特流（Bit-Stream Imaging）获取等。逻辑获取是指仅收集案件调查需要的文件，稀疏获取是指仅收集数据存储区域未分配区的数据碎片，多用来恢复删除的信息，这两种类型只在取证调查人员对案情有足够了解的情况下才出现。比特流获取又称为克隆（Clone），是一种对整个存储设备进行逐位复制的方法，克隆的副本（称为镜像，Image）包含存储媒介的所有扇区和簇。取证调查人员通常对副本进行操作，以避免对原始存储媒介的污染。克隆方式一般有 2 种：磁盘到磁盘和磁盘到镜像文件。

按照获得数据的先后次序分类，获取的方法分为两类：静态获取（Dead Acquisition）和动态获取（Live Acquisition），或称为关机状态和开机状态的获取。静态获取主要针对文件系统数据、存储设备等，主要包括电子邮件、临时文件、文字处理文档、电子表格、引导扇区（Boot Sector）、松弛空间（Slack Space）、

未分配磁盘空间、各种删除的文件、系统注册表、事件/系统日志、Web 浏览器缓存、Cookie 和隐藏文件等。动态状态主要针对系统数据，包括当前系统配置、运行状态、日期和时间、在线数据、运行进程、登录用户、DLL 和共享库、交换文件和临时文件以及网络数据等，也包括路由表、ARP（地址解析协议）缓存数据、网络配置、网络连接等。

数据的易失性决定了采用哪种方法。按照 RFC3227，以下是一个典型系统的易失性顺序（Order of Volatility）的例子：注册表和缓存；路由表、进程表、内核状态、内存；临时系统文件；磁盘或其他存储媒介；远程和监控数据；物理配置、网络拓扑；存档媒介。由前到后，易失性降低，意味着在收集时的紧迫性优先等级降低。

按照获得数据的先后次序分类，获取的方法分为两类：远程获取和本地获取。远程获取（Remote Acquisition）是指通过网络实时远程计算机系统的数据，这种方法不同于本地的离线直接获取，会涉及相关的法律授权、访问授权以及在获取过程中无法使用只读锁等问题。本地获取是指要获取的数据就在现场的存储器中。在本地获取中，通常使用一种称为写保护（Write Blocker）或者只读锁（Read-only blocker）的硬件设备，把要调查的硬盘连接到取证计算机上，以防止对原始数据的修改。

3.计划与保护

计划是指按照程序规范和技术标准，结合案情和取证目标的不同，制定科学和可行的取证方法和操作流程。此部分需要考虑多个因素，包括规范与标准、软件和设备、环境与条件等，并进行实施前的评估。保护是指对于数字证据的载体的完整性和可用性，采取防止损毁、破坏和修改的一系列措施和机制。在任何一个案件调查中，这个工作都应该贯穿取证的全部环节，并在数字证据的整个生命周期进行维持。

证据监督链（Chain of Custody）或简称 COC，是实施有效保护的一种机制。证据监督链由一系列相关联的、针对特定设备的记录文档、照片、视频等构成，从收集阶段就开始建立，并与各个阶段的转换同步，直到数字证据的结果展现和应用，即数字取证任务的终结。建立并维持证据监督链的目的是为了在任何一个给定时间点，能够识别载有潜在数据证据的媒介的访问和移交等活动的有效性。证据监督链记录至少要包含下面的信息：

（1）唯一证据标识符；

（2）是谁在使用，呈现的状态、使用的时间和地点；

（3）是谁，以及何时交接证据；

（4）交接的原因和相关授权；

（5）任何不可避免的状态改变、对此负责的人以及引起改变的理由。

完整性检验是实现数据完整性保护的一种方法，在数字取证中得到广泛应用。在数字证据的移交、保管等过程中经常需要对电子数据完整性进行校验。基本方法之一是使用密码学哈希函数（Cryptographic Hash Function），或称为散列算法。例如在获取中生成的镜像文件，需要应用哈希函数计算该文件的哈希值（Hash Value），或称校验值，然后在后续检验分析等各个环节，可以随时再次计算其哈希值。如果哈希值不变，则可以认为镜像文件就是原始的文件，内容没有被修改；否则，就应该质疑该镜像文件是否发生了内容的变化。这种完整性的检验是由哈希函数本身的属性决定的。

在收集和获取工作中，要特别注意保护电子数据的完整性。在《关于办理刑事案件收集提取和审查判

断电子数据若干问题的规定》的第 5 条和第 6 条，列举了 5 种可以采用的方法：

（1）扣押、封存电子数据原始存储介质；

（2）计算电子数据完整性校验值；

（3）制作、封存电子数据备份；

（4）冻结电子数据；

（5）对收集、提取电子数据的相关活动进行录像。

还规定还明确了在初查过程中收集、提取的电子数据以及通过网络在线提取的电子数据，可以作为证据使用。

作业指导书（SOP，Standard Operating Procedure），或称标准作业程序、标准操作指南，指将某一事件的标准操作步骤和要求以统一的格式描述出来，用于指导和规范日常的工作。每个数字取证实验室或机构都要依据数字取证的国家或行业标准，结合本机构的实际情况，例如使用的设备、软件和环境，将数字取证任务进行切分、量化和细化，抽象出不同取证工作中的相同操作，制定出相对固定的作业指导书，目的是提升工作效率，规范操作行为，减少人为因素导致的取证结果偏差。机构应该根据科学发展和技术进步对数字取证带来的影响，定期对作业指导书进行系统的评估和审核，减少作业指导书因稳定性导致的滞后性。

4.检验与分析

这是数字取证中工作量最大的部分，包括数据恢复与文件雕刻、数据解析与重组、证据关联与提取、数据解密与保护、代码分析与功能验证、差异对比、数据过滤和汇总等。这些技术也是本书的主要内容。

检验与分析往往是交叉和循环的工作，以发现数字痕迹（Artifacts），回答需要调查的问题。根据案件和目标的不同，检验与分析工作采用的方法和手段也有很大区别。例如在识别和提取数字痕迹工作中，用到的基本方法包括但不限于：

（1）关键字搜索以定位包含特定串的文件。但对于某些情况，关键字搜索有时或失效，如加密文件、压缩文件、字符编码误用的文件、被插入格式标签的文本。

（2）文档检索（Document Retrieval）以定位特定话题的文件。

（3）元数据属性匹配（Metadata attribute matching）以定位符合满足元数据条件的文件。例如文件元数据中包含特定的修改时间。

（4）匹配给定的文件属性。例如利用搜索哈希值已知的文件。

（5）检验包含特定内容的文件，以发现需要的信息。例如通讯录文件。

（6）检验恢复的文件或数据记录，以发现被删除的内容。

在分析取证结果中有一些重要的考虑因素，例如：

（1）取证调查人员是否理解数字痕迹与案件或事件的关系。

（2）采取了哪些步骤或方法，以发现和消减可能出现的偏差。

（3）有无反取证（Anti-forensics）的因素导致取证难以进行。

（4）是否存在与系统时间相关的问题需要分析。

（5）数字痕迹或者对应的用户行为是否能与本地或外部来源关联？例如某个操作是来自计算机的本

地用户，还是来自远程黑客入侵或恶意代码。

在检验与分析中除了手工进行，还会借助各种不同的取证工具和软件，例如时间线（Time-line）分析工具能把事件列入一个时间序列，以帮助理解事件之间的顺序；关联分析工具能从大规模案件分析中帮助厘清实体之间的关系等。这些工具功能很强大，但可能并不完美，例如人脸或语音识别软件，在某些情况下结果会有明显的误差和偏离，甚至存在争议。所以取证调查人员应该非常清楚地理解这一点，确保以正确的方式、在合适的场景下使用取证软件和工具。

5.验证与总结

验证是指对取证技术、取证软件的功能和性能的测试和分析，包括对错误率与错误源及其影响的分析，以及如何有效地避免或减少等。国外的权威机构对此进行了大量的工作，例如，美国商务部的国家标准与技术研究院（National Institute of Standards and Technology，NIST）等，但国内的相关工作刚刚起步。

当取证结果出现明显异常时，可以利用多种软件提取或分析同一数据源，进行交叉验证，分析原因。这就涉及到验证取证软件或技术的可靠性问题。从对数字取证技术的验证（Validation）和审核（Verification）的角度看，有下面几种方法考查软件或技术算法的可靠性。

（1）检查数字取证技术用到的算法是否合适及其局限性；

（2）某种算法的一般用途是明确的，但实现细节可能未知，所以直接分析其细节通常不可行，但可以通过评估性能来测试其是否实现了目的；

（3）部分验证过程应该包含对可能出现错误的分析；

（4）需要测试算法实现中的错误以及在给定运行环境（硬件和操作系统）中的异常；

（5）测试一种技术在哪些条件下正常，哪些条件下不正常；

（6）在取证技术的实现中，突出的特点之一是其运行日志应该是丰富和完备的，以便于发现和跟踪问题。

总结是对取证调查结果的确认、复核与归档，以满足提交呈供电子数据的需要。总结的过程包括以客观通俗、严谨规范的语言描述事实，整理结果并进行分类归档和妥善保存，以作为呈供法庭的诉讼证据。总结的内容包括但不限于涉案电子设备、原始数字的检查结果；规范表达的日期和时间、硬盘的分区情况、操作系统和版本；数据信息和操作系统的完整性、计算机病毒评估情况、提取的相关文件种类、软件许可证以及对电子证据的分析结果和评估报告等所有相关信息。

1.3 数字取证的技术标准与规范

1.3.1 国家技术标准

国家标准是指由全国信息安全标准化技术委员会（www.tc260.org.cn）制定的标准。截至 2022 年 9 月，共推出 5 个标准，分别是：

1. GB/T 29360-2012 《法庭科学 电子物证数据恢复检验规程》

2. GB/T 29361-2012 《法庭科学 电子物证文件一致性检验规程》