伯克利

数字开源调查规程

关于有效利用数字开源信息调查违反国际刑法、人权法和人道法行为的实用指南

HUMAN RIGHTS CENTER
UC Berkeley School of Law

联合国人权
高级专员办事处

伯克利

数字开源调查规程

关于有效利用数字开源信息调查违反国际刑法、人权法和人道法行为的实用指南

HUMAN RIGHTS CENTER
UC Berkeley School of Law

联合国人权高级专员办事处

纽约和日内瓦，2023 年

HR/PUB/20/2
ISBN: 978-92-1-154248-6
eISBN: 978-92-1-005347-1
销售编号 : C.20.XIV.4

本作品由联合国代表联合国人权事务高级专员办事处 (人权高专办) 与加州大学伯克利分校法学院人权中心联合出版。

联合国文件编号由大写字母和数字构成。凡提及这种格式的编号，即指联合国某一文件。

封面图片来源：Ahmed Elgamal 使用人工智能平台 Playform 制作的深伪卫星图片。

加州大学伯克利分校法学院人权中心对以下捐赠者提供的财政支持表示感谢：西格丽德 • 劳辛信托基金、橡树基金会、加州大学伯克利分校的个人捐赠者、开放社会基金会和洛克菲勒基金会贝拉吉奥中心。

目 录

前 言

自 1990 年代初以来，数字工具和互联网，如同之前的照相机和电话一样，彻底改变了我们获取、收集和传播有关侵犯人权和其他严重违反国际法行为的信息，包括国际犯罪的信息。

今天，调查人员可以从大量公开的卫星图像、视频和照片，包括从智能手机上传到互联网的材料和社交媒体平台上的发帖中，获取可能存在的侵犯人权和其他严重违反国际法行为的数据，包括有关国际犯罪的数据。这一发展有助于调查人员绕过政府和其他传统的信息守门人，去接触甚至是实时获取到涉及不法行为的关键信息，否则这些信息将隐藏在公众视野之外。

然而，由于人权组织、政府间机构、调查机制和法院有时难以调整其工作做法，以纳入新的事实寻找和分析的数字方法，因此数字开源信息的使用基本上是临时性的。上述机构面临的最大挑战之一是在越来越多的在线信息中发现并核实相关材料，特别是用智能手机和其他移动设备拍摄的照片和视频，其中一些可能被篡改或被张冠李戴。

同时，国际性刑事法院和调查机制以及国家惩治战争罪单位的出现，进一步提升了对于获取、保存和分析可作为刑事审判证据的开源信息的通用标准的需求。为使开源信息作为证据被法庭采信，检察官和律师通常必须能够确立其真实性和保管链。适当地对这些材料进行的处理和加工，将大大增加检察官和律师使用这些材料的可能性。然而，如果收集和保存方式不当，则此等信息将不能被视为确证案件事实的可靠信息。如有明确标准用于评估开源信息作为联结因素或犯罪证据权重，将有益于法院和调查机制的工作。认证及核查的通用方法论标准同样将服务与人权实况调查任务，这些任务也越来越多地将数字开源材料纳入其调查工作。调查委员会、维和行动的人权部分、联合国人权事务高级专员办事处（人权高专办）的外地办事处以及联合国其他的人权监测和调查工作都将受益于健全的方法论原则和办法，以支持其调查结果的有效性和重要性。

为满足这一需求，我们两家机构，即加州大学伯克利分校法学院人权中心和人权高专办联合出版了《伯克利数字开源调查规程：关于有效利用数字开源信息调查违反国际刑法、人权法和人道法行为的实用指南》。本出版物最初源起于 2009 年的伯克利校园，当时人权中心将法律专家、技术专家、记者和活动人士聚集在一起，制定使用数字技术和方法以揭示和记录侵犯人权行为的战略。此后，人权中心召开了一系列跨学科研讨会，与广泛的技术、法律和方法学专家，包括来自人权高专办的专家合作，集思广益，开发新工具并确定和提炼相关准则、标准和方法，用于发掘、评估、核实和保存数字开源信息，以记录侵犯人权行为并将行为人绳之以法。这一进程与人权高专办制定指南和工具以支持联合国各调查委员会和实况调查团及人权高专办工作人员在实况调查和其他调查工作中越来越多使用开源信息的努力高度一致。

《伯克利规程》的制定得益于拥有不同专业视角、法律和文化背景、性别和国籍的个人作出的贡献，与专家进行了 150 多次咨商，包括联合国人权调查员在内的主要利益攸关方提供了投入。本规程还借鉴了人权高专办方法、教育和培训科以及国际刑事法院检察官办公室的专门工作组的专门知识。依照制定新方法的国际标准，人权高专办和人权中心对《伯克利规程》运用严格的审查、修订和核证程序。

在这种合作方式基础上，《伯克利规程》包含对涉嫌违反国际人权法以及国际人道法和刑法的行为进行在线研究的国际标准。它还提供方法和程序方面的指导，以采取专业、合法且合乎道德的方式收集、分析并保存数字信息。最后，《伯克利规程》规定了在线调查员可以采取的措施，其目的是保护在线调查员自身以及他人的数字、身体和社会心理安全，包括证人、受害人和第一反应人 (如公民、活动人士和记者) 等，这些人员冒着危及自身福祉的风险记录侵犯人权和严重违反国际法的行为。

《伯克利规程》追随联合国两项更早规程的步伐，即《关于调查潜在非法死亡的明尼苏达规程》(1991 年，2016 年更新) 和《酷刑和其他残忍、不人道或有辱人格的待遇或处罚的有效调查和文件记录手册》(《伊斯坦布尔规程》) (1999 年，2004 年更新)。《明尼苏达规程》是由 1980 年代参与寻找失踪人员的律师和法证科学家制定的，确立对可疑死亡或无人在场的死亡进行法医学调查的国际标准和程序，并作为评估此类调查可信度的一种手段。同样，《伊斯坦布尔规程》为医疗从业人员和律师提供指导，说明如何识别和记录酷刑造成的身体和社会心理后遗症，以便文件记录可以作为法庭或包括人权调查和监测在内的其他情形下的有效证据。所有这三项规程都建立一个信念上，即科学、技术和法律可以而且必须合力为人权服务。与之前的规程一样，《伯克利规程》将以联合国正式语言提供，以促进其在全世界的使用和效用。

我们希望，在一个日益数字化的世界里，《伯克利规程》将帮助在线调查人员——无论是法律专业人士、人权维护者、记者还是其他人——制定和实施有效流程，以记录并核实违反国际人权法及国际人道法和刑法的行为，最好地利用数字开源信息，以便将应对此等违法行为负责之人公平地绳之以法。

Eric Stover
加州大学伯克利分校
法学院人权中心
教务主任

Michelle Bachelet
联合国人权事务
高级专员

执行摘要

开源调查是指全部或部分依靠可公开获得的信息，对涉嫌的不法行为进行正式和系统的在线调查。今天，大量可公开获取的信息可以通过互联网获取，快速演变的数据场景带来了新的信息类型和来源，有助于调查涉嫌侵犯人权和严重国际罪行。调查此类指控的能力对于无法及时进入犯罪现场的调查人员特别有价值，这种情况在国际调查中很常见。

开源信息可以提供线索，支持情报产出，并在法庭上作为直接证据。然而，为使开源信息能用于正式调查进程，包括用于法律调查、实况调查任务和调查委员会，调查人员必须采用连贯一致的方法，这既能加强其调查结果的准确性，又能让法官和其他实况调查人员更好地评估调查进程本身的质量。制定《伯克利数字开源调查规程》的目的是给国际刑事司法和人权领域的调查员提供国际标准和指导。这些调查员来自一系列机构，包括媒体、民间社会团体和非政府组织、国际组织、法院以及国家调查机构和国际调查机构。制定一致、可衡量的标准来支持这一多学科领域，是使开源调查实践专业化的一种手段。

虽然关于特定工具和软件的使用指南和培训是提高数字开源调查质量的重要组成部分，但《伯克利规程》关注的重点不是具体的技术、平台、软件或工具，而是即便技术本身发生变化也能持续适用的基本原则和方法。这些原则概要规定了进行有效开源调查的最低法律和伦理标准。调查人员遵循《伯克利规程》的指导，将有助于确保其工作质量，同时最大限度地减少对自己和他人可能造成的人身、社会心理和数字风险。

《伯克利规程》被设计为一个教学工具和供开源调查员使用的参考指南。导言部分之后的三个章节专门介绍总体框架，包括原则、法律考虑因素和安全。其余章节则侧重于调查过程本身。《伯克利规程》的这一部分从关于准备工作和战略规划的章节开始，然后用一章专门介绍所需的各种调查步骤，即在线查询、初步评估、收集、保存、核实和调查分析。最后一章阐明报告开源调查结果的方法和原则，作为结束。

撰稿人和参与者

伯克利规程协调委员会

Lindsay Freeman, 高级法律研究员，加州大学伯克利分校法学院人权中心

Alexa Koenig, 执行主任，加州大学伯克利分校法学院人权中心

Eric Stover, 教务主任，加州大学伯克利分校法学院人权中心

伯克利规程编辑委员会

Sareta Ashraph, 高级法律顾问；大律师，Garden Court Chambers 事务所；前高级分析员，促进对达伊沙 / 伊拉克和黎凡特伊斯兰国所犯罪行追究责任的联合国调查组

Alix Dunn, 执行主任，The Engine Room 组织

Richard Goldstone, 前法官，南非宪法法院；前首席检察官，前南斯拉夫问题国际法庭和卢旺达问题国际法庭

Brenda J. Hollis, 共同检察官 (国际)，柬埔寨法院特别法庭；前首席检察官，塞拉利昂问题余留事项特别法庭

Tanya Karanasios, 方案主任，WITNESS 组织

Enrique Piracés, 媒体和人权方案主管，卡内基梅隆大学人权科学中心

Beth Van Schaack, 人权问题客座教授，斯坦福法学院；战争罪行问题无任所大使的前副手，美国国务院全球刑事司法办公室

Michel de Smedt, 司长，国际刑事法院检察官办公室调查司

Alan Tieger, 高级检察员，科索沃专家检察官办公室；前高级出庭律师，前南斯拉夫问题国际法庭

Christian Wenaweser, 列支敦士登常驻联合国代表；前主席，《国际刑事法院罗马规约》缔约国大会

Alex Whiting, 调查负责人，科索沃专家检察官办公室；实务教授，哈佛法学院；前检察事务协调人和调查事务协调人，国际刑事法院检察官办公室

Susan Wolfinbarger, 外交事务干事和分析团队负责人，美国国务院；前高级项目主任，地理空间技术项目，美国科学促进会

伯克利规程咨询委员会

Federica D'Alessandra, 执行主任，牛津国际和平与安全项目，牛津大学；国际公法和公共政策小组出版物《民间社会记录严重侵犯人权行为手册：原则和最佳做法》编辑

Stuart Casey-Maslen, 荣誉教授，比勒陀利亚大学法学院；《关于调查潜在非法死亡的明尼苏达规程》撰稿人 (2016 年)

Alison Cole, 人权问题专家顾问，新西兰内务部

Francoise Hampson, 名誉教授，埃塞克斯大学法学院；布隆迪问题调查委员会成员

Christof Heyns, 人权法教授，比勒陀利亚大学；人权委员会成员；曾任法外处决、即决处决或任意处决问题特别报告员；《关于调查潜在非法死亡的明尼苏达规程》(2016 年) 协调人

Vincent Iacopino, 高级医疗顾问，人权医生组织；《酷刑和其他残忍、不人道或有辱人格的待遇或处罚的有效调查和文件记录手册》(《伊斯坦布尔规程》) 的主要撰稿人

Kelly Matheson, 高级律师和项目管理人，WITNESS 组织；《视频作证实地指南》的作者

Hanny Megally, 阿拉伯叙利亚共和国问题独立国际调查委员会专员；高级研究员，纽约大学国际合作中心

Juan Méndez, 人权法驻校教授，华盛顿法学院； 曾任酷刑和其他残忍、不人道或有辱人格的待遇或处罚特别报告员；调查采访和程序保障通用规程协调人

Aryeh Neier, 名誉会长，开放社会基金会

Navi Pillay, 会长，国际反死刑委员会；曾任联合国人权事务高级专员；前法官，国际刑事法院；前庭长，卢旺达问题国际法庭

Paulo Sérgio Pinheiro, 阿拉伯叙利亚共和国问题独立国际调查委员会主席；曾任布隆迪人权状况特别报告员；曾任缅甸人权状况特别报告员

Thomas Probert, 特别讲师，比勒陀利亚大学人权中心；研究助理，剑桥大学治理与人权中心；《关于调查潜在非法死亡的明尼苏达规程》(2016 年) 撰稿人

Stephen Rapp, 杰出研究员，美国大屠杀纪念博物馆西蒙－斯克约防止灭绝种族中心；前战争罪行问题无任所大使，美国国务院全球刑事司法办公室；前检察官，塞拉利昂问题特别法庭

Cristina Ribeiro, 调查协调人，检察官办公室，国际刑事法院

Patricia Sellers, 国际刑事法院的检察官性别平等问题特别顾问；客座研究员，牛津大学凯洛格学院；前法律顾问和出庭律师，前南斯拉夫问题国际法庭和卢旺达问题国际法庭

研讨会参与人

新法证学研讨会：利用开源信息调查严重犯罪 (2017 年，意大利贝拉焦)

Hadi Al Khatib, 叙利亚档案

Stuart Casey-Maslen, 比勒陀利亚大学

Yvan Cuypers, 国际刑事法院

Scott Edwards, 国际特赦组织

Lindsay Freeman, 加州大学伯克利分校法学院人权中心

Alexa Koenig, 加州大学伯克利分校法学院人权中心

Steve Kostas, 开放社会正义行动

Andrea Lampros, 加州大学伯克利分校法学院人权中心

Kelly Matheson, WITNESS 组织

Félim McMahon, 国际刑事法院

Julian Nicholls, 国际刑事法院

Thomas Probert, 剑桥大学

Cristina Ribeiro, 国际刑事法院

Gavin Sheridan, Vizlegal 组织

Eric Stover, 加州大学伯克利分校法学院人权中心

Alan Tieger, 前南斯拉夫问题国际法庭

Mark Watson, 国际正义与问责委员会

Guy Willoughby, 战争罪研究协会

构建开源调查的伦理框架研讨会（英国埃塞克斯大学，2019 年）

Fred Abrahams, 人权观察

Leenah Bassouni, 加州大学伯克利分校法学院人权中心

Federica D'Alessandra, 牛津大学

Sam Dubberley, 国际特赦组织

Jennifer Easterday, JustPeace Labs

Scott Edwards, 国际特赦组织

Lindsay Freeman, 加州大学伯克利分校法学院人权中心

Geoff Gilbert, 艾塞克斯大学

Christopher "Kip" Hale, 国际正义与问责委员会

Evanna Hu, Omelas

Gabriela Ivens, Mozilla 研究员和 WITNESS 组织

Alexa Koenig, 加州大学伯克利分校法学院人权中心

Matt Mahmoudi, 剑桥大学

Lorna McGregor, 艾塞克斯大学

Daragh Murray, 艾塞克斯大学

Vivian Ng, 艾塞克斯大学

Enrique Piracés, 卡内基梅隆大学人权科学中心

Zara Rahman, The Engine Room 组织

Sasha Robehmed, The Engine Room 组织

Ilia Siatitsa, 国际隐私组织

人权高专办代表, 来自方法、教育和培训科

开源调查产生的法律问题圆桌会议（2019 年，海牙）

David Akerson, 促进对达伊沙 / 伊拉克和黎凡特伊斯兰国所犯罪行追究责任的联合国调查组

Sareta Ashraph, Garden Court Chambers 事务所

Danya Chaikel, 科索沃专家检察官办公室

Alan Clark, 国际刑事法院

Federica D'Alessandra, 牛津大学

Nico Dekens, Bellingcat 调查机构

Chris Engels, 国际正义与问责委员会

Lindsay Freeman, 加州大学伯克利分校法学院人权中心

Emma Irving, 莱顿大学

Michelle Jarvis, 协助调查和起诉自 2011 年 3 月以来在阿拉伯叙利亚共和国境内犯下国际法所规定最严重罪行者的国际公正独立机制

Edward Jeremy, 国际刑事法院

Ashley Jordana, Global Rights Compliance 基金会

Sang-Min Kim, 加州大学伯克利分校法学院人权中心

Alexa Koenig, 加州大学伯克利分校法学院人权中心

Nicholas Koumjian, 缅甸问题独立调查机制

Bastiaan Van Der Laaken, 协助调查和起诉自 2011 年 3 月以来在阿拉伯叙利亚共和国境内犯下国际法所规定最严重罪行者的国际公正独立机制

Dearbhla Minogue, Global Legal Action Network

Nick Ortiz, 莱顿大学

Matevz Pezdirc, 欧盟刑事司法合作局反灭绝种族罪网络

Sanja Popovic, 科索沃专家检察官办公室

Steven Powles, Doughty Street Chambers 事务所；国际律师协会战争罪委员会

Stephen Rapp, 美国大屠杀纪念博物馆西蒙一斯克约防止灭绝种族中心

Cristina Ribeiro, 国际刑事法院

Mark Robson, 国际正义与问责委员会

Brad Samuels, SITU 调查

Dalila Seoane, Civitas Maxima 组织

Carsten Stahn, 莱顿大学

Melinda Taylor, 国际刑事法院

Alan Tieger, 科索沃专家检察官办公室

Raquel Vázquez Llorente, 酷刑目击者组织

其他专家评审员

Elise Baker, 加州大学伯克利分校法学院人权中心

Sean Brooks, 加州大学伯克利分校长期网络安全中心

Stephanie Croft, 加州大学伯克利分校法学院人权中心

Sam Dubberley, 国际特赦组织

Thomas Ewing, 高级防务研究中心

Christopher "Kip" Hale, 国际正义与问责委员会

Gabriela Ivens, 人权观察

Felim McMahon, 加州大学伯克利分校法学院人权中心

Daragh Murray, 艾塞克斯大学

Yvonne Ng, WITNESS 组织

Zara Rahman, The Engine Room 组织

Mark Robson, 国际正义与问责委员会

Justin Seitz, Hunchly

Andrea Trewinnard, 加州大学伯克利分校法学院人权中心

Steve Trush, 加州大学伯克利分校长期网络安全中心

Raquel Vázquez Llorente, 酷刑目击者组织

特别致谢

特别感谢国际刑事法院检察官办公室在线调查工作组成员。

还要感谢人权高专办的许多同事，因为他们的努力，这份联合出版物才得以完成。*

* 依照人权高专办的政策，为本出版物提供的文稿不归属于受人权高专办雇用的人员。

缩写和简称

HTML	超文本标记语言
ICRC	红十字国际委员会
信通技术	信息和通信技术
IP	因特网协议
ISP	互联网服务提供商
NGO	非政府组织
人权高专办	联合国人权事务高级专员办事处
PDF	可携式文件格式
URI	统一资源标识符
URL	统一资源定位符
VPN	虚拟专用网络

导　言

本章摘要

- 宗旨
- 受众
- 定义

1. 《伯克利数字开源调查规程》阐述了在识别、收集、保存、分析和展示数字开源信息以及在国际刑事和人权调查中使用此类信息时应适用的专业标准。开源信息是任何公众成员均可观察、购买或请求获得的信息，不需要通过特殊法律地位或未经授权的方式获取。数字开源信息是公开可用的数字格式信息，通常从互联网上获取。数字开源信息包括用户生成的数据和机器生成的数据，可包括例如，在社交媒体上发布的内容；网站和信息共享平台上的文件、图像、视频和音频记录；卫星图像；以及政府发布的数据。[1] 数字开源调查是基于数字开源信息开展的调查。为便于阅读，本规程下文将数字开源信息和调查分别称为“开源信息”和“开源调查”。

2. 虽然在调查中使用开源信息并非新现象，但由于利用互联网和其他数字资源共享信息的情况日益增多，包括社交媒体数量激增，公开来源的数量和多样性都在扩展。本规程既探讨了在处理数字信息时面临的复杂性问题，也探讨了在评估信息来源和验证来自开放在线论坛的信息时遇到的独特挑战。

3. 如今越来越多的国际刑事和人权调查员利用互联网协助开展工作，但目前并没有关于开源调查的通用参考、准则或标准。本规程力图通过规定原则和做法填补这一空白，这些原则和做法将帮助调查员按照专业标准开展工作，并酌情促进保存开源信息，供问责机制使用。

4. 本规程特别关注为确保国际司法和问责制而进行的开源调查，其中大体包括：人权文件编制、信息保存、证据收集和真相调查；调查委员会和实况调查团进行的调查；[2] 其他类型国际授权的调查和询问；[3] 真相与和解进程；民事诉讼；以及包括国际刑事诉讼在内的刑事审判。由于开源调查有助于开展不同类型的努力以确保问责，[4] 所以本规程中概述的方法和文件要求可能比新闻和人权倡导等其他领域历来所用方法和文件要求更加严格。无论调查出于何种目的，开源调查人员通过遵守本规程所述、依照共同法律标准制定的方法原则，将确保开展高质量的工作，并最大限度地利用在法院、法庭和其他程序中收集的信息，进而确保问责。

1 本清单并非详尽无疑。

2 调查委员会和实况调查团是政府或国际组织为调查各种问题而设立的机构。调查委员会或实况调查团报告事实调查结果，得出法律结论并提出建议。虽然国际调查委员会或实况调查团的调查结果不具法律约束力，但可以产生很大影响。不过，在某些司法管辖区，国家调查委员会的调查结果可能具有约束力。关于国际调查委员会和实况调查团的进一步信息，见人权理事会，“国际调查委员会、人权委员会、实况调查团和其他调查”。可查阅 www.ohchr.org/zh/hr-bodies/hrc/co-is.

3 例如，见联合国人权事务高级专员根据人权理事会第 39/1 号决议提交的关于委内瑞拉玻利瓦尔共和国人权状况的报告 (A/HRC/41/18)。另见人权理事会第 41/2 号决议，其中理事会请高级专员编写一份关于菲律宾人权状况的报告。

4 例如，缅甸问题国际独立实况调查团在核查过程及调查结果和结论中，除使用第一手资料和其他信息外，还使用了开源信息。实况调查团的最后报告 (A/HRC/42/50) 是促使人权理事会设立缅甸问题独立调查机制的一个因素，该机制已获授权开展司法调查。实况调查团还获得授权向缅甸问题独立调查机制移交资料，包括移交开源调查的内容。冈比亚向国际法院起诉缅甸违反《防止及惩治灭绝种族罪公约》的案件也立足于实况调查团的报告。这表明为某一目的收集的信息最终可能对另一个法律问责程序有所帮助。

5. 此外，本规程着重强调对违反国际法行为（包括侵犯人权行为）和违反国际刑法行为（包括战争罪、危害人类罪和灭绝种族罪）进行调查的标准。而且本规程提供的指导可适用于其他类型的调查，包括国家法院或地市法院开展的调查。

6. 最终，本规程旨在协助开源调查人员按照与法律要求和道德规范基本一致的专业方法开展工作。本规程还旨在帮助调查过程的不同最终用户，包括律师、法官和其他决策者，更好地理解和评价开源调查技术。本规程同样旨在为有经验的从业人员提供资源，并为希望了解如何对涉嫌违反国际法的行为进行开源调查的人员提供培训和教学工具。[5]

A. 宗旨

7. 虽然调查人员长期以来一直依赖开源信息，但在二十世纪早期至中期才加快步伐有系统地利用此类信息，重点是从外国电台广播和报纸中提取情报。[6] 随着1990年代万维网的出现，以及2000年代初社交媒体和智能手机的普及，开源信息的数量和质量发生了巨大变化。如今，任何拥有智能手机和能够访问互联网的个人都可以在全球范围内创建和发布数字内容，尽管这些内容的质量、真实性和透明度各不相同。数据量的增长以及此类数据传输与共享的速度加快为开源调查人员收集和分析有关国际犯罪和侵犯人权行为的信息创造了新机会。而与此同时，内容创建者现在可以相对容易地传播虚假信息和操纵数字数据。本规程试图应对这一新环境以及在处理这些机遇和挑战时遇到的复杂性问题。

8. 开源信息对各类调查都很有裨益，但在国际刑事和人权调查中发挥的作用尤为关键。之所以如此，有几点原因。首先，国际授权的调查，包括由联合国调查委员会和实况调查团开展的调查，或者由国际刑事法院授权的调查，都有赖于允许进行调查的法律和政治程序。[7] 因此这些调查通常在事件发生后很久才进行。第二，国际调查往往可能无法进入被调查事件发生的实际地点，例如，由于一国拒绝合作或拒绝准许进入。第三，即使获准进入某一地区或领土，调查人员实际进入相关地点的机会也可能有限，或可能因保护方面的考虑而被阻止进行现场调查或面谈。最后，大多数调查人员在指称罪行或违法行为发生地没有充分的执法权，因此可能无法收集必要信息。即使在有国家合作的情况下，跨境取证也可能是一个艰巨的过程，受到繁琐官僚程序的羁绊。所有这些因素都表明为什么开源调查技术不仅强大而且必要，既可以远程开展调查，又可以在事件发生的同时进行调查。

9. 本规程针对在不同环境中工作的各类调查人员群体，他们的任务、调查权力和资源各不相同。因此，规程采取了

[5] 本规程还提供了开源调查的一些模板和一个术语表（见下文第八章）。

[6] Nikita Mehandru and Alexa Koenig, “ICTs, social media, & the future of human rights”, Duke Law & Technology Review, vol. 17, No. 1, p. 129.

[7] 联合国授权的调查委员会和实况调查团是由安全理事会、大会、人权理事会和秘书长等设立。就国际刑事法院而言，检察官办公室可根据缔约国或安全理事会移交的案件发起调查，或自行发起并经法官授权进行调查。

灵活办法，不预设调查人员以相同方式开展工作，而是根据每个独特的工作环境酌情调整方法。此外，由于协助开展开源调查的技术、工具和技能不断演变，本规程关注的重点不是会发生变化的具体工具、平台、网站、软件或来源，而是应当用于指导开源调查的基本原则和程序。

10. 本规程旨在制定标准化程序，为不同的调查、机构和司法管辖区域提供方法指导，从而帮助开源调查人员了解以下事项的重要性：

(a) 在可能情况下，追踪在线内容的出处，确定其原始来源；

(b) 评价在线来源的可信性和可靠性；

(c) 核实在线内容并评估其真实性和可靠性；

(d) 遵守法律要求和道德规范；

(e) 将自身、所在组织和第三方受到伤害的任何风险降至最低；

(f) 加强对信息源的人权保护，包括对隐私权的保护。

B. 受众

11. 本规程的目标受众包括为确保司法和问责而通过识别、收集、保存和（或）分析开源信息对国际犯罪或侵犯人权行为进行调查的个人和组织。这包括在以下机构工作的调查人员、律师、档案员和分析人员：国际、区域和混合刑事法庭、国家战争罪机构、调查委员会、实况调查团、独立调查机制、国际组织、过渡时期司法机制及非政府组织。其他可能的受益者还有对违反国际法的行为进行司法和准司法开源调查的各类国际和区域机制的工作人员。[8] 本规程还可能对数字信息的第一反应人具有指导意义，如社区组织和独立研究人员，他们通常最先公布基于开源信息的调查结果，其工作往往对于确立其他正式授权的开源调查至关重要。目标受众还包括支持受害人对施害者个人或国家提起民事求偿的个人和组织。本规程通常还可帮助那些根据开源调查得出事实结论或法律结论的人员，使他们更好地对所依据或所评价的任何开源调查内容作出评估。

12. 其他潜在利益攸关方可能包括社交媒体平台等基于网络的服务提供方，他们存储大量数据并在数据保存中发挥关键作用，还可能包括为支持开源调查技术和流程提供软件的开发人员。

C. 定义

13. 为给开源调查提供实用的标准和指导，调查人员必须对特定术语形成共同理解。本节阐明了规程通篇所用关键术语的含义，包括经常混淆的术语之间的区别。[9]

[8] 例如，见人权理事会特别程序的来文和访问报告。可查阅 www.ohchr.org/zh/special-procedures-human-rights-council. 另见安全理事会设立的制裁委员会的工作。可查阅 www.un.org/securitycouncil/zh/content/repertoire/sanctions-and-other-committees.

[9] 关于相关术语和定义的更全面汇编，见第八章。

1. 开源信息与闭源信息

14. 开源信息包括任何公众成员均可观察、购买或请求获得的公开可用信息，无需特殊法律地位或未经授权的访问。闭源信息是指访问受限或访问受法律保护的信息，[10] 但可通过私人渠道（如司法程序）合法获取或由信息源自愿提供。尽管定义简单，但从在线内容中确定什么构成开源信息比初看上去要复杂得多。互联网上有越来越多的数据未经所有者同意而被公开，如因存在安全漏洞而遭黑客窃取、泄露和暴露的信息，或未经适当许可由第三方发布的信息。尽管这些信息公开可用，因而可在形式上视为开源，但某些类型的最终用途可能受到法律和道德的限制。此外，数字信息可由掌握专业技术技能和受过专业培训的人取得，他们可以访问普通人无法访问或不太可能访问的网络和数据。[11] 仅能从暗网上获取的信息就是一个实例，暗网是互联网中只能通过 Tor 浏览器等特定软件访问的部分。[12] 尽管暗网因提供匿名服务使其对非法活动具有吸引力，但在大多数国家，使用 Tor 浏览器和搜索暗网是合法行为。只要信息不涉及未经授权而访问，本规程就将此类信息归入“开源”范畴。最明显的区别是，开源信息不涉及与互联网个人用户互动或向其索取信息。[13] 通过与其他互联网用户沟通而从他们那里获取的信息被视为闭源信息。

15. 数字开源信息 [14] 是指互联网上的开源信息，可通过公共网站、互联网数据库或社交媒体平台访问。下文介绍获取开源信息的不同方式。

2. 获取数字开源信息

(a) 观察

16. 许多平台上的内容仅需使用任意数量的免费网络浏览器前往相关网站即可获得。其他在线平台则需用户登录或注册才能访问和查看内容。只要这些过程在司法管辖区内对所有用户开放且访问相关内容系属合法，并且在访问或查看时没有违反隐私或安全控制，这些内容即可视为开源。不过，一些符合这项定义的内容也可能不属于开源信息，例如特许信息、机密信息或其他受法律保护的信息。在这种情况下，尽管任何公众成员均可看到此类信息，但这些信息在司法程序中作为证据的用途可能受到限制。依赖此类材料还可能引起道德伦理或方法上的关切，例如无法确定内容出处或核实内容。

[10] 例如特许信息和机密信息。

[11] 有些行为可能违反网站的服务条款，但本身并不违法。例如，违反网站的服务条款收集数据是未经授权的行为，可能会导致被禁止使用该网站。

[12] 暗网是指互联网中只能通过专门软件访问的那一部分。Tor 浏览器就是这种软件的一个例子。

[13] 尽管从私人数据库购买信息或向公共政府机构提交获取信息的申请需要某种程度的在线交流，但这通常是一个自动过程，与此处所述同互联网其他个人用户进行的互动类型不同。

[14] 开源信息在本规程中也可称为在线内容、在线材料或在线数据。

(b) 购买

17. 开源调查的一些数据源位于付费平台上，或者采用免费和增值相结合的模式，即额外的功能和数据访问要收取财务费用。越来越多的企业收集公共数据并提供免费和付费的数据访问服务。开源调查人员认为有用的许多信息都存在于只有付费才能访问的数据库中和平台上。就本规程而言，开源信息包括向全体公众成员提供的付费服务，但不包括仅限特定群体（如执法人员或有执照的私人调查员）访问的服务。

(c) 申请

18. 此处“申请”一词是指任何个人均可根据信息自由或信息获取方面的法律向国家机构提出获取公共信息的申请。申请不是指要求个人、公司或组织自愿交出其信息，而仅限于向有法律义务以相同方式回应所有人的国家实体提出申请。开源调查可能引起其他在线调查活动，例如使用信息收发服务、聊天室、论坛或电子邮件与外部信息源互动接触。此类互动接触不在本规程所述开源调查的范围内。

3. 开源情报

19. 开源情报是开源信息的一个子类别，是为协助制定政策和作出决策的特定目的而收集、使用，最常见于军事或政治环境中。虽然开源信息包括任何人均可合法获取的所有公开可用信息，但开源情报是这些信息的子集，“为满足特定情报需求而及时收集、利用并向适当受众传播”。[15] 在国际刑事和人权案件中，开源情报用于为决策职能提供背景信息，例如，为安全相关活动（如保护进入现场的证人和团队成员或跟踪相关人员）提供信息，而不用于履行与调查过程（如确定各种犯罪的要素）有关的信息收集职能。

4. 开源调查

20. 开源调查是指利用开源信息履行信息收集和证据收集职能。

5. 开源证据

21. “证据”一词应与“信息”相区别。[16] 不同司法管辖区域通常将证据定义为在调查中使用或在审判等司法审讯中提出的事实证明。开源证据是具有证据价值的开源信息，可被采纳用于在法律诉讼中确立事实。在泛指“信息”时，不得误用也不得过度使用“证据”一词，这一点非常重要。

6. 开源信息与开源软件

22. “开源”一词通常用于描述可自由使用和重新发布的软件或代码，此类软件或

[15] National Open Source Enterprise, Intelligence Community Directive No. 301, 11 July 2006, p. 8（脚注省略）。

[16] Federica D'Alessandra and others, eds., *Handbook on Civil Society Documentation of Serious Human Rights Violations: Principles & Best Practices* (The Hague, Public International Law and Policy Group, 2016), p. 17.

代码不受版权、专利或其他法律控制的限制。开源软件在源代码的基础上建构，任何有访问权限的人均可加以检查、修改和加强。[17] 开源软件通常对用户并不可见，但可由计算机程序员进行调整和修改。尽管开源调查人员经常使用开源软件和工具寻找、收集、保存和分析开源信息，但开源软件与开源信息并不相同。

7. 可信性与可靠性

23. 就国际刑事审判中的证词证据而言，法官会评估“证人的可信性”及“其证词的可靠性”。[18] 对于联合国调查委员会和实况调查团开展的调查以及类似调查，指导意见规定“访谈者应评估受访者的可信性和可靠性”。[19] 指导意见阐明，“评价将考虑信息对调查主题的相关性；还将考察信息源的可靠性以及信息的有效性或真实性。”[20] 本规程中这些术语的用法如下：

(a) “可信性”是指可以相信或值得信赖；

(b) “可靠性”是指一贯、可靠或按预期运作的能力；

(c) “真实性”或“有效性”是指准确、诚实或与事实相符。

17 见 Opensource.com，“What is open source?”。

18 International Criminal Court, *Prosecutor v. Bosco Ntaganda*, Case No. ICC-01/04-02/06, Judgment of 8 July 2019, para. 53.

19 人权高专办，《国际人权法和人道法调查委员会和实况调查团：指导意见和实践》(2015 年，纽约和日内瓦)，第 52 页。可查阅 www.ohchr.org/Documents/Publications/CoI_Guidance_and_Practice.pdf.

20 同上，第 59 页。

二

原 则

本章摘要

- 为遵守与数字开源调查相关的专业原则，调查人员必须确保做到负责、称职和客观，并确保各项工作依法开展，同时充分考虑到安全关切。
- 调查人员还必须考虑在整个调查周期各个阶段使用的方法。相关的方法原则至少包括准确性、数据最少化、数据保存和安全设计。
- 最后，所有调查人员都应以伦理道德考量为指引。至少要考虑到保护所有参与调查或与调查有关的个人的尊严，以及确保谦逊、包容、独立和透明。

24. 虽然开源调查使用的技术、工具和技能会发生变化，但某些总体方法原则和伦理原则应当一以贯之。确定这些原则是努力实现开源调查领域专业化的重要步骤。以下各项原则对于确保开源调查的质量至关重要，进而将增强开源调查的可信性、可靠性和潜在效用，以期确保问责并最大限度减少可能对不同利益攸关方造成的伤害。

A. 专业原则

1. 问责

25. 开源调查人员必须对其行为负责，而清晰的文件编制、记录保存和监督通常可以保证这一点。保持调查方法和程序透明是确保问责的重要因素。因此，在可能且合理的范围内，开源调查人员应保留活动记录。开源调查的各个步骤，从识别相关材料到收集、分析和报告，都应连贯清晰地记录在案。任何参与收集或处理在线信息的个人都应认识到其方法可能受到质疑，包括可能被传唤出庭作证。开源调查的文件记录既可手动完成，也可使用各种软件提供的自动化流程完成。只要文件记录连贯一致并足够详尽，手动或自动方法均可使用。自动化流程和软件必须能让用户理解，也能由用户或开发人在法庭上作出解释。此外，开源调查人员应记录工作过程中使用的任何工具或软件。

2. 能力

26. 开源调查人员必须经过适当培训并具备适当技术技能，以执行所从事的各项活动。他们必须以专业和道德的方式开展在线活动，避免盗用他人的工作成果，(在安全且参与者希望的情况下) 对所有参与调查者给予认可，准确报告数据，包括确认在线内容中可能存在的任何差距。开源调查人员和调查流程也必须保持灵活，跟上新的发展步伐，并酌情采用新技术和新技能。此外，各组织和调查组应建立机制，确保程序得到连贯执行和遵守。

3. 客观

27. 客观是适用于所有调查的基本原则，无论是在线调查还是离线调查。开源调查人员应理解个人、文化和结构性偏见可能对其工作产生影响，并应理解有必要采取对策以确保客观性。开源调查人员必须确保客观地进行调查，提出和运用多种工作假设，而非倾向于用任何特定理论阐释其处理的案例。客观性对于在线进行的开源调查尤为重要，这是由互联网上信息的结构和呈现给用户的方式所决定。即使基本查询相同，但所用浏览器、搜索引擎、检索词和句法都可能导致查询结果差异悬殊。搜索引擎和网站使用的互联网架构和算法中的固有偏见可能危及搜索结果的客观性。[21] 搜索

[21] 见 Safiya Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York, New York University Press, 2018);Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York, Picador, 2019)。

结果也可能受到许多技术因素的影响，包括所用设备及其位置，以及用户以往的搜索历史和互联网活动。开源调查人员应采用各种方法抵消此类偏见，确保搜索结果尽可能多样化，例如，进行多项搜索查询并使用各种搜索引擎和浏览器。[22] 调查人员应认识到，搜索结果也可能受到其他因素的影响，如受到数据环境差异的影响，在这种情况下，在线信息可能不均衡地来自特定社会群体或社会阶层。[23] 最后，调查人员应始终努力了解并纠正自身偏见，无论是有意识的偏见还是存在于潜意识中的偏见。[24]

4. 合法

28. 开源调查应遵守适用的法律，这意味着调查人员需对适用于其工作的法律有基本了解。特别是，调查人员应了解数据保护法和受国际人权法保护的隐私权。[25] 即使信息可能已公开，但这并不意味着信息的收集和使用不涉及隐私问题。开源调查人员必须考虑其行为对隐私的影响，包括个人在不同数字空间对隐私的合理期望。调查人员还应该意识到马赛克效应，即如果有足够多包含类似或互

22 例如见 Paul Myers, “How to conduct discovery using open source methods”, in *Digital Witness, Using Open Source Information for Human Rights Investigation, Documentation and Accountability,* Sam Dubberley, Alexa Koenig and Daragh Murray, eds. (Oxford, Oxford University Press, 2020)(讨论搜索引擎和搜索词的选用如何导致开源调查的结果产生偏差)。

23 例如见 Alexa Koenig and Ulic Egan, “Hiding in plain site: using online open source information to investigate sexual violence and gender-based crimes”, in *Technologies of Human Rights Representation,* James Dawes and Alexandra S. Moore eds. (即将出版)(讨论妇女相对缺乏使用智能手机的机会以及性暴力和性别暴力幸存者在网上使用密语会如何导致与此类犯罪相关的开源信息的数量和可获得性下降，还讨论了技术相关职位和战争罪调查员普遍由男性担任如何负面影响到自动和 (或) 人工检测进程找出性别犯罪相关开源信息的可能性)。关于偏差的进一步讨论，见下文关于道德原则的第二章 C 节和下文关于数据环境评估的第五章 B 节。

24 例如见 Forensic Science Regulator, *Cognitive Bias Effects Relevant to Forensic Science Investigations, FSR-G-217* (Birmingham, United Kingdom, 2015)(讨论可能对调查质量产生负面影响的各类认知偏差，包括期望偏差、证实偏向、锚定、情境偏差以及角色和重建效应)；Wayne A. Wallace, *The Effect of Confirmation Bias on Criminal Investigative Decision Making* (Minneapolis, Walden University ScholarWorks, 2015)(对证实偏向加以解释：调查人员寻找或相信支持自己所偏好案件理论的信息，“同时忽略或忽视相反的证据”)；Michael Pittaro, “Implicit bias within the criminal justice system”, Psychology Today, 21 November 2018 (讨论能总体影响刑事调查的偏见，并提出已知的消除偏见的技巧)；Jon S. Byrd, “Confirmation bias, ethics, and mistakes in forensics”, Forensic Pathways, 21 March 2020 (讨论可能扭曲法证分析的各种认知和道德误差，以及避免这些误差的技巧)。另见 Yvonne McDermott, Daragh Murray and Alexa Koenig, “Digital accountability symposium: whose stories get told, and by whom? Representativeness in open source human rights investigations”, Opinio Juris, 19 December 2019 (讨论开源调查的方法如何对“所报告的侵害行为类型、有机会发声的受害人和证人、以及对大规模侵犯人权行为事件的叙述构建方式”产生负面影响)；由 Yvonne McDermott 牵头的题为“人权调查的未来：利用开源情报改变侵犯人权行为的记录和发现”的项目。

25 《世界人权宣言》第十二条规定，任何人的私生活、家庭、住宅或通信不得任意干涉，他或她的荣誉和名誉不得加以攻击。人人有权享受法律保护，以免受这种干涉或攻击。《公民及政治权利国际公约》第十七条规定，任何人之私生活、家庭、住宅或通信，不得无理或非法侵扰，其名誉及信用，亦不得非法破坏。第十七条还规定，对于此种侵扰或破坏，人人有受法律保护之权利。

补信息的数据集被公布或合并，那么即使是匿名的公开数据也可能易于重新识别身份。[26] 此外，调查人员应意识到，在一些司法管辖区域，对个人进行持续不断的在线监测，或系统收集并长期保留个人数据，可能需要额外的许可和保障措施，因为此类活动会引起更严重的隐私问题。[27]

5. 安全意识

29. “安全设计”[28] 针对的是调查及任何附带活动的架构和基础设施，而安全意识原则侧重于个人在工作过程中必须考虑的因素，特别是对其在线行为的意识。所有进行在线调查的个人都应具备基本的操作安全意识，以确保他们尽量减少自己的数字痕迹，并认识到潜在的风险。进行开源调查的组织应确保其调查人员接受信息安全培训，以便了解自己可能面临的风险，并了解信息安全的三大核心支柱：(a) 保密性 (例如只允许已获许可的用户查阅数据)；(b) 完整性 (确保数据不被未经授权的用户篡改或以其他方式更改)；(c) 可用性 (确保系统和数据可在获授权用户需要时供其使用)。培训还应该关注互联网的治理结构。在开始在线调查活动之前，应进行威胁和风险评估，并应定期审查评估结果，必要时进行修正。安全是每个人的责任，而不仅是信息技术部门或安全风险管理人员的责任。

B. 方法论原则

1. 准确性

30. 出于方法论和道德操守的要求，必须仅依靠可信材料以确保调查的准确性，因而也能确保质量。开源调查人员在其调查过程中以及陈述任何结果时，应尽可能真实准确，特别是在承认基础数据或整个案件的欠缺方面。通常通过使用并测试多个暂定假设和 (或) 同行审议来提高准确性，这两种方法都有助于尽量减少在选择、解读和呈现数据时有偏见的可能性。不应夸大或夸张分析性结论。使用清晰、客观、基于事实的语言并避免使用情绪化语言，能保护调查及其结果的实际客观性，并让人们感知到这种客观性。

---

26 “马赛克效应的概念源自情报收集的马赛克理论，在该理论中，不同的信息片段尽管单独使用时效用有限，但在与其他类型的信息结合时变得有意义 (Pozen 2005)。马赛克效应的概念应用到公共使用的数据时，表明即使是孤立看来无害的匿名数据，如果公布了足够多包含类似或互补信息的数据集，也可能变得易于重新识别身份。”见 John Czajka and others, *Minimizing Disclosure Risk in HHS Open Data Initiatives* (Washington, D.C., Mathematica Policy Research, 2014), appendix E, p. E-7。可查阅 https://aspe.hhs.gov/system/files/pdf/77196/rpt_Disclosure.pdf. 另见 David E. Pozen, “The mosaic theory, national security, and the Freedom of Information Act”, *Yale Law Journal,* vol. 115, No. 3 (December 2005), pp. 628-679。

27 例如，在大不列颠及北爱尔兰联合王国，法律规定“为……执法目的而处理的个人数据的保存时间不得超过该数据处理目的所需的时间”(Chapter 12 of the Data Protection Act 2018, part 3, chap. 3, sect. 39 (1))。根据欧洲议会和欧盟理事会 2016 年 4 月 27 日关于在处理个人数据和自由转移此类数据方面保护自然人，同时废除第 95/46/EC 号指令的第 2016/679 号条例 (《通用数据保护条例》)，只能为“特定、明确和合法的目的”收集个人数据，必须仅限于为收集目的所必需的信息，并且只能在收集目的所必需的时间内保持可识别性 (第 5-6 条)。

28 见下文第 33 段。

2. 数据最少化

31. 数据最少化原则规定，只有满足下列条件才可收集和处理数字信息：(a) 出于可阐明的目的，具有正当理由；(b) 为实现该目的所必需；(c) 与实现该目的的能力相称。[29] 在开源调查方面，只有在涉及具体调查时才应收集在线内容。这一原则倾向于逐项手动收集，而不是批量自动收集，但同时指出后一种方法在某些情况下可能是适当的。将这一原则应用于在线内容的收集有助于避免过度收集，这一点的重要性有几个原因。在使用自动收集流程时，过度收集问题特别引起关切，因其可能会造成或加剧安全漏洞，[30] 当过度收集导致调查人员不了解其所掌握的信息有哪些类型时，尤其如此。如果自动化流程不根据信息的内容类型对其进行区分，过度收集还可能会引起隐私和数据保护问题。最后，避免过度收集的实际目的是最大限度降低储存费用并防止在调查周期不同阶段形成下游瓶颈，例如审查阶段、分析阶段，如果调查导致法律诉讼，还包括披露阶段。

3. 保全

32. 防止对相关信息收集不足与避免过度收集同样重要。对网上的信息而言，这一点尤其值得关注，原因是网上信息的持久性和可获得性往往不稳定。保全原则旨在避免收集不足，以便有价值和可能有证明力的证据不致丢失。例如，社交媒体平台可能会删除违反其服务条款的内容，即使该内容对调查人员有潜在价值。除非及时向平台提出保全请求，或由调查人员以其他方式保存内容，否则这些信息可能会永远丢失。此外，用户可以选择删除或编辑自己的内容，使曾经公开的信息无法查看。另外，互联网上的信息很容易被抽离情境、丢失、删除或损坏。如果数字资料要在未来的问责机制中保持可访问性和可用性，就需要在短期和长期积极认真地保存这些资料。[31]

4. 安全设计

33. 安全设计原则要求尽可能确保数字信息和在线操作在默认设置下是安全的。开展在线开源调查的组织应投资于并实施适当的技术和结构措施，以确保在默认设置下，基础设施（包括硬件和软件）在调查人员上网时适当匿名化且不可溯源。所有设备都应配备最新的软件以防止恶意软件，并进行适当的隐私和安全设置。安全措施应在在线调查活动开始前到位；应根据需要不断监测、更新和调整这些措施。调查人员、调查团队或组织应安排持续测试，包括渗透测试，[32] 以确保其安全系统按设计运行。

[29] 本规程从欧洲联盟《通用数据保护条例》借鉴了数据最少化原则，但对其进行了调整，以适应开源调查的情况（见《条例》第 5 条）。

[30] 关于安全漏洞的例子，见下文关于安全的第四章。

[31] 详情见下文关于保全的第六章 D 节。

[32] 渗透测试是为了测试系统的安全性而授权进行的模拟网络攻击。

C. 道德原则

1. 尊严

34. 在进行调查时，应意识到并敏感地注意任何与尊严有潜在关联的问题，特别是受国际人权法保护的利益。例如，调查人员应坚持不歧视原则，因为这可能影响到调查的内容以及调查的实施者或负责人人选，并应纳入关于证人、幸存者、其他调查人员、被告和其他可能受负面影响者的数字安全、身体安全和心理社会安全的保障措施。遵守尊严原则也可能影响到调查中的哪些内容，包括书面和任何视觉材料可以公开分享，举例而言，如非必要，不全面显示痛苦或暴力内容。这一原则确保人权规范成为开展符合道德操守的开源调查的一套指导性标准。

2. 谦逊

35. 开源调查人员应该谦逊，认识到自己的局限性，并意识到自己对哪些情况不了解。要正确了解和解读开源信息，可能需要专门培训或向专家咨询。谦逊也意味着对错误负责。如果调查人员发现自己犯了错误，应该纠正这一错误，或将错误报告给那些能够将所造成伤害降到最低的人。理想情况下，特别是对公开且广泛传播的调查，应该有一个报告错误和发布更正的机制。

3. 包容

36. 开源调查人员必须确保将各种观点和经历纳入调查。可能影响在线调查整体包容性的考虑因素包括调查的地理范围、被调查的违法行为和（或）国际罪行以及是否认识到关于社会不同阶层的网上信息存在不均衡性。[33] 调查团队也应保持多样性，包括性别上保持平衡。此外，包容原则和尊严原则可能会影响到调查人员在调查中对材料收集和使用的选择，以及向不同受众呈现这些材料的方式。

4. 独立

37. 开源调查人员应该保护自己及其调查免受不当影响，应查明并避免任何切实存在或被认为存在的利益冲突，并制定保障措施减缓无法避免的冲突。程序、方法和供资的透明化有助于评估独立性，并保护调查的实际独立性和可被感知的独立性。

5. 透明

38. 问责原则要求调查人员的方法和结果保持透明，而道德方面的透明原则涉及开源调查人员在网上和对外界的行为方式。这意味着避免不实陈述。[34] 尽管匿名和不可溯源，包括使用虚拟身份，[35] 可能对安全而言很重要，但调查人员应意识

[33] 见下文关于数据环境评估的第五章 B 节。

[34] 例如，以虚假借口试图加入封闭小组或在社交媒体上建立联系。

[35] 关于虚拟身份的讨论，见下文关于基础设施相关考虑因素的第四章 C 节。

到不实陈述可能带来负面影响，例如损害调查、团队或组织的声誉和可信度，或污染已收集的信息。通过不实陈述获取信息可能会侵犯目标对象的隐私权并(或)给调查带来污点，尤其在相关司法管辖区域规定不实陈述非法的情况下。

三

法律框架

本章摘要

- 确定适用哪些法律对于决定拟收集的内容和最佳收集方式至关重要。适用的法律因调查人员的身份、调查对象的身份、调查目的以及调查人员、调查对象、数据和法律程序所在的司法管辖区而有所不同。
- 采取维护数字材料真实性和记录保管链的方式保存数字材料，会让数字材料更有可能作为证据被法庭采信。
- 确定调查的类型及其最终目标（例如刑事诉讼、民事诉讼、过渡期正义程序等）将决定适用的证据门槛。
- 侵犯个人隐私权可能导致证据被排除。

39. 开源调查人员必须了解其开展工作所处的法律框架，包括了解与其调查有关的适用部门法以及调查活动所在的司法管辖区的法律框架。了解适用于调查的实体法，包括潜在违法行为[36]或犯罪的要素，以及责任模式，[37]可以使调查更有针对性，使收集的信息和得出的分析结论更可能有助于确保伸张正义和追究责任的努力。同样，了解相关司法管辖区的程序法和证据规则能让调查员在开展工作时采取适当方式，以符合对开源信息用于诉讼程序的相关要求。

40. 对于国际刑事调查，法律框架由相关法庭、法院或法院系统的法定文书作出规定。[38]对于国际授权的调查，例如调查委员会，确立调查的机制除规定其他因素外，还规定适用的部门法律以及调查的地理范围和时间范围。[39]对于其他调查，包括非政府组织进行的调查，调查实体本身可确定自己的法律框架。[40]

41. 本章旨在帮助开源调查人员更好地领会和理解其工作的潜在最终用途并相应地调整其调查技术。鉴于适用法律因司法管辖区域、调查类型和调查实体的法律授权而不同，以下各节概述了调查潜在违反国际法行为时的主要考虑因素。建议在可行的情况下，调查人员从熟悉相关司法管辖区和事由的律师那里获得专家法律咨询意见。

A. 国际公法

42. 本规程重点关注存在大量重叠的三类国际公法：国际人道法、国际人权法和国际刑法。这三类法律相辅相成；实际上，

36 例如，在调查仇恨言论和煽动暴力行为时，调查人员应了解达到《公民及政治权利国际盟约》第二十条第二款规定的高门槛的行为类型。见关于禁止构成煽动歧视、敌意或暴力罪的鼓吹民族、种族或宗教仇恨言论的拉巴特行动计划 (A/HRC/22/17/Add.4，附录)，第 11 和 29 段，及其基于人权的门槛测试，该测试有 32 种语文版本。可查阅 www.ohchr.org/zh/freedom-of-expression. 关于仇恨言论，见联合国消除仇恨言论战略和行动计划 (2019 年)。可查阅 www.un.org/en/genocideprevention/hate-speech-strategy.shtml.

37 在刑法中，可根据相关法规所界定的若干责任模式，追究行为人的责任。这些责任模式包括直接和间接实施、共同实施、协助和教唆以及指挥责任。见 Jérôme de Hemptinne, Robert Roth and Elies van Sliedregt, eds., *Modes of Liability in International Criminal Law* (Cambridge, United Kingdom, Cambridge University Press, 2019)。

38 例如见国际刑事法院，《程序和证据规则》(2013 年)；前南斯拉夫问题国际法庭，《程序和证据规则》(2015 年 7 月 8 日)；卢旺达问题国际刑事法庭，《程序和证据规则》(2015 年 5 月 13 日)；塞拉利昂问题余留事项特别法庭，《程序和证据规则》(2018 年 11 月 30 日)；黎巴嫩问题特别法庭，《程序和证据规则》(2019 年 4 月 10 日)；柬埔寨法院特别法庭，《内部规则》(2011 年 8 月 3 日)。

39 例如，2019 年 9 月成立的委内瑞拉玻利瓦尔共和国问题国际独立实况调查团的任务授权是调查自 2014 年以来发生的法外处决、强迫失踪、任意拘留以及酷刑和其他残忍、不人道或有辱人格的待遇等行为，并向人权理事会提交一份调查结果报告 (人权理事会第 42/25 号决议，第 24 段)。成立于 2011 年的阿拉伯叙利亚共和国问题独立国际调查委员会的授权任务是调查自 2011 年 3 月以来在阿拉伯叙利亚共和国境内发生的所有涉嫌违反国际人权法的行为，查证可能构成这种违反行为和犯罪行为的事实和背景情况，并在可能的情况下查明责任人 (人权理事会 S-17/1 号决议，第 13 段)。2017 年派往刚果民主共和国开赛地区的一队国际专家的任务授权是收集和保存关于开赛地区侵犯和践踏人权的指控以及违反国际人道法的指控的信息，并将此调查结论递交刚果民主共和国司法当局 (人权理事会第 35/33 号决议，第 10 段)。

40 包括非政府组织在内的一些组织往往有自己的内部方法，要求其注重某一特定法律领域，例如涉及酷刑或性暴力和性别暴力的法律，这些法律也将为调查重点提供指导。

国际人道法和（或）国际刑法的适用并不免除各国根据国际人权法应履行的义务。下文概述了每个实践领域，包括其法律渊源和不同实践领域之间的区别，以便开源调查人员了解在工作中应遵循哪些参考资料的指导。

1. 国际人道法

43. 国际人道法即“武装冲突法”规范的是如何开展敌对行动，并解决冲突中出现的人道主义问题，这可能是国际性冲突或是非国际性冲突。[41] 国际人道法在武装冲突开始时被触发，并一直延续到实现和平为止，但武装冲突与和平的分界并不总是清晰或明确。[42] 国际人道法的主要渊源是 1899 年和 1907 年海牙公约、[43] 1949 年 8 月 12 日的日内瓦四公约[44] 及其 1977 年的附加议定书，[45] 以及规范某些类型武器使用的若干条约。[46] 习惯法也是国际人道法的一个重要来源，因为习惯法填补了条约留下的空白。习惯国际人道法对冲突各方都有约束力，对非国际性武装冲突尤其重要，因为其相关规则比基于条约的国际人道法规则更详细。[47] 直到 1990 年代初，国际人道法的主要执行机制是国家军事法庭，各国在这些法庭上追究其本国士兵和军官的责任。随着国际刑事法庭的兴起，某些严重违反国际人道法的行为在这些法庭的成立规约中被定为战争罪，[48] 为在

41 国际性武装冲突与非国际性武装冲突之间的区别基于两个因素：所涉各方的结构和地位。国际武装冲突的参与方是主权国家。相比之下，非国际性武装冲突的参与方是国家和有组织的武装团体。见 Andrew Clapham, Paola Gaeta and Marco Sassòli, eds., *The 1949 Geneva Conventions, A Commentary* (Oxford, Oxford University Press, 2015), chaps. 1 and 19。

42 国际性冲突的起始相对明确，因其触发因素是两个国家之间任何使用武力的行为，非国际性武装冲突的起始则不那么直观。非国际性武装冲突只有在武装团体组织严密并且暴力达到一定强度时才存在，而这两个因素需要逐案进行详细的事实分析。见 Sylvain Vité, “Typology of armed conflicts in international humanitarian law: legal concepts and actual situations”, *International Review of the Red Cross*, vol. 91, No. 873 (March 2009), pp. 72 and 76-77。关于武装冲突何时算结束、何时算实现和平也存在争议。停火或和平协议虽可以帮助证明武装冲突已结束，但不是决定性的。人们提出了武装冲突结束的各种检验标准，即：达成总体和平协定后军事行动立即全面结束，存在和平解决方案，认定存在冲突的标准停止存续。见 Nathalie Weizmann, “The end of armed conflict, the end of participation in armed conflict, and the end of hostilities: implications for the detention operations under the 2001 AUMF”, *Columbia Human Rights Law Review*, vol. 47, No. 3 (2016), pp. 221-224。

43 分别是《关于陆战法规和习惯的公约》（《海牙第二公约》）和《陆战法规和惯例公约》（《海牙第四公约》）。

44 见《改善战地武装部队伤者病者境遇之日内瓦公约》（《日内瓦第一公约》）；《改善海上武装部队伤者病者及遇船难者境遇之日内瓦公约》（《日内瓦第二公约》）；《关于战俘待遇之日内瓦公约》（《日内瓦第三公约》）《关于战时保护平民之日内瓦公约》（《日内瓦第四公约》）。

45 见 1949 年 8 月 12 日《日内瓦四公约关于保护国际性武装冲突受难者的附加议定书》（第一议定书）；1949 年 8 月 12 日《日内瓦四公约关于保护非国际性武装冲突受难者的附加议定书》（第二议定书）。

46 例如见《关于禁止细菌（生物）及毒素武器的发展、生产及储存以及销毁这类武器的公约》、《禁止或限制使用某些可被认为具有过分伤害力或滥杀滥伤作用的常规武器公约》、《关于禁止发展、生产、储存和使用化学武器及销毁此种武器的公约》、《关于禁止使用、储存、生产和转让杀伤人员地雷及销毁此种地雷的公约》、《集束弹药公约》。另见 International Committee of the Red Cross (ICRC), “Weapons”, 30 November 2011。可查阅 www.icrc.org/en/document/weapons.

47 见 ICRC, “Customary international humanitarian law”, 29 October 2010。可查阅 www.icrc.org/en/document/customary-international-humanitarian-law-0. 另见 ICRC, “Welcome to the Customary IHL Database”。可查阅 https://ihl-databases.icrc.org/customary-ihl/eng/docs/home.

48 例如：《国际刑事法院罗马规约》第 8 条将国际人道法写入战争罪的定义。

国际一级执行国际人道法提供了一条新途径。一些国家还将战争罪编入国家立法，[49] 以便可在其常规法院系统内审理此类案件，而不是在军事法院审理。国内案件可在冲突发生国审理，也可以根据普遍管辖原则在其他国家审理，后一种情况越来越多。[50] 一些国家设立了专门的战争罪部门来起诉这类案件。国际刑事法庭和国家法院为国际人道法判例的不断增加作出了贡献，这些判例也是重要的法律渊源，其规则可能具有约束力，取决于司法管辖区域。

2. 国际人权法

44. 根据国际法，各国有尊重、保护和实现人权的义务和责任。1948 年通过的《世界人权宣言》为国际人权法奠定了基础。尽管《宣言》是一种期望，不具有法律约束力，但其中一些条款已成为习惯国际法的一部分。[51] 《宣言》还促成了两项公约和许多人权条约。[52] 各国只受其已签署、批准公约和条约的约束，除非这些文书所载规范已获得习惯国际法的地位。[53] 国际人权法也已成为许多国际性刑事法庭法定框架的组成部分。此外，根据国际公约设立了若干区域人权法院，负责裁决起诉公约缔约国违反国际人权法的案件，其中包括非洲人权和民族权法院、[54] 欧洲人权法院 [55] 和美洲人权法院。[56] 区域一级还有其他人权机构，包括非洲人权和民族权委员会、欧洲社会权利委员会和美洲人权委员会，所有这些机构都在持续发展国际人权法的判例法。

---

49 例如见：澳大利亚 (经修正的 1945 年《战争罪法》，第 7 节)；波斯尼亚和黑塞哥维那 (《刑法》第 171-184 条)；肯尼亚 (2008 年《国际犯罪法》，第 6(1)(c) 和 (2)-(4) 条)；新西兰 (2000 年《国际罪行和国际刑事法院法》，第 11 节)；南非 (2012 年《日内瓦四公约执行法》)。

50 根据“普遍管辖权”，国家法院可以起诉在境外犯下危害人类罪、战争罪、灭绝种族罪和酷刑罪等严重违反国际法罪行的个人，所依据的原则是，这些罪行损害国际社会和国际秩序本身，而各国可以采取行动保护国际社会和国际秩序。见国际正义资源中心，“普遍管辖权”。可查阅 https://ijrcenter.org/cases-before-national-courts/domestic-exercise-of-universal-jurisdiction.

51 许多国家、联合国官员和学者都指出，《世界人权宣言》的大多数条款 (即使不是全部) 构成习惯国际法。具体而言，人们公认，《世界人权宣言》规定禁止奴役、任意剥夺生命、酷刑、任意拘留和种族歧视的条款构成习惯国际法。见 Hurst Hannum, “The status of the Universal Declaration of Human Rights in national and international law”, *Georgia Journal of International and Comparative Law,* vol. 25, No. 1 (1996), pp. 322-332 and 341-346。

52 见《消除一切形式种族歧视国际公约》、《公民及政治权利国际公约》、《经济社会文化权利国际公约》、《消除对妇女一切形式歧视公约》、《禁止酷刑和其他残忍、不人道或有辱人格的待遇或处罚公约》、《儿童权利公约》。关于联合国核心人权条约的更多资料，见人权高专办，“核心国际人权文书及其监测”。见 www.ohchr.org/zh/node/100004.

53 习惯国际法是指由既定国际惯例产生的国际义务，而不是由正式书面公约和条约产生的义务。它源于各国出于法律义务感而遵循的普遍、一贯的做法。强行法是习惯国际法的一个基本组成部分，系指国际法的某些基本的首要原则。例如见 Legal Information Institute, “Customary international law” and “Jus cogens”, Cornell Law School。可查阅 www.law.cornell.edu/wex.

54 根据《非洲人权和民族权宪章》(《班珠尔宪章》) 设立。

55 根据《保护人权与基本自由公约》(《欧洲人权公约》) 设立。

56 根据《美洲人权公约》(《哥斯达黎加圣何塞公约》) 设立。

45. 国际组织在习惯国际人权法的发展和标准制定方面也发挥着关键作用。[57] 联合国人权事务高级专员办事处（人权高专办）以及其他国际实体就有助于标准制定和软法发展的法律领域发表专题报告。人权条约机构[58]提供报告、[59]判例法[60]和其他形式的指导，包括有助于拟订和理解各自条约所载条款的一般性评论和一般性建议。[61] 同样，人权理事会特别程序在国际人权法标准设定规范的演变过程中发挥了作用，[62] 其他机制也是如此，包括实况调查团和调查委员会等。

46. 与国际人道法的情况类似，国际人权法已成为许多国家法律框架的组成部分，这些国家或是遵循一元论法律传统在国家领域内直接适用国际义务，或是采取将国际法直接融入国内立法或运用普遍管辖权的方式，从而发展了关于国际人权法的重要法理学。[63]

3. 国际刑法

47. 国际刑法在和平时期和武装冲突期间都适用，对犯下战争罪、危害人类罪和灭绝种族罪等国际法上罪行的个人追究刑事责任。[64] 这些罪行有时统称为“暴行罪”[65]或“严重国际罪行”，大多已列入《罗马规约》，该规约被广泛认为反映了习惯国际刑法。国际刑法还包括一些未列入《罗马规约》的罪行，如恐怖主义行为。[66] 国际刑法和跨国刑法相关领域可能有一些重叠，跨国刑法将跨境贩运人口、毒品、武器和其他非法货物

[57] 国际组织的例子包括国际刑事法院、国际移民组织和禁止化学武器组织以及人权机制，例如人权理事会的特别程序和调查委员会或类似机构。特别程序针对联合国所有会员国执行任务，不依赖于会员国对某个特定条约的批准。这些人权机制的法律规范和机制各不相同，收集资料的方法和标准也存在差异。例如：任意拘留问题工作组的主要工作方法是向有关个人、其家属或代表、政府、非政府组织和国家机构收集关于具体案件的资料。然后，工作组调查来文所报告的案件，包括为此进行国家访问。关于工作组的最新工作方法，见 A/HRC/36/38。相比之下，调查委员会则由人权理事会临时设立，通常根据其任务规定自行开展调查，为此常常开展国家访问，在访问期间，除其他外，进行证人访谈。例如见布隆迪问题调查委员会的职权范围。可查阅 www.ohchr.org/Documents/HRBodies/HRCouncil/CoIBurundi/TermsofreferenceCOIBurundiENGL.pdf.

[58] 例如见人权高专办，“人权条约机构”。可查阅 www.ohchr.org/zh/node/101034.

[59] 报告可以采取结论性意见的形式，条约机构据此审议缔约国和其他利益攸关方提交的关于缔约国履行特定条约义务情况的报告。一些条约机构也能够发布调查报告。例如见消除对妇女歧视委员会，“调查程序”。可查阅 www.ohchr.org/zh/treaty-bodies/cedaw/inquiry-procedure.

[60] 条约机构针对具体案件发表关于个人申诉的意见。大体上见人权高专办，“人权条约机构——个人来文”。可查阅 www.ohchr.org/zh/treaty-bodies/human-rights-treaty-bodies-individual-communications.

[61] 见人权高专办，“人权条约机构——一般性评论”。可查阅 www.ohchr.org/zh/treaty-bodies/human-rights-treaty-bodies-general-comments.

[62] 大体上见人权高专办，“人权理事会特别程序”。可查阅 www.ohchr.org/zh/special-procedures-human-rights-council.

[63] Amnesty International, *Universal Jurisdiction: A Preliminary Survey of Legislation Around the World – 2012 Update* (London, 2012), pp. 1-2.

[64] Robert Cryer, Darryl Robinson and Sergey Vasiliev, *An Introduction to International Criminal Law and Procedure*, 4th ed. (Cambridge, United Kingdom, Cambridge University Press, 2019), chap. 15.

[65] 虽然“族裔清洗”一词未列入《罗马规约》，也不是国际法定义的独立罪行，但被认为属于“暴行罪”的范畴。在这方面，请见联合国，“暴行罪分析框架：预防工具”，第 1 页。可查阅 www.un.org/en/genocideprevention/documents/about-us/Doc.3_Framework%20of%20Analysis%20for%20Atrocity%20Crimes_EN.pdf.

[66] 见安全理事会第 1757(2007) 号决议，附件，附文（《黎巴嫩问题特别法庭规约》），第 2 条。

等行为定为犯罪。[67] 与国际人道法和国际人权法不同的是，国际刑法的重点是个人刑事责任，而不是国家责任。国际刑法案件可在国家刑事法院、混合刑事法庭、[68] 国际性刑事法院或法庭[69]（包括国际刑事法院或行使普遍管辖权的国内法院）审理。国际刑法的渊源包括法院和法庭的组成文件（如安全理事会决议、规约、程序和证据规则以及法院条例）以及对国际犯罪行使管辖权的国家的国内法。国际刑法的另一个重要渊源是判例法，判例法可具有约束力或劝导力，视具体的司法管辖区而定。[70]

B. 管辖权和问责

48. 管辖权是一个法律术语，系指授予法律实体（如法院或法庭）规定、裁决和执行法律的权力。本规程中关于司法和问责的定义非常宽泛，指不同类型的司法和非司法程序。追究国际犯罪和违反国际人权法及（或）国际人道法行为的责任，可通过刑事、民事或行政性质的法律程序，也可通过不具法律约束力的程序，如国际人权调查报告（包括调查委员会和实况调查团的报告）以及其他过渡期正义机制（包括着重调查真相的举措）。调查人员应尽可能考虑到可能据以追究责任的管辖权范围。

49. 开源调查人员应确定与其工作相关的问责机制，并确定所收集的证据在何处可以或可能得到采信用于确证事实。然而，在国际调查的初期阶段可能并不了解或不清楚这方面的情况。如果罪行发生地国没有正常运作的司法系统，或者国际社会尚未充分着手调查此事，情况尤其如此。此外，不可能预测未来可能相关的所有管辖权。开源调查人员如果不了解具体机制或管辖权，则应尽力收集和保存信息，以便在最广泛的潜在相关管辖权范围内最大限度地利用信息。调查人员如果了解案件最终审判地点的相关要求，则应根据这些具体要求调整其程序。

50. 管辖权可以通过以下方式确定：

(a) 属地管辖权系指法院审理与既定地域范围内所发生行为有关的案件之权力。对国际性法庭而言，属地管辖权通常限于已批准法庭创设条约的各国领土；

(b) 属时管辖权系指法院审理规定时间内据称发生的行为所涉案件的权力；

(c) 属人管辖权系指法院对诉讼程序中某一当事方拥有作出裁决的权力；

(d) 属事管辖权系指法院审理特定类型案件或与特定事项有关案件的权力；

[67] Cryer, Robinson and Vasiliev, *An Introduction to International Criminal Law and Procedure*, chap. 15.

[68] 这一术语除其他外包括柬埔寨法院特别法庭、塞拉利昂问题特别法庭、黎巴嫩问题特别法庭、科索沃问题特别法庭和检察官办公室以及中非共和国特别刑事法庭。

[69] 这一术语包括常设国际刑事法院和前南斯拉夫问题特设国际法庭、卢旺达问题国际刑事法庭和国际刑事法庭余留事项国际处理机制。

[70] 见 Rosa Theofanis, “The doctrine of res judicata in international criminal law”, *International Criminal Law Review*, vol. 3, No. 3 (2003)。

(e) 普遍管辖权系指法院所主张的审理被告人的权力，而不论被控罪行实施地点，也不论被告人国籍、居住国或与起诉机构存在任何其他关系。

C. 调查权力和职责

51. 正式调查权系指法律赋予特定实体在特定管辖范围内进行调查的权力。与司法管辖权的限制非常相似，司法或检察机构只能在法律授权的范围内进行调查。[71] 调查权力可包括传唤证人、通过传票要求出示记录和执行搜查令的能力。法律可能要求调查实体遵循严格的程序，或者在某些情况下，调查实体可以决定自己的程序。[72]

52. 调查违反国际法行为的大多数其他机构一般不具有调查权力或可强制执行的证据收集手段，如传票或搜查令。因此，这些机构可能完全依赖开源信息和自愿提供的信息，如各种文书、数字化文件和证人证词。

53. 一般而言，调查权力伴随着明确的职责。[73] 一些调查人员可能没有警察权力或其他法律权力，但建议所有调查人员尽可能遵守法律调查人员的主要职责，以确保调查质量。法律调查人员和检察官的共同职责和义务包括调查定罪和免罪情形的职责、保护证人职责、保全证据职责、确保诉讼公正的职责和尊重被告人权利的义务。

54. 在刑事审判中，检察官也有义务向辩护方披露相关信息和证据。[74] 这不仅包括审判时得到采信的证据，还包括调查过程中收集的任何证明有罪或无罪的信息，包括与证人可信度有关的信息。[75] 对于特许不予披露的信息或可能使人面临风险的信息，存在某些例外情况。法院可命令不披露可能因披露而受到危害的被害人或证人的身份，但这一点无法保

71 见 Justia，“Agency investigations”。可查阅 www.justia.com/administrative-law/agency-investigations.

72 同上。

73 例如，《罗马规约》第五十四条界定了检察官的调查职责和权力，确定检察官除其他外进行调查、收集和审查证据、约谈被害人和证人并与各国和国际组织合作的能力。

74 例如见前南斯拉夫问题国际法庭，《程序和证据规则》，规则 66 (A)；卢旺达问题国际刑事法庭，《程序和证据规则》，规则 66 (A)；黎巴嫩问题特别法庭，《程序和证据规则》，规则 110 (A)。

75 例如见国际刑事法院，《程序和证据规则》，规则 76-84；前南斯拉夫问题国际法庭，《程序和证据规则》，规则 66 (A) (ii)；卢旺达问题国际刑事法庭，《程序和证据规则》，规则 66 (A) (ii)；塞拉利昂问题特别法庭，《程序和证据规则》，规则 66 (A) (ii)；黎巴嫩问题特别法庭，《程序和证据规则》，规则 110 (A) (ii)；东帝汶重罪特别审判小组，《过渡刑事诉讼规则》，第 24.4 节。

证。[76] 许多刑事管辖区都有披露规则，要求检察官交出任何可能开脱罪责的材料。[77] 任何案件的开源调查人员在开展工作时都应该考虑到这些披露义务，即使该案件最终提交法庭的可能性很小。[78] 调查人员应当考虑披露信息的可能性，还因为其他几个原因。例如，如果要求检察官审查调查工作收集的所有材料，调查人员应谨慎开展批量收集资料工作，因为大量资料可能会造成过重负担，甚至导致无法审查。当涉及到所收集信息的保存和存储时，这一点也很重要，包括进行适当标记，这将极大便利日后的材料检索和审查工作。

D. 程序和证据规则

55. 在开展法律调查工作时，开源调查人员的主要任务是收集相关的真实信息，可用于得出事实和法律结论。特别是在国际性法院和法庭中，调查人员必须致力于确保所收集的任何开源证据均可采纳，并且具有相关性、可靠性和证明力。刑事调查有别于为其他目的进行的调查，刑事调查适用的证据标准更高，[79] 程序和证据规则更严格，包括更严格的可采性，以保护被告人的正当程序和公平审判权利。[80] 虽然在国际性刑事法院和法庭上，证据可采性受到的阻却通常低于一些国家法院，但证据收集方法仍将影响法官对证据的权重衡量。所有司法管辖区域都是如此。在一个数字信息（包括错误信息和虚假信息）激增的时代，[81] 调查人员必须能够确定开源信息是否真实，并以足够的准确性证明或否定其真实性。[82]

56. 就司法程序而言，可采性是指诉讼一方提交的物项是否可采纳为证据记录在案。

[76] 例如见国际刑事法院，《程序和证据规则》，规则 81(4)；前南斯拉夫问题国际法庭，《程序和证据规则》，规则 69；卢旺达问题国际刑事法庭，《程序和证据规则》，规则 69；塞拉利昂问题特别法庭，《程序和证据规则》，规则 69；黎巴嫩问题特别法庭，《程序和证据规则》，规则 115-116；东帝汶重罪特别审判小组，《过渡刑事诉讼规则》，第 24.6 节。

[77] 例如见前南斯拉夫问题国际法庭，《程序和证据规则》，规则 68；卢旺达问题国际刑事法庭，《程序和证据规则》，规则 68；塞拉利昂问题特别法庭，《程序和证据规则》，规则 68；黎巴嫩问题特别法庭，《程序和证据规则》，规则 113；《国际刑事法院罗马规约》，第 67(2) 条；东帝汶重罪特别审判小组，《程序和证据规则》，第 24.4(c) 条。脱罪证据是可能免除被告罪责的证据。在美国，布雷迪规则是美国最高法院制定的一项审前证据开示规则，要求检方在刑事案件中将所有脱罪证据移交给辩护方。见 Brady v. Maryland, 378 U.S. 83 (1963)。

[78] 由于披露义务可能要求将收集到的部分或全部材料移交给辩护方，开源调查人员保护身份和其他敏感信息的能力可能会被否定。

[79] 例如，虽然国际法院通常会适用“排除合理怀疑”的刑法证据标准，但调查委员会和类似机构最常采用的是“合理理由相信”这一较低标准，基于该标准作出调查结论。详见人权高专办，国际人权和人道法调查委员会和实况调查团：指导和实践，第 62-63 页。

[80] International Criminal Court, *Prosecutor v. Jean-Pierre Bemba*, Case No. ICC-01/05-01/08 A, Judgment on the Appeal of Mr Jean-Pierre Bemba Gombo against Trial Chamber III's "Judgment pursuant to Article 74 of the Statute", 8 June 2018, Appeals Chamber, Separate Opinion of Judge Van den Wyngaert and Judge Morrison, para. 5.

[81] 错误信息是指不真实但无意造成伤害的信息。例如：不知道某条信息为不实信息的个人可能会在社交媒体上传播该信息，以期有所帮助。虚假信息是以造成伤害为明确目的而故意制造或传播的不实信息。虚假信息的制造者通常有政治、经济、心理或社会动机。见 Claire Wardle, "Information disorder: the essential glossary" (Cambridge, Massachusetts, Shorenstein Center on Media, Politics and Public Policy, 2018)。可查阅 https://firstdraftnews.org/wp-content/uploads/2018/07/infoDisorder_glossary.pdf?x32994.

[82] 同上。

一般而言，国际性刑事法庭使用三要素测试方法来评估所提供物项的可采性：(a) 相关性；(b) 证明价值；(c) 相较于证据对审判公正性的不利影响，衡量其证明价值。[83] 如果物项有助于增加或减少某个事实的可能性，则其具有相关性；物项证明价值则指该物项是否有助于证明或证伪案件中的争议事实。在非司法调查的情况下，运用类似于可采性的评估。对每条信息都应根据其可靠性、相关性和证明价值进行评估，以确定是否以及应如何用于确定法律和（或）事实结论。[84]

57. 权重指的是赋予某个物项的价值，以及在得出法律或事实结论时最终依赖该物项的程度。权重的确定应采用整体评估方法，部分依赖可能支持、证实或反驳有关事实的其他信息。在许多法律诉讼中，对可采性和权重分别进行评估。在其他情况下，如果证据可采性不是一个因素，人权调查人员将采用类似方法评估信息的权重。

58. 适用于国际刑事诉讼的程序和证据规则载于每个法院的组成文书，最常见的是其程序和证据规则。判例法提供了进一步指导。根据调查的性质，有时应联系法律专家征求咨询意见。如果一项调查旨在促进法院的诉讼程序，则尤应如此。

59. 开源信息可以是书面证据和证词证据的结合体。例如，个人作出陈述的视频需要进行认证，陈述内容也需要单独核实。[85] 因此，确证数字物项作为文件的真实性或评估数字物项作为证词证据所

[83] 根据《罗马规约》（第六十四条第九款第 1 项和第六十九条第四款），国际刑事法院审判分庭“有权应当事一方的请求或自行决定……裁定证据的相关性或可采性……除其他外，考虑到证据的证明价值，以及这种证据对公平审判或公平评估证人证言可能造成的任何不利影响，依照《程序和证据规则》。”

[84] 例如见人权高专办，《国际人权和人道法调查委员会和实况调查团：指导和实践》，特别是关于信息收集与评估的第四 .C 章。

[85] 见 Human Rights Center, University of California, Berkeley, School of Law, “Digital fingerprints: using electronic evidence to advance prosecutions at the International Criminal Court” (Berkeley, 2014)。可查阅 www.law.berkeley.edu/files/HRC/Digital_fingerprints_interior_cover2.pdf. 传闻证据是指作证证人直接知悉范围以外的信息。在一些司法管辖区，传闻证据不予采纳，除非其符合特定的例外要求。在其他司法管辖区，传闻证据可以采纳，但由于无法在控方或辩方的交叉诘问中对其进行适当检验，因此赋予其的权重微不足道。根据欧洲安全与合作组织的说法，“虽然在普通法管辖区，除非有特殊情况，一般不采纳传闻证据，但在大陆法管辖区或国际法庭并不禁止传闻证据”。见 Organization for Security and Cooperation in Europe, Mission to Bosnia and Herzegovina, *Investigation Manual for War Crimes, Crimes Against Humanity and Genocide in Bosnia and Herzegovina* (Sarajevo, 2013), p.26。可查阅 www.osce.org/bih/281491?download=true. 尽管在大陆法管辖区和国际法庭没有这种障碍，但作为一般规则，传闻证据被视为一种特别不可靠的间接证据类别，法官往往赋予其相对很小的权重。

具有的可靠性和可采性的各种手段可以适用。调查人员应了解相关司法管辖区处理各类证据的方法。即使书面证据的作者不明或无法作证，书面证据也可能经常被采信。文件无须通过能够确证文件真实性的证人提交也可能被采信，前提是提供该文件的一方能够明确、具体地显示文件契合案件之处以及如何适用于本案。[86]

60. 在犯罪和违法行为的责任归于指挥结构中更高层官员的情况下，收集的信息不仅可用于确定“犯罪基础事实”(见下文)，而且对于证明被指控个人施害者的责任模式[87]可能也有关。[88]当犯罪或违法行为的每一个要素，包括身体行为(犯罪行为)和被告的精神状态(犯罪意图)，都被证明符合适用的证明标准时，就可以认为个人负有责任。为了做出这一判断，事实调查人员将针对违法或犯罪行为的每个要素审查所纳入的信息。调查人员应熟悉可指控的犯罪或违法行为、每种犯罪或违法行为的要素、被控实施这些行为的主体以及依据的是何种责任理论。在国际刑法案件中，从业人员往往将“犯罪证据”与“关联证据”分开。现将这两个概念解释如下：

(a) 犯罪证据是指控所依据的罪行证据，包括关于何人、何事、何地和何时的信息。[89]例如，如果被控施害者被指控实施了构成危害人类罪的谋杀，则任何证明已发生谋杀的信息均被视为犯罪证据；

(b) 关联证据是指被控施害者对所犯罪行负有责任的证据，如果施害者没有直接犯罪，这一点尤其重要。[90]

86 例如见，前南斯拉夫问题国际刑事法庭，检察官诉 Pavle Strugar 案，案件编号 IT-01-42-T，关于可接受某些文件的裁决，2004 年 5 月 26 日，第二审判分庭，以及检察官诉米兰·米卢蒂诺维奇等人案，案件编号 IT-05-87-T，关于采纳书证的检方动议的裁决，2006 年 10 月 10 日，审判分庭；卢旺达问题国际刑事法庭，检察官诉 Edouard Karemera 等人案，案件编号 ICTR-98-44-T，关于 Joseph Nzirorera 请求采纳律师席文件的动议的裁决：公开声明和纪要，2009 年 4 月 14 日，第三审判分庭；国际刑事法院，检察官诉托马斯·卢班加·迪伊洛案，案件编号：ICC-01/04/-01/06，关于采纳“律师席”所提交材料的裁决，2009 年 6 月 24 日；前南斯拉夫问题国际法庭，检察官诉拉多万·卡拉季奇案，案件编号 IT-95-5/18-PT，关于检方提出的澄清请求和对审判行为准则所作提议的命令，2009 年 10 月 20 日，审判分庭，以及检察官诉拉多万·卡拉季奇案，案件编号 IT-95-5/18-T，关于检方第一次采纳证据动议的裁决，2010 年 4 月 13 日，审判分庭；国际刑事法院，检察官诉热尔曼·加丹加和马蒂厄·恩乔洛·楚伊案，案件编号 ICC-01/04-01/07，关于检察官采纳证据动议的裁决，2010 年 12 月 17 日，第二审判分庭。

87 Cryer, Robinson and Vasiliev, *An Introduction to International Criminal Law and Procedure*, chap. 15.

88 见人权高专办，《谁的责任？在联合国调查委员会、实况调查团和其他调查中追究违反国际人权和人道法的个人责任》(纽约和日内瓦，2018 年)。可查阅 https://ohchr.org/Documents/Publications/AttributingIndividualResponsibility.pdf.

89 Kelly Matheson, *Video as Evidence Field Guide* (WITNESS, 2016), p. 42. 可查阅 https://vae.witness.org/video-as-evidence-field-guide.

90 同上。

换言之，关联证据是将责任方与犯罪行为联系起来的证据。例如，在指控上级未能防止或惩罚其所知晓的被控违法行为的案件中，关联证据就是证明这种知晓状态或证明上级“有效控制”直接行为人的证据。

E. 隐私权和数据保护

61. 隐私权是一项基本人权。[91] 隐私权的一项重要内容是保护个人数据的权利，各种数据保护法对此都有明确规定。[92] 特别是，在利用数字信息和通信技术（信通技术）的调查中，数据保护和隐私法越来越重要。下文简要概述了隐私方面的国际人权的概念，以及开源调查人员应了解的数据保护、数据安全和数据共享的全球框架。在数据环境中，信息隐私（包括已经存在的或可以得到的关于某个人的信息）特别重要。[93]

62. 开源调查人员必须尊重人权并应对隐私权特别敏感，这一问题经常在数字信息背景下提出。例如，侵犯隐私权是法官可用于在国际刑事法院排除证据的少数理由之一。[94] 隐私权支撑并保护人的尊严和其他关键价值包括结社自由和言论自由。欧洲人权法，以迅速扩大的处理数字权利问题的判例法体系对隐私法做出了一些最有力的解释。侵犯这些基本权利将不可避免地导致被告方在刑事诉讼中提出质疑，甚至可能导致对调查方提起民事诉讼的案由。除了隐私法以外，众多数据保护法律和法规也有助于确保个人数据的安全。尤其是，开源调查人员应了解欧洲议会和理事会 2016 年 4 月 27 日在处理个人数据和此类数据自由流动方面保护自然人的第 2016/679 号条例、废除第 95/46/EC 号指令（通用数据保护条例），及其个人数据保护办法，因为这部法律设定了很高的标准，其他国家正在考虑通过类似立法。[95] 但是，各国的数据保护法规各不相同，差异很大，有时甚至会出现直接冲突的规则。开源调查人员应咨询法律专家，以了解与其运营所在地司法管辖区相关的适用数据保护法律和法规。

63. 最后，开源调查人员应了解对未经授权访问数据和网络的一般性禁止。例如，这包括使用在被破坏的数据集中发现的已泄露密码访问受限制材料，以及通过

[91] 隐私权被纳入众多人权文书和 130 多个国家的宪法规定。例如，见《美洲关于人的权利和义务宣言》，第五条；《欧洲人权公约》，第 8 条；《美洲人权公约》，第 11 条；《儿童权利公约》，第 16 条；《保护所有移徙工人及其家庭成员权利国际公约》，第 14 条；《非洲儿童权利与福利宪章》，第 10 条；《阿拉伯人权宪章》，第 16 和 21 条；《东盟人权宣言》，第 21 条。另见 Privacy International，“What is privacy?”，23 October 2017。可查阅 https://privacyinternational.org/explainer/56/what-privacy.

[92] 100 多个国家和众多国际和区域文书都有数据保护法律。例如，见经济合作与发展组织，《关于保护隐私权与个人数据跨国际流动管理公约》；欧洲委员会，《关于在自动处理个人数据方面保护个人的公约》、《欧洲联盟基本权利宪章》；《亚洲太平洋经济合作组织隐私权框架》；《西非经共体内部个人数据保护补充法》。

[93] 大致见 A/HRC/39/29, 第 5 段。

[94] 见《罗马规约》，第六十九条第七款。

[95] 该条例规定，自然人拥有与保护个人数据、保护个人数据处理和个人数据在欧盟内部不受限制流动有关的权利。《关于在自动处理个人数据方面保护个人的公约》、特别是其 2018 年议定书也规定了类似的权利。该公约不仅对欧洲委员会成员国具有约束力，而且对其他一些国家也具有约束力。

欺骗和其他形式的社会工程获得对受限制信息的未经授权的访问。[96]

F. 其他相关法律考虑

64. 在开源调查过程中，其他法律可能相关。以下是开源调查人员应该了解的一些法律考虑事项的非详尽列表。

1. 违反服务条款

65. 一些常见的开源调查技术涉及违反网站或平台的服务条款。例如，刮取数据或使用虚拟身份（而非真实身份）违反了平台服务条款尤其是社交媒体平台服务条款。[97] 违反服务条款是违约行为。调查人员应核实此类行为在其工作所在地司法管辖区内是否也可能构成非法行为。必须在有必要维护通过使用虚拟身份而提供的安全原则与在此等情况下违约可能造成的损害之间取得平衡，最常见的补救办法是禁止用户访问平台。然而，如上所述，虽然虚拟身份在用于纯开源搜索和监测时是必需的，但不应将虚拟身份用于试图访问社交媒体上分享的、必须遵守限制性访问控制要求的内容；或以虚拟身份为借口，打着虚假身份的幌子直接从某人处获取信息。此等行为会导致调查人员脱离开源调查的范畴，会违反道德原则[98]并可能触犯法律。[99]

2. 知识产权法

66. 调查人员应了解他们可能需要获得的任何知识产权许可才能合法发布、分发和（或）以其他方式使用他们在调查过程中收集的信息。相关法律因司法管辖区而异，但大多数司法管辖区（至少）为内容（例如网上共享的视频、照片或文本）的创作者提供某种形式的版权保护。“创作者”通常被定义为实际创造素材的人。–例如通过拍摄照片、录制视频或撰写原始文本创造素材的人，而不是上传素材的人，尽管他们可以是同一个人。最终用户可能需要获得创作者的同意才能进行拟议的使用，以避免侵犯版权（例如，如果在公开报告或新闻报道中使用相关内容）–如果上传者不是创作者，那么获得上传者的同意通常不足以避免违法。这也是试图找到调查人员可能获取的每一项内容的原始来源的另一个原因。一些（但不是所有）司法管辖区规定了需要获得同意的例外情况，通常称为“合理使用”

[96] 根据美国国家标准和技术研究所的说法，社会工程是“通过与个人交往以获得信心和信任，进而欺骗个人透露敏感信息的行为”(Paul A. Grassi, Michael E. Garcia and James L. Fenton, *Digital Identity Guidelines* (Gaithersburg, Maryland, National Institute of Standards and Technology, 2017), p.54。另见 Michael Workman, “Gaining access with social engineering: an empirical study of the threat”, *Information Systems Security*, vol. 16, No. 6 (2007)。关于未经授权和欺骗性访问的进一步讨论，见下文第 65 段。关于用户伪装的讨论，见下文第 107 段。

[97] 例如：脸书的服务条款要求用户“使用与您在日常生活中所用名字相同的名字”、“提供关于您自己的准确信息”、“只创建一个账户（您自己的账户）并将您的时间安排用于个人目的”。见 www.facebook.com/terms.php. 冒充行为违反推特的规则和政策。见“Impersonation policy”，网址：https://help.twitter.com/en/rules-and-policies/twitter-impersonation-policy.

[98] 关于虚假陈述的讨论，见上文关于道德原则的第二章 C。

[99] 见上文关于隐私权和数据保护的第三章 E。

或“合理处理”例外情况，即将录像、照片、文字和其他信息用于某些有益于社会的目的（如教育、执法或新闻）的情况。然而，这些例外情况在适用时通常非常狭窄，因此，在没有仔细审查的情况下，决不能假定某一特定用途属于此类例外。有些机制有时有助于将侵权的可能性和（或）范围降到最低，如，在数字报告中嵌入原始内容的链接而不将其从原始来源中移除；认可创作者的功劳；只使用原始内容的一小部分——不过这也需要视具体情况和司法管辖区而定。受知识共享许可或其他免费许可约束的信息可能有广泛的免费许可用途。不过，如果适用此类免费许可，就必须遵守许可条件且不要将内容视为免于许可。

四

安　全

本章摘要

- 每个人都有责任确保调查和受调查影响者的安全，而不仅仅是信息技术专业人员的安全。
- 安全考虑应包括双重的：(a) 与基础设施有关，包括硬件、软件和网络；(b) 与行为有关，包括调查人员以及与之互动的所有人员的行为。
- 安全评估应在三个层面进行包括在组织、具体调查 / 案件和具体活动 / 任务层面进行。
- 保护措施应旨在减轻风险评估调查所确认的风险和威胁。
- 安全评估应考虑所有类型的伤害，包括数字、财务、法律、物理、社会心理和声誉伤害。
- 开源调查中某些最大的漏洞与互联网连接 /IP 地址、设备及其功能和用户行为有关。
- 调查人员和调查机构应参与持续的安全培训，并部署随着威胁或薄弱环节性质的变化而发展的保护措施。

67. 本章概述了与开源调查相关的在线和离线安全考虑因素。通过适当的准备、投资和对威胁评估和风险缓解的关注，开源调查人员应该能将人员、数据和其他资产可能遭受的危害风险降至最低。包括硬件和软件在内的安全基础设施以及用户行为规程应尽可能在开始调查之前就位、定期予以评估并在必要时进行更新。一个组织的规模和资源可能会对某些保护措施的可行性产生影响；因此，本章载有灵活的标准，应根据组织和调查的具体需要对标准进行调整。开展高风险调查（例如涉及特别脆弱受害者的调查或在被控施害者是国家行为体和（或）被单独指认的情况下开展的调查）的组织应该聘请有经验的网络安全专业人员提供服务。此外，一个健全的安全框架应包括某种独立的审计机制和持续的培训，以便用户能够紧跟新技术发展和最佳实践的步伐。

A. 最低标准

68. 由于安全基础设施和用户行为的最佳实践在不断变化，该规程提供了总体原则，以帮助指导开源调查人员思考安全问题。调查人员必须对自己的安全负责，包括评估其行为造成的风险程度并采取适当的减少风险和保护措施。尽管需要对安全采取定制的个性化方法，但为了遵守安全原则，开源调查人员应始终在工作中应用一些最低标准：

(a) 开源调查人员应避免向第三方披露有关其自身、其所在组织和任何合作伙伴或来源的可识别要素，除非这是调查目标或义务。因此，调查人员应在网上保持匿名，并尽可能确保其网上活动无法溯源；

(b) 开源调查人员开展网上活动时应预见到此类活动可能受到第三方的监测和分析。因此，他们在进行网上活动时，应与其虚拟身份保持一致，且不得泄露其身份或调查目标，也不得危及其信息来源人或其他第三方；

(c) 开源调查人员应当意识到，过度利用单一网上信息源（如特定网站）可能会增加第三方监测和分析的风险。因此，他们应采取措施尽量降低这种可能性，例如数字来源多样化等做法；

(d) 开源调查人员应避免可识别或可预测的行为模式（如在可识别设备上的重复搜索模式），这些模式可能有助于第三方确定调查目标，并使调查人员更容易成为网络钓鱼攻击和其他形式社会工程的目标；[100]

(e) 开源调查人员应始终将其专业工作与个人网上活动分开。不应使用个人网上账户，并尽可能不使用个人设备开展专业调查，且决不能将专业设备用于个人网上活动；[101]

[100] 见下文有关网络钓鱼攻击和社会工程的说明。

[101] 如果使用个人设备不可避免，则用户应在分隔开的网上环境中进行专业调查和个人活动，例如通过使用虚拟机进行调查。

(f) 开展多项调查的开源调查人员不应将各项调查混在一起。因此，他们应为每项调查活动保存不同的起始时间，在不同地点维护各项调查的数据和文件，并在必要时使用不同的虚拟身份；[102]

(g) 开源调查人员应使用旨在将可能引入的敌对或恶意软件或活动期间可能遇到的其他破坏性影响降到最低的技术系统或环境的设计。

B. 安全评估

69. 为了开发一个适当有效的安全框架，开源调查人员必须了解网络安全和风险管理的关键概念。他们还必须能够识别需要保护的资产和潜在的危害，并评估潜在的威胁、风险和漏洞。

70. 风险是指薄弱环节被威胁因素所利用而导致资产丧失、损害或损毁的可能性。这些术语的定义见下文。在互联网上开展的开源调查涉及与传统调查不同的信息收集方法，因此产生了不同类型的风险。这些风险的识别和评估是调查规划和准备工作的重要组成部分。开源调查中的一些常见风险例子包括：调查对象或支持调查对象的实体的技术能力和技术意识，这些对象和实体可能逃避或误导调查；用于调查的网上环境技术配置中存在的问题，这些问题可能导致信息泄露，从而损害调查；可能损害调查人员计算机系统、活动、身份或收集数据的恶意软件或代码；或可能会损害调查活动的技术功能，例如跟踪装置、cookies、信标和分析。

71. 下一节解释关键术语及其在开源调查中的应用，从而为评估威胁和风险提供路线图。

1. 资产

72. 资产系指需要受到保护的任何事物，包括人员、[103] 财产和信息。在进行开源调查时，需要保护的人员可能包括调查人员或调查小组，包括调查人员或调查小组与之一起工作的任何人（即内部同事和外部伙伴，当地和实地工作人员）、信息作者或信息来源、证人、受害者、被控施害者和旁观者。财产包括可以被赋予价值的有形和无形物项。[104] 有形资产包括建筑物、设备和文件，而无形资产则包括声誉和专有信息，如数字数据、元数据、数据库、软件代码和记录。

2. 伤害

73. 伤害系指对资产的物理或精神损害或损伤，或资产的损毁，可能涉及数字、财务、法律、物理、社会心理或声誉方面的伤害。

[102] 除了尽量减少混淆不同调查的风险外，这种做法还有助于有效维护监管链。

[103] 只有在进行安全评估时才将人员称为资产。

[104] 见Threat Analysis Group, “Threat, vulnerability, risk - commonly mixed up terms”。可查阅www.threatanalysis.com/2010/05/03/threat-vulnerability-risk-commonly-mixed-up-terms.

(a) 数字伤害

74. 数字伤害系指对任何数字信息或基础设施的损害。潜在的数字伤害可能包括破坏、操纵或丧失数据访问权限，或中断计算机系统和平台的服务。

(b) 财务伤害

75. 财务伤害可能源自多个方面，包括与调查相关的法律和声誉伤害。调查人员、调查目标和无关人员都可能遭受这种伤害。此外，如果调查人员未能充分评估调查的长期成本，也可能会造成财务伤害。

(c) 法律伤害

76. 开源调查人员可能会因其工作过程或产出而承担法律责任。调查人员应当了解法律所允许的界限和自身行为的法律后果，以尽量减少自己和（或）第三方承担法律责任的风险。调查也可能会对调查对象乃至无关人员造成法律伤害，他们可能会卷入调查过程中发现的法律错误。[105]

(d) 物理伤害

77. 物理伤害可能包括对人身或财产的损害。虽然开源调查人员通常不用深入实地，而是在办公室或家里工作，但仍应将物理伤害作为线上活动的一个潜在后果予以评估。网络空间的行为可能导致现实后果，调查人员应当对此有意识，并做好准备。例如，开源调查人员应当意识到可能身处不安全的环境，并可能会因调查人员线上行为而面临物理伤害风险的个人，不管是，同事、涉事国家的线上用户或者是其他人。线上调查人员在伦理上对他人负有注意义务，有时在法律上也负有注意义务，[106] 应当确保那些面临物理伤害风险的人员不会因调查活动而陷入更大的危险。在开始工作之前，应当将物理风险作为全面威胁评估的一部分加以考虑，并在调查的整个过程当中重新予以评估。

(e) 社会心理伤害

78. 社会心理伤害的范围从心理困扰到心理创伤不等，可能影响到调查团队的任何成员和（或）以其他方式参与调查的人员或受调查影响的人员，包括调查对象和无关人员。除了从道德或伦理角度出发必须保护自己和他人免受心理伤害之外，还应意识到，人有时可能是任何组织有效运作过程中最脆弱的环节。正在经历社会心理伤害的人员可能尤为脆弱，给威胁行为体创造新的可乘之机，或给物理和数字安全带来其他风险，特别是在负面的心理影响导致功能受损的情况下，例如，安全规程未能像往常一样得到严格遵守。众所周知，观看大量暴力和其他视觉冲击较大的视频画面尤其令人难以承受，会造成心理困扰或创伤，可能需要专业支持。继发性创伤的迹象可能包括行为改变、情绪波动、饮食习

[105] 关于相关法律考虑因素的更多讨论，另见前文第四章 E 和 F 节。
[106] 《罗马规约》，第五十四条第（一）款第 2 项。

惯变化、失眠、嗜睡或做噩梦。[107] 关于制定和创建韧性计划和自我护理的一节描述了减轻社会心理伤害的策略。[108]

(f) 声誉伤害

79. 在开源调查中，开源调查人员和（或）其组织可能会遭受最为严重的声誉伤害，例如，如果调查人员发布错误信息、违反伦理或以其他方式制作有问题的内容，其声誉就会受到伤害。调查对象也可能会因其被指控的行为公之于众而面临污名化，从而遭受声誉伤害。当针对个人或组织的指控事后证明错误时，这种声誉伤害可能尤为令人担忧。

3. 保护措施

80. 保护措施是为防止或最大限度减少漏洞而采取的努力，可能包括物理、技术和政策措施。物理保护可能包括为储存有敏感资料的建筑、房间或柜子上锁。技术措施可能包括在设备上设置密码、使用加密和多重要素验证，或控制数据系统的访问权限。政策措施包括内部和外部的规则、法律和执行机制，例如，禁止用工作电邮向个人电邮发送内部工作成果的规则，或禁止在办公电脑上使用个人社交媒体账户的政策。

4. 威胁

81. 需要对资产加以保护，使其免遭威胁。威胁系指任何可能有意或无意地利用漏洞获取、损害或损毁资产的东西。威胁可能来自组织或调查的内部或外部，可由个人、团体、机构或网络执行。开源调查人员应当重点注意下列威胁。

(a) 分布式拒绝服务攻击

82. 分布式拒绝服务攻击是旨在破坏目标访问机器或网络的能力的网络攻击。针对面向公众的资产，如网站和其他远程访问门户，应当建立缓解此类攻击的系统。此外，应当建立事件日志系统，在发生攻击时，用以记录所有行动和相关行为体。

(b) 网络钓鱼攻击

83. 网络钓鱼系指在电子通信中伪装成值得信赖的实体，企图骗取敏感信息，如用户名、密码和详细的信用卡信息。[109] 网

[107] 见 Dart Center for Journalism and Trauma, "Working with traumatic imagery", 12 August 2014 (可查阅 https://dartcenter.org/content/working-with-traumatic-imagery); Sam Dubberley, Elizabeth Griffin and Haluk Mert Bal, *Making Secondary Trauma a Primary Issue: A Study of Eyewitness Media and Vicarious Trauma on the Digital Frontline* (Eyewitness Media Hub, 2015) (可查阅 http://eyewitnessmediahub.com/research/vicarious-trauma); Sam Dubberley and Michele Grant,"Journalism and vicarious trauma: a guide for journalists, editors and news organisations"(First Draft News, 2017) (可查阅 https://firstdraftnews.org/wp-content/uploads/2017/04/vicarioustrauma.pdf); Center for Human Rights and Global Justice, "Human rights resilience project launches new website", 21 May 2018(可查阅 https://chrgj.org/2018/05/21/human-rights-resilience-project-launches-resources-for-resilience-website); Keramet Reiter and Alexa Koenig,"Reiter and Koenig on challenges and strategies for researching trauma", Palgrave MacMillan (可查阅 www.palgrave.com/gp/blogs/social-sciences/reiter-and-koenig-on-researching-trauma)。

[108] 关于自我护理的更多信息，见下文第五章 D 节。

[109] 见 Phishing.org, "What is phishing?"。可查阅 www.phishing.org/what-is-phishing.

络钓鱼或电话诈骗被用来获取保密信息或骚扰调查人员。个人账户遭遇网络钓鱼的风险通常高于专业账户；因此，使用个人账户可能会危及调查或工作成果。

(c) 中间人攻击

84. 中间人攻击系指恶意行为体将自己插入两方之间的对话，冒充对话双方，并获取双方试图发送给对方的信息的一种网络攻击。[110] 在中间人攻击中，恶意行为体能够在外部的任何一方不知情的情况下，拦截、发送和接收原本意在传送给他人的数据或根本无意发送的数据，而外部的任何一方知道时已经为时太晚。[111]

(d) 社交工程

85. 社交工程是指对人们进行心理操纵，左右其采取可能有害的行动，如泄露保密信息。社交工程有多种不同的形式，如鱼叉式网络钓鱼。[112] 由于社交工程的伎俩在不断调整和演变，调查人员应当不断接受培训，了解如何发现和避免已经发现的社交工程伎俩。

(e) 恶意软件

86. 恶意软件是具有险恶意图的软件的简称，是指设计目的是不经用户同意而侵入并损坏计算机的计算机程序。恶意软件有若干类型，包括间谍软件和勒索软件。

5. 威胁行为体

87. 威胁行为体或恶意行为体系指对某一事件或事故负有责任的个人或实体，且该事件或事故对另一实体或行为体的安全或保障产生影响，或有可能产生影响。在国际刑事和人权调查中，威胁行为体可能是被控施害者、调查目标（包括政府）或其支持者。开源调查人员必须确定潜在的威胁行为体，了解其能力和发动攻击的可能性。

6. 漏洞

88. 漏洞指保护措施的薄弱之处或短板，既可能存在于数字领域，也可能存在于物理领域。在线上活动方面，漏洞可能包括安全保护措施的薄弱环节，可被加以利用，在未经授权的情况下访问资产，漏洞也可能包括软件的安全缺陷、不安全的设计以及权限过多的用户和代码。在线下环境中，漏洞也可能包括人的弱点，如容易受到勒索或胁迫的团队成员，或因过度暴露于视觉冲击较大的材料或其他艰难的工作条件而变得脆弱的团队成员。[113] 向调查目标公开正在进行的调查，或者透露调查范围，可能会产生新的漏洞。最后，安全漏洞可能来自外部威胁，如新的恶意软件和病毒，调查人员应当对此有所了解。安全摸排和风险评估工作应当考虑到这些类型的漏洞。

[110] 见 Veracode，“Man in the middle (MITM) attack”。可查阅 www.veracode.com/security/man-middle-attack.
[111] 同上。
[112] 鱼叉式网络钓鱼是指发送看似来自熟知或可信发件人的电子邮件，以诱使目标个人透露保密信息的欺诈做法。
[113] 关于韧性和自我护理的更多信息，见下文第五章 D 节。

89. 开源调查人员还应当注意下列线上漏洞。

(a) Cookies

90. Cookie 是一个小型文件，通常通过网站发送，或存储于用户计算机内存之中，或写入计算机磁盘，以供浏览器使用。Cookie 通常是网站正常运行所必需的，例如，通过存储网站偏好和详细身份信息，以便用户无需在后续访问中重复输入数据。开发 cookie 的初衷，就是为了收集和存储有关访客及其访问活动的重要且通常较为敏感的数据。有些 cookies 已经发展成为集中化工具，可用来收集数据，帮助刻画用户的浏览兴趣和习惯。Cookie 在过期或被用户删除之前，可能会一直存在于计算机上。

(b) 跟踪器

91. 跟踪器是 cookie 的一种类型，它利用浏览器的功能记录访问过的网页、以往输入的搜索标准等。跟踪器是持久存在的 cookie，无间断保留网站访客的行为日志。跟踪器的形式可以十分简单，通过给用户的浏览器分配一个独特的身份，然后将该身份与所有后续浏览和搜索活动（搜索标准、访问过的网页、访问过的网页顺序等）关联在一起。如此一来，跟踪器的主人就能够将网站（或一组附属网站）的先后访问联系在一起，从而详细掌握用户及其浏览习惯的全貌。跟踪器通常内置在广告之中，而广告则分布在多个网站上，让跟踪器有更多机会捕捉用户的活动和行为。即使用户访问的是一个“可信的”网站，也可能导致用户电脑被安装跟踪器，使得用户在互联网上的后续活动被跟踪。

(c) 信标

92. 信标是一种跟踪用户活动和行为的机制。信标由网页中的一个不显眼（通常不可见）的小元素构成（与一个透明的像素一般大小），经由浏览器渲染后，会导致该浏览器和附属计算机的详细信息被发回第三方。信标可与 cookie 一起使用，以触发数据收集和传输，用独特方式识别用户身份和记录用户的浏览习惯。信标与社交媒体网站密切相关，对这种网站而言，识别个中关系和网络是其关键基石。最后，信标可以在基于 HTML 的电子邮件中使用，以收集和报告用户身份，访问以往存储在该计算机上的任何 cookie。

(d) 其他代码和脚本

93. 越来越多的网站正在利用由访客的浏览器下载的小段代码，这些代码能够存储相关访问的信息。这种代码可以影响网站的显示方式、网站对输入的反应和浏览器对网站的反应。代码还可存储访客凭证、活动等方面的敏感数据。数据收集可能是持久性的，收集后还可能将数据发送给第三方。

C. 基础设施方面的考虑因素

94. 基础设施系指开源调查所需的结构、设施和系统，包括软件和硬件。基础设施应当提供（并具备）充足的安全措施，以保护和保存组织的资产和数据。为了提高基础设施的韧性，缓解措施应当到位，以确保在发生下列任何情况时连续运行：

(a) 互联网连接中断或丧失；

(b) 对已储存数据的访问权限中断或丧失；

(c) 数据丢失、受损或破坏；

(d) 软件服务中断或丧失；

(e) 硬件损坏或丢失；

(f) 未经授权访问设备；

(g) 未经授权访问网络；

(h) 意外删除或篡改数据；

(i) 故意破坏或篡改数据；

(j) 数据泄漏或数据遭到“挟持”。

95. 调查所必需的架构由将要进行的线上调查活动规模、调查性质和相关对象以及可用于根据需要建立、维持和修改基础设施的资金决定。

1. 基础设施

96. 用于开源调查的基础设施将至少包括下列部分，以及具体调查策略所涉及的其他功能。

(a) 设备

97. 开源调查人员必须拥有访问线上内容的设备，如台式电脑、笔记本电脑、平板电脑或智能手机。硬件和设备应当设有密码保护，启用全盘加密，并且最好使用多重要素验证。[114] 所有设备都应定期备份。未在使用的硬件应当安全存放，仅限用户和经批准的人员访问。个人设备不应用于工作相关活动。同样，与调查有关的设备也不应用于个人活动，否则，就会有将个人社交媒体与为调查目的打造的虚拟身份关联起来的风险。[115]

(b) 网络连接

98. 在理想情况下，调查人员会有信号较强、稳定、私密的互联网连接，应当避免使用公用 Wi-Fi。虽然免费的公用 Wi-Fi，包括如酒店或网吧提供的网络等半私密网络，提供了便捷的选择，但这类网络很不安全，容易受到多种威胁，其中最大的威胁就是黑客能够将自己安插在用户与连接点之间。使用有密码保护的个人热点确实需要投入资金，但这对于进行安全的线上调查活动至关重要。此外，虽然互联网连接并不总是在调查人员的掌控之下，但从功能和安全角度出发，最好还是选择信号较强、稳定的互联网连接。如果要在网络连接不稳定的情况下使用虚拟专用网络 (VPN)，调查人员应建立故障保险机制，确保在连接中断时不会暴露 IP 地址。

(c) 网页浏览器

99. 网页浏览器是线上调查使用的核心工具之一，用于查询、搜索和访问互联网上

[114] 多重要素验证是一种安全增强手段，要求用户在登录一个账户时出具两种凭证，例如，既要提供一个密码，还要提供一个生物特征识别 (指纹) 或智能卡。见 United States, National Institute of Standards and Technology, “Back to basics: multi-factor authentication (MFA)”。可查阅 www.nist.gov/itl/applied-cybersecurity/tig/back-basics-multi-factor-authentication.

[115] 这一建议在差旅期间可能难以遵守，因为许多调查人员会随身携带办公设备，但又想或需要在下班之后处理个人事务。因此，进行开源调查的组织应当制定合理的差旅政策。

发布的网站。浏览器是调查人员与互联网之间的主要界面，但作为一个风险来源，却经常遭到忽视。现代浏览器在不断迭代，具备广泛的内置功能，以适应多种要求。对于意图实施监视或向对手发动攻击的人而言，浏览器也是一个关键目标，因为浏览器的功能可以被滥用，还可以相对容易地添加功能。浏览器能够同时访问互联网和计算机，因而，也能够访问可用于识别用户身份的信息。通过浏览器泄漏的数据，可能会披露足以警示调查对象的数据。现代浏览器有若干内置功能，还可添加多个额外功能，即所谓的浏览器插件，插件在单独或合并运行时有可能泄露数据，导致调查工作、调查人员或查询线索及相关搜索活动暴露。浏览器在默认情况下还能下载和执行网站派生的计算机代码。对调查人员而言，计算机代码的存在和（或）功能可能并不明显，但代码可能会改变传送给调查人员的数字内容，访问调查人员计算机上的功能和数据，甚至导致计算机偏离设想异常运行。开源调查人员应当设法将这些风险降到最低，确保使用经过定期筛查的、安全、更新版浏览器，并安装能够减轻上述部分风险的适当软件和插件。[116]

2. 安全措施

100. 基础设施的上述基本要素可用来识别用户的身份和位置。为了遵守匿名和不溯源原则，调查人员应当采用下列策略伪装其互联网连接。这种策略能对位置和IP地址进行掩码，对机器进行伪装，遮蔽其识别特征、操作系统和浏览器。

(a) 伪装连接

101. IP地址可能会泄露信息，而这些信息可能会被用来攻击组织的基础设施。开源调查人员应当设法使用VPN、代理或其他软件屏蔽其计算机的IP地址，也就是说，要确保向互联网披露的IP地址与调查人员或其组织没有关联。VPN还能为调查人员的计算机与VPN服务器之间的通信建立加密通道，如此一来，连接信号穿越的任何网络/节点都只能看到经过加密的数据，从而提供一层额外的保护。不过，某些VPN被一些国家和网站禁用，使用这种VPN可能会使调查活动引起第三方的怀疑。在理想情况下，VPN应能够让调查人员使用多个IP地址，并在必要时能够迅速切换IP地址。IP地址应不可追踪至一个国家，而应分割开来，反映世界各地的多个地点。

(b) 伪装机器

102. 为了遮罩可用于识别用户身份的某些特征，调查人员可使用虚拟机，即具备独立计算机的行为能力的软件程序或操作系统。使用虚拟机实际上是在计算机中创建一个新的计算机——一个与计算机其他部分完全独立的环境。虚拟机也能够执行任务，如运行应用程序和计算机

[116] 关于浏览器和其他操作安全措施方面的最新指导，见Computer Security Resource Center of the United States National Institute of Standards and Technology (https://csrc.nist.gov)。

程序。虚拟机如同一台完全独立的计算机，[117] 能让使用它的调查人员看起来像一个不同的线上主体。在使用虚拟机时，调查人员可通过一个系统改变浏览器、用户代理、软件、打开的端口、操作系统和机器的其他相关信息，这样一来，用户每次上线时都显示为一个不同的主体。在理想情况下，基础设施会允许调查人员使用虚拟机，以遮罩实际使用的机器。虚拟机可以销毁和重新创建、恢复至以往节点、改变配置方式、复制用于新案件或保存以备今后之需。调查人员也可采取比较繁琐但也相对有效的替代方法，每次上线时使用不同的浏览器，改变设置以限制所用机器的指纹的独特性，使用防止跟踪的插件，从而手动改变线上表象。

3. 其他基础设施

103. 调查人员应在开始工作前考虑用其他基础设施保护其网络和基础设施，包括以下系统：

(a) 备份系统；

(b) 审核活动及跟踪用户操作的记录系统；

(c) 隔离的存储系统及合适的存储位置，用于收集搜索过程中发现的数字材料。为了保护数据不受外部影响，各组织应拥有与主要网络分离的平台（如证据存储库、数据库或其他信息管理系统）。平台应有两大部分：一个部分连接到互联网，另一个部分断开连接。在某些情况下，妥当的做法可能是，尽快将数据从连接互联网的基础设施转移到更安全的网络／存储库，以便可以安全地审阅信息。

D. 与用户相关的考虑因素

104. 任何安全框架最薄弱的环节之一都是用户。即使有完善的基础设施，但如果不通过定期培训和监督来调整用户行为，安全原则也不会得到遵守。安全是每个人的责任。如果没有接受过有关如何降低这些风险的适当培训，个人不应参与可能使数据或人员处于风险之中的活动。应对调查人员进行培训，使之能评估在进行不同在线活动时，哪些行为是适当的。

105. 若威胁行为人试图追溯活动的源头直至网络或用户，在此种情形下，匿名有助于最大限度地降低危害。[118] 任何在线活动都容易成为第三方跟踪的目标；因此，调查人员在开展网上活动时应假定存在这样的威胁。最常见的跟踪对象包括 IP 地址、浏览器和屏幕分辨率（用于识别设备）、网站导航时间和活动（例如输入的搜索词或访问的页面）。威胁行为人可能会试图识别在线活动的源头。如果有人试图追溯，应引导威胁行为人远离调查人员或调查实体的真实位置或身份。这可以通过一些措施来实现。例如，使

[117] 见 Techopedia，“Virtual machine (VM)”，2020 年 5 月 21 日。可查阅 www.techopedia.com/definition/4805/virtual-machine-vm.

[118] 追溯是指通过追踪信息或一系列事件的线索来发现某人或某事的起源点。

用 VPN，这样看起来就像是从其他地点上网；或是创建和使用虚拟身份，就像是另一个人在上网。[119]

106. 屏蔽在线调查中使用的连接和机器能提供重要保护，但如果用户通过在网站上进行自我身份识别、使用个人信息注册或登录社交媒体平台或其他私人账户等方式暴露自己，这种保护可能就会受到破坏。调查人员不得使用其个人账户进行调查，不得在用于开源调查的浏览器中登录个人账户。某些账户可能要求在创建时使用照片、电话号码或电子邮件。绝不可使用调查人员或其他人的个人照片、电话、电子邮件或数据，也不得使用可归因于调查人员或其他人的照片、电话、电子邮件或数据。

用户伪装

107. 虚拟身份[120]是一种虚假的在线身份或个人资料，可用于在社交媒体平台和其他需要用户登录才能查阅内容的开放式网络平台上进行安全调查活动。这还可能包括使用虚假在线身份而非真实身份的虚拟账户、电子邮件、消息服务、数据库、产品或应用程序。从安全角度而言，开源调查人员应当为开源材料的在线调查活动创建和使用虚拟身份。这是为了确保在威胁行为人试图追溯该个人资料的在线活动时，他们将找到基于虚拟身份的一致和令人信服的信息，而不会泄露调查人员或调查实体的真实信息，或关于调查内容或重点的信息。这也是保护为调查提供支持的人员的一项重要安全措施。应对虚拟个人资料和账户以及使用这些资料和账户进行的活动进行规划，[121]应保留用于创建账户的信息记录，并记录使用此类账户的活动，以便在日后需要时（例如在法庭上）进行解释。[122]

[119] 关于虚拟身份的讨论，另见上文第二章 C 节、第三章 F 节、第四章 A 和 C 节。
[120] 任何对虚拟身份的使用都应平衡安全需要与关于透明度的道德原则。见上文第二章 C 节（道德原则）。
[121] 见下文第五章 C 节（在线调查计划）。
[122] 见下文第六章 D 节（保存）。

五

准　备

本章摘要

- 准备和战略规划是进行彻底、安全调查的关键。
- 准备包括三个过程：(a) 评估威胁和风险，制定减轻这些威胁和风险的计划；(b) 评估信息格局；(c) 制订调查计划。这些过程可能在调查的整个生命周期内相互重叠和 (或) 重复。
- 准备工作包含制定计划以处理调查的任何负面社会心理问题，例如，可能因接触图示资料或其他可能造成创伤的资料而产生的负面社会心理问题。
- 准备工作包含制定计划说明如何处理在调查的整个生命周期中收集的任何信息，包括何时以及在何种条件下删除信息、如何以及在何种条件下共享信息、谁应有权查阅信息。
- 准备工作应包含对可能有用的软件和其他工具进行评估。调查人员应了解商业资源、定制资源和开源资源之间的权衡。

108. 开源调查人员应在采取某些准备措施之后才开始在线调查活动。准备步骤应包括进行数字威胁和风险评估以及数据格局评估。[123] 调查人员随后应制定在线调查计划整合这些评估的见解。下文详细介绍了每项活动。

109. 在组织层面，收集和保存信息之前还必须制定关于数据保留、数据删除、数据查阅和数据共享的政策，详情如下。

A. 数字威胁和风险评估

110. 考虑潜在的威胁并采取策略管理风险（无论是物理、数字还是心理社会风险），将确保安全和道德原则得到遵守。首先，应进行数字威胁和风险评估，确定可能因在线活动（尤其是访问目标网站、对特定来源进行持续监控或从社交媒体平台抓取数据）而产生的一般性威胁和具体案件相关威胁。评估应包括传统威胁分析的要素，如查明所有潜在的威胁行为人，评估这些威胁行为人的兴趣点和能力，确定攻击的可能性，考虑薄弱环节，并采取保护措施，尽量减少漏洞。与安全专家、特别是那些具有网络安全专门知识的专家协商，或征询这些专家的意见，将有助于评估。[124] 应定期审查评估结果并在必要时进行更新。此外，可能还需要针对特定类型的在线活动或新的潜在威胁行为人的介入进一步作出评估。[125]

B. 数据格局评估

111. 开源调查人员应了解所调查情形的数据环境。可用技术和已使用技术的类型，包括由谁使用，将对可获得的数字数据类型产生影响。这需要确定在被调查的地理区域内最常用的在线平台、通信服务、社交媒体平台、移动技术和移动应用程序。例如，在战争罪调查中，调查人员需要了解武装冲突各方以及旁观者或其他证人使用的交通工具、信通技术和数字媒体的类型，以便了解哪些类型的信息最有可能被获取并在网上传播。

112. 调查人员应调查在该地理区域内使用或接触到这些技术的人群类别。在这方面，调查人员应意识到，用户生成的公开数字内容，包括社交媒体帖子和通过网络平台分享的信息，可能无法平等捕捉到针对所有个人和团体的侵害行为的全部范围。这是因为数字技术的使用可能因性别、[126] 民族、宗教、信仰、年龄、社

[123] 见下文附件二（数字威胁和风险评估模板）和附件三（数据格局评估模板）。

[124] 关于开源调查中威胁和风险的一般信息，见上文第四章（安全）。

[125] 见下文附件二（数字威胁和风险评估模板）。

[126] 例如，妇女、女童和男女同性恋、双性恋、跨性别者、间性者可能无法使用或持有家里的移动电话。关于所谓“性别数字鸿沟”的进一步讨论，见 A/HRC/35/9。另见人权理事会第 32/13 号决议，以及 Araba Sey 和 Nancy Hafkin（编辑），*Taking Stock: Data and Evidence on Gender Equality in Digital Access, Skills, and Leadership*（中国澳门，EQUALS 全球伙伴关系和联合国大学，2019 年）。可查阅 www.itu.int/en/action/gender-equality/Documents/EQUALS%20Research%20Report%202019.pdf.

会经济地位、种族、语言、[127] 族裔或宗教少数群体成员身份、土著身份、移民身份和地理位置等因素而有所不同。[128] 造成这种不平衡的原因可能是缺乏设备、设施或资源获取途径，[129] 这些个人没有机会在网上创建或上传信息，披露他们的问题或相关侵害行为。另一个因素可能是，除其他外，这些人可能没有获得平等的教育，因此在技术技能方面能力较弱。由于交叉形式的歧视，某些社会阶层可能在网上双重隐形。例如，关于上述边缘群体之一的妇女和女童的信息在开源资料中可能更少出现。这些因素可能意味着，这些人既不是内容创建者，也不是内容所述对象，因而导致在线调查的结果出现偏斜。

113. 此外，社会各阶层获得技术的机会不平等，也可能导致不仅对在线内容所代表的对象，而且对在线可获取的侵害行为类型的关注发生偏斜，涉及用户生成的内容时尤其如此。例如，当妇女共用其男性家庭成员拥有的移动电话，或与他人共用一个账户时，她们可能不会讨论性暴力和性别暴力等敏感问题，或是性健康和生殖健康方面的问题。此外，社交媒体上用户生成的内容，包括照片和视频，可能更容易反映某些侵害行为而非其他侵害行为。例如，性暴力和性别暴力可能发生在私人场合，与驱逐相比，更难通过照片反映出来。

114. 虽然上述部分因素可以通过查阅多种类型在线资料 (而不仅仅是用户生成的内容) 来减轻，但在分析其他类型的开源信息时还是必须考虑同样这些因素。例如，在查阅政府生成的数据和统计数据时，调查人员应始终质疑数据是否准确地反映了社会所有阶层、所有方面的情况。[130] 有若干关键事项和技术可以评估，取决于根据一项具体调查的地域和时间范围确定的相关性。调查人员应考虑性别、年龄、地理位置、社会经济差异和其他相关人口统计学信息。这项评估的目的是提高调查人员对所调查情况的了解，以便设计有效的在线调查战略，并迫使调查人员考虑可用的在线数据中可能存在的偏差。不是所有调查都涉及这些类别，因此调查人员应调整数据格局评估，使之与具体情况相宜。[131] 关于可列入数据格局评估的各类信息的完整清单，见下文附件三。

[127] 例如，那些属于语言少数群体的人，在访问通常以主导语言运行的在线空间时可能面临障碍的群体。然而，一些语言少数群体也可能拥有自己的在线空间，以他们自己的语言运行或使用自己的语言。因此，调查人员可能需要通过少数民族语 (包括土著语言) 进行搜索。

[128] 例如，在农村地区，互联网连接可能较少。

[129] 例如，无法物理访问高速互联网连接点，或者买不起设备或支付订阅费。

[130] 一般见人权高专办，一种基于人权的数据方法－ 2030 年可持续发展议程“不让任何一个人掉队”(2018 年，日内瓦)。可查阅 www.ohchr.org/sites/default/files/Documents/Issues/HRIndicators/GuidanceNoteonApproachtoData_CH.pdf.

[131] 模板见下文附件三。

C. 在线调查计划

115. 在开始开源调查前，应制定在线调查计划。[132] 该计划应涵盖：(a) 总体调查策略；(b) 具体在线调查活动。如果在线调查是使用传统技术 (如听取证人证词或收集物证) 进行的更广泛调查的一部分，则应将在线调查计划纳入主调查计划。调查人员应将性别观点纳入调查计划，确保调查涵盖所有与性别相关的问题，并考虑到技术获取性质的差异性。[133] 在线调查计划应涉及以下主题。

1. 目标和计划活动

116. 计划应具体说明开源调查的目标和优先事项、实现这些目标的拟议战略及落实目标的时限。

2. 风险管理策略

117. 计划应包括上述数字威胁和风险评估的主要结果 (如可能存在的网络威胁) 以及风险管理战略，包括如何识别、应对入侵或攻击并从中恢复。

3. 摸排行为人及合作机会

118. 开源调查人员可能希望对正在进行类似或重合调查的其他行为人的情况有所了解，从而评估他们的活动可能会如何相互影响，并探索潜在的伙伴关系和合作机会。这可能包括识别数字档案管理员、记者或其他保存可能与调查有关的在线内容的团体或个人。这种情况摸底还应考虑到其他行为人可能有的偏见和局限，这些偏见和局限可能导致第三方的调查结果不能充分反映特定情况的复杂性，或是如上所述，可能由于未调节对数字领域固有偏差而排除某些群体。如果建立了这种伙伴关系，则有必要为信息共享建立书面协议。

4. 资源

119. 计划应确定开展计划中活动所需的资源，包括人员配备、培训、工具和设备。对人员配备需求的评估可包括：执行任务所需的团队成员数量、能力、团队成员的包容性和多样性、对额外培训需求的评估。这可能包括评估必要的软硬件基础设施及长期保存数字资料的财务成本。

[132] 见下文附件一 (在线调查计划模板)。

[133] 关于如何纳入性别观点的进一步指导，见《将性别视角纳入人权调查：指导和实践》(联合国出版物，出售品编号：19.XIV.2)。

计划还应确保有专用资源，用于为调查人员提供对性别问题有敏感认知的心理健康专用资源，特别是在以下情况下：开源调查涉及图示内容；或者在调查人员或受牵连第三方很可能会由于身份或隐私暴露而遭到报复。[134]

5. 作用和责任

120. 如果进行团队合作或与外部伙伴合作，要明确界定开源调查人员的作用和职责应考虑到协调各项活动的必要性，包括避免重复活动和数据收集。此外，计划的这一部分应考虑具体调查可能需要哪些专门领域的专业知识，以及如果现有小组中没有专家，调查人员是否需要咨询或聘请专家。专业知识领域可能包括数字取证、卫星图像分析和数据科学。在某些专门知识领域，可能需要未雨绸缪，先找到来自不同性别背景和其他背景的专家，确保调查小组及其分析具有包容性和多样性。

6. 文件

121. 开源调查应以便于有效管理和遵守问责原则的方式记录在案。在法律诉讼中，这些文件应使调查人员能够证明所收集的证据具有相关性和证明力，并解释在线活动中采取或未采取的步骤及其原因。无论是自行执行任务还是主管授命执行任务，系统都应具备为特定调查活动（包括在线活动）创建任务的机制，如请求调查特定人员或进行其他查询。任务结果（包括报告）应说明所使用的方法和技术。报告时应将操作信息和调查信息分开。为保护调查来源和调查方法，可能需要对操作信息保密；而调查信息则必须在法律诉讼中披露。

122. 应定期审查在线调查计划，并在必要时进行修改。见下文附件一（在线调查计划模板）。

D. 复原力计划和自我护理

123. 虽然开源调查人员可能不会进行面对面的访谈或亲自到犯罪现场，但数字研究的特殊性意味着他们可能会观看、收集和分析大量的图示资料或其他会造成创伤的数字信息，这可能会导致二次创伤等问题。开源调查人员应了解自我护理原则，[135] 调查管理人员应营造重视自我护理、敏感对待性别和文化问题的组织环境。这项工作应在调查的准备阶段就

[134] 例如，调查人员可能受到网上仇恨言论的攻击或受到骚扰，这些攻击可能带有性别歧视色彩（例如，妇女和男女同性恋、双性恋、跨性别者、性别奇异者以及间性者调查人员可能面临高于平均水平的网上仇恨言论、人肉搜索、强奸威胁和其他基于性或性别的暴力威胁）。例如见大赦国际，“Toxic Twitter – a toxic place for women”。可查阅 www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/.

[135] 关于自我护理对人权调查领域工作人员的重要性的进一步讨论，见人权高专办，《人权监测手册》(2011 年，日内瓦)，第 12 章（创伤和自我护理），第 20 至 39 段。可查阅 www.ohchr.org/Documents/Publications/Chapter12-MHRM.pdf.

开始，通过制定计划以培养韧性，减轻调查的负面心理社会影响，这种影响可能因性别、文化和年龄而异。出于伦理道德原因，这样一项计划是必不可少的，这是促进和尊重调查小组每名成员人权的一部分。最大限度地增强人身和数字安全也至关重要。即使经过适当的培训，个人在压力之下也可能成为团队安全、信息安全和工作质量的隐患。应划出专门的时间和资源来确保计划得以妥当执行，特别是在预计在线调查可能涉及查看大量图形图像，包括暴力或其他令人不安的内容时。有多种策略可减轻观看图示内容可能带来的负面影响，这些策略往往归为三类：个人认识、尽量减少接触的策略和群体支持。

124. 首先，调查人员应了解自己和队友的基线行为，包括工作、娱乐、睡眠、饮食的模式，以便发现偏差并进行处理。制定调查人员两人一组的政策有助于发现问题，因为个人可能认识不到或不想承认自己的行为变化，而这些变化可能更容易被其他人注意到。团队成员应对可能引发强烈情绪的图示资料和其他材料的反应差异保持敏感和尊重，并认识到此类差异可能因个人、性别和文化群体而异，具体个人的反应也会随着时间的推移因所承受压力的程度和其他环境因素而发生变化。调查人员还应该认识到，对图示资料或令人震惊的内容做出情绪反应通常是很正常的，并不是软弱的表现，而是功能健康的体现，甚至是力量的标志。

125. 第二，应采取策略尽量减少接触有害内容。这方面常见的策略包括：当第一次观看可能含有图示资料的内容时，关闭音频；或者在音频并非即时分析任务所必需时，关闭音频，因为声音中含带许多能激起强烈情感的内容；尽可能缩小屏幕尺寸；在分析特定行为的背景而不是行为本身时，覆盖图示信息；对数据集中包含的任何图示内容进行标记，使个人不会在事先不知道将要看到什么的情况下看到该内容；在共享图示内容时彼此发出警告，以便减轻意外因素；两人一组进行工作；避免在孤立环境中工作或在深夜工作；根据需要，定时休息。

126. 第三，个人和组织应该在团队成员中培养一种群体意识，这可以起到保护作用——本质上是再现了在实地进行调查时可能缔结的同道情谊。这可以通过以下方式实现：定期情况汇报，可以减少孤立，帮助调查人员更好地了解其工作的积极影响；团队外出活动，包括庆祝调查工作的重要里程碑进展；关于复原力战略的团队培训。如果从个人、文化和结构层面着手，提高韧性的努力可能会特别有效。例如，增强个人能力，使之能批判性地思考自己在参与调查工作时的心理社会需求，并营造环境，认真对待调查工作的心理社会方面，明确和含蓄地鼓励支持性做法，接受包容性和多样性。

E. 数据政策和工具

127. 在调查过程中，应制定、实施和遵守关于数据处理、保存和销毁的政策。组织应酌情制定关于保留信息的政策（保留政策）和删除信息的政策（删除政策），并制定有关查阅信息（内部）和信息共享（外部）的策略。此外，就创建和使用虚拟身份、获得经批准的软件和所用工具而制定具体政策，也可能是有益的。

1. 数据政策

(a) 数据保留政策

128. 为了遵守众多数据保护法律和数据保留规定，数据保留政策非常重要。在某些情况下，对数据必须保留多长时间有最低要求，而在其他情况下，对数据最多可以保留多长时间设有限制。政策应概述持久数据存储和记录管理的方法，以满足法律和业务数据归档要求。不同的数据保留政策对法律和隐私问题与经济关切和需要知道的问题进行权衡，确定保留时间、存档规则、数据格式以及允许的存储、访问和加密方式。[136] 理解适用的规则对于制定此类政策是必要的。

(b) 数据删除政策

129. 在没有明确的删除和保留政策、没有日志记录删除内容、操作者、删除时间和目的情况下，删除数据集的一部分会造成严重问题，特别是在资料可能用于法庭审讯的情况下。调查人员应遵守有关删除数字数据的适用法规，并意识到，使用一种方法而非另一种方法可能会造成法律问题。

(c) 数据访问政策

130. 收集和处理数据（尤其是敏感数据）的组织应制定明确的政策，规定哪些人可以查阅哪种类型的数据。数据库或系统中的任何设置都应反映这一政策。

(d) 数据共享政策

131. 组织可能需要考虑制定与外部行为人共享数据的政策。如果与外部伙伴合作，则应签订共识备忘录或合同，确保合作伙伴遵守此类政策。

2. 信息管理

132. 在从事开源调查前，特别是在收集和保存数字材料前，调查人员、团队和组织应建立信息管理系统。此类系统有多种方案可供选择，本规程并不主张采用某一特定方案，而是在下文介绍对调查过程有帮助的主要功能，这些功能在某些情况下可能必不可少。此外，如第四章所述，应建立安全基础设施和规程。

(a) 调查管理系统

133. 调查管理系统用于记录作为调查一部分而开展的各项活动。并非所有从事调查的组织都有此类系统，但强烈建议建立这一系统，特别是对于较大的组织或调查团队而言。此类系统可用于分配任务和报告活动情况，有助于减少重复劳动，因而使整个流程结构清晰并尽可能高效。

[136] Yvonne Ng, "How to preserve open source information effectively", in *Digital Witness, Using Open Source Information for Human Rights Investigation, Documentation and Accountability*, Sam Dubberley, Alexa Koenig and Daragh Murray, eds. (Oxford, Oxford University Press, 2020), pp. 143–164.

(b) 信息和证据管理系统

134. 信息管理系统用于存储作为调查的一部分而收集的数据。信息管理系统应发挥两种不同功能：(a) 跟踪材料的收集和处理情况；(b) 分离可能用作证据的材料。

3. 基础设施——后勤和安全考虑

135. 无论是为从事开源调查的组织设计基础设施，还是决定独立调查员应使用哪些工具，都要考虑一些重要的后勤和安全问题。一般而言，系统开发有三种方法：(a) 定制系统和工具；(b) 使用互联网上现有开源或免费工具和软件；(c) 从第三方购买商业产品。每种方法各有利弊，其成功与否取决于调查人员工作的具体环境和背景。同样，本规程并不主张一种方法优于另一种方法，而是介绍每种方法的利弊，以及在决定使用哪种产品时应考虑的具体因素。

(a) 商业产品

136. 商业产品的优点是，私营企业可能有更好的安全基础设施，并能够提供持续和一致的技术支持。然而，商业产品的明显缺点是费用。此外，与第三方互动和对第三方的依赖对于试图对调查保密的组织而言可能是一个问题。许多商业产品都有闭源代码以保护其知识产权。商业产品也可能引起对数据所有权、数据可携性和可输出性以及与其他系统之间互操作性的关切。此外，公司可能会回应政府索取私人信息的压力。一个关键问题是，尽管公司有安全团队保护其产品和用户，但用户必须相信公司已经正确设计并将正确维护其系统，而且后期阶段不存在隐藏费用。

(b) 定制或自定义工具

137. 从头开始定制工具或对现有工具进行自定义设置的优点是，调查人员或组织对整个系统和全部数据保持控制，因此可避免与第三方互动。定制系统也更易于与其他定制系统整合。缺点是建立和支持此类系统需要时间、成本和专业知识，因而对大多数组织是一项挑战。此外，测试人员和用户有限的封闭系统可能难以识别漏洞，或难以获得足够反馈以实现功能最大化。

(c) 开源和免费工具

138. 开源工具是开发者公开发布源代码的工具，任何人均可自由使用或修改。一些商业产品有开源代码，而一些免费工具有闭源代码，但这些都是例外。最常见的情况是，开源工具是免费工具。对于预算有限的小型组织和有繁琐付费产品采购程序的大型组织而言，免费工具是值得考虑的重要代替方案。然而，对用户免费的工具可能会以其他方式盈利，例如向用户出售数据和分析，从而引发安全和隐私问题。此外，使用这些工具需要事先调研，以了解工具由谁创建，是否经过独立审计，以及是否可以持续。所有三个方面都可能破坏调查的可信性。特别是如果某一案件进入审判阶段，而工具受到对方质疑，工具在法律方面可能会出现问题。此外，需要为这些软件系统和工具创建备份计划及数据迁移和备份系统，以防系统和工具过时或没有可用的开发人员。虽然开源工具可能对

组织具有吸引力，部分是因为其他观点相似的团体正在使用，但调查人员必须对这些工具的运作方式以及使用这些工具在特定环境中可能产生的影响进行全面、独立的评估。

139. 在决定是否定制工具、使用免费试用或开源软件还是购买产品时，调查人员应遵循下文附件五提供的尽职调查指南。

六

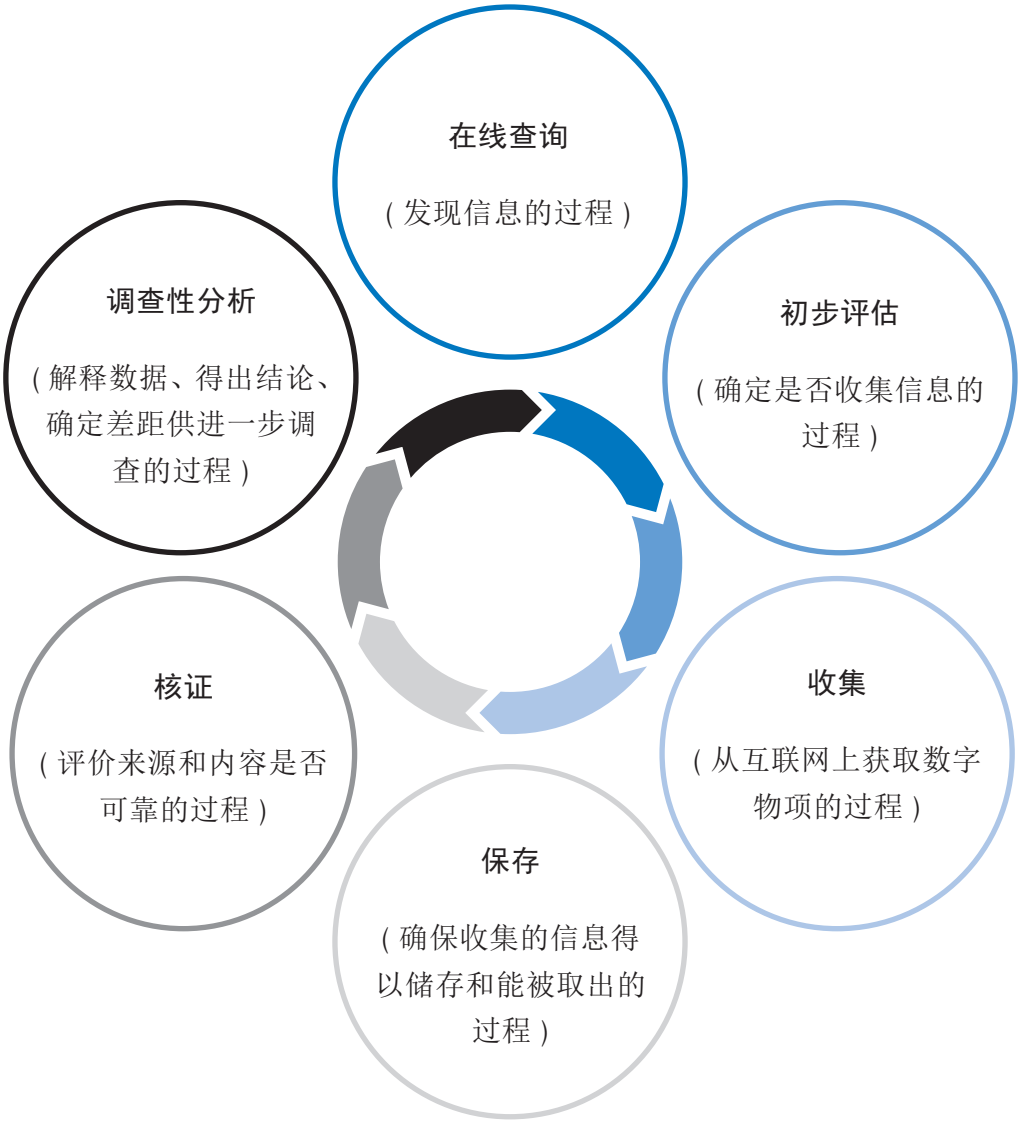
调查流程

本章摘要

- 调查流程有六个主要阶段，即：(a) 在线查询；(b) 初步评估；(c) 收集；(d) 保存；(e) 核证；(f) 调查性分析。总体而言，这些阶段是调查周期的一部分，在整个调查过程中可能重复多次，因为新发现的信息会引向新的调查方向。
- 调查人员应记录每一阶段开展的活动。这将使调查（包括保管链）便于理解、保持透明，并有助于确保调查的效率和成效，包括确保调查完整性和团队成员之间的沟通。

140. 开源调查需要仔细观察和系统性的查询，以便在复杂和动态的数字环境中确立事实。开源调查人员必须运用批判的视角审查在线内容，并能评估数字材料可以被扭曲或被操纵的方式。他们还应采用结构化的方法查询互联网，同时考虑到算法偏差以及与特定群体和在线信息的动态性质有关的开源信息可得性不平等问题。每个指称的事实都应经过严格审查。本章介绍了结构化的开源调查方法。下图显示了开源调查周期。需要注意的是，开源调查很少是线性的，并且通常需要重复这一流程鉴于案件构建的周期性特征。而且偏离这一顺序也可能有正当理由的。

开源调查周期



A. 在线查询

141. 在线查询有两个主要过程：(a) 搜索，即通过使用一般或高级搜索方法发现信息和信息源；(b) 监测，即连贯持续地审查一组固定来源，从中发现新信息。

1. 搜索

142. 在线搜索是以任务为导向的活动，旨在发现与既定目标或研究问题相关的新信息。搜索应做到结构化和系统化，包括以明确的研究问题、搜索参数以及关键词和运算符开展调查。[137] 不同的搜索引擎、搜索工具、搜索词和运算符会产生不同结果；因此，调查人员必须发挥一定程度的创造力和坚韧精神，利用各种途径和渠道寻找相关信息。除了利用搜索引擎在编入索引的网站上查找信息外，结构化搜索也可用于在社交媒体平台和数据库中查询。由于需要采取各种各样针对具体案件的方法，调查人员应仔细记录工作流程，以便可以在调查报告的方法部分进行解释或在法律程序中作证。这可能是一个追溯的过程，不一定需要与研究本身同步进行。不过，文件记录应尽可能同期进行。对结构化搜索的文件记录应包含以下信息：

(a) 目标和研究问题：阐明在线搜索力图解决的问题，同时铭记上文所述客观性原则；

(b) 事实、假设和未知事项：如果事实已经确立，就将已知事实作为出发点。也可以在线索信息或逻辑假设的基础上开始工作，即使这些信息和假设尚未得到验证。然而，任何假设都必须如实记录。最后，在调查之初阐明知识差距或其他“未知数”可能有所裨益。明确区分信息类别将有助于通过理清搜索词及其依据来防止出现偏颇或扭曲的结果。

(c) 搜索词和关键词：为了进行有针对性的搜索，调查人员应根据案件涉及的一种或多种理论，建立符合客观性原则的关键词清单。调查人员最好使用所有相关语言和文字中的关键词，并谨慎对待可能出现的搜索结果所含内容过多或不足的情况。尽管案件存在差异，但有些一般主题应纳入关键词清单，如重要地点、姓名、组织、日期和相关话题标签。在具体调查的背景中，确定哪些信息可能符合认定有罪和无罪的条件也可能有所帮助。

(d) 搜索和搜索引擎：调查人员应追踪所进行的搜索，记录获得相关材料的途径，包括得出相关内容所用的术语、运算符和搜索引擎。调查人员无需记录全部搜索结果，因为这将造成过度的负担，而且鲜有证据价值。

2. 监测

143. 监测是指随着时间的推移跟踪某一确定的信息源，例如某一特定主题。其目的

[137] 布尔操作符是一些简单词语，如“和”、“或”、“不”，可用于“在搜索中合并或排除关键词，从而产生更有针对性和更有成效的结果”。见 Alliant International University Library, “What is a Boolean operator?” 可查阅 https://library.alliant.edu/screens/boolean.pdf.

是跟踪固定来源产生的不断变化的内容。在线监测应是一项结构化的活动，利用已知和曾评估过的在线来源列表，如网站或社交媒体账户，并针对既定目标持续进行的搜索查询。例如，见以下各类信息来源：

(a) 网站和社交媒体账户：调查人员应维护被监测的网站和账户资料的工作列表，其中应包括实施监测的理由、监测工作的负责人、监测工作的执行者以及监测的频率。

(b) 话题标签和关键词：调查人员还应该维护并定期更新被监测的话题标签和关键词的工作列表。

(c) 自动化：监测可能涉及使用自动化工具，例如，定期对特定网站进行搜索或运用某些参数进行搜索。应始终记录此类工具的使用情况，包括工具名称和版本以及输入其中的信息。

3. 偏见

144. 在进行结构化搜索和监测活动时，开源调查人员必须始终对偏见保持警惕，既包括自身的认知偏见，也包括在线信息中的固有偏见。例如，如果调查人员正在搜索关于强奸的信息，所得到的大部分数据或网上讨论的问题很可能涉及对育龄妇女的婚外强奸。搜索结果可能会少报不太明显或不太见诸报道的强奸类型，如男子和男童、男女同性恋、双性恋、跨性别者和间性者、老年妇女遭受的性暴力以及婚内强奸案件。

145. 另一个实例是对网上仇恨言论煽动的暴力进行的调查，因为此类言论通常包含和依赖不易被人工调查人员或机器发现的暗语和符号。特别是如果调查人员来自目标社区之外，他们可能不了解用于煽动仇恨或暴力的术语和符号在特定文化和背景中的使用情况。使情况更为复杂的是，网上的仇恨言论往往是为避免被机器或人工监测员发现而故意设计，以免被从网络平台上删除，虽然事实上其目的是为了煽动对目标人群的暴力或歧视。为帮助克服在发现煽动歧视、敌意或暴力方面的困难，调查人员应采用基于人权的测试，例如像《拉巴特行动计划》中所提供的对构成煽动歧视、敌意或暴力的鼓吹民族、种族或宗教仇恨言论的禁止。[138]

146. 归根结底，调查人员对抗“机器中的偏见”和自身偏见的最佳方式是意识到存在这种偏见的可能性，承认风险，并在可能情况下采取积极步骤，通过研究与特定背景或与犯罪或事件有关的术语和符号，并通过扩大在线查询的规模和多样性抵消偏见。在涉及性暴力和性别暴力案件中，以及在幸存者被污名化和使用暗语的任何其他犯罪中，调查人员应咨询专家，这些专家或能识别和分享幸存者和施害者在网络空间交流时经常使用的暗语和交流方式。[139]

[138] 见人权高专办，“表达自由与煽动仇恨：人权高专办和《拉巴特行动计划》”。见 www.ohchr.org/zh/freedom-of-expression.

[139] 例如，见 Koenig and Egan, “Hiding in plain site: using online open source information to investigate sexual violence and gender-based crimes”。

B. 初步评估

147. 在从互联网上收集内容前，开源调查人员应初步评估所发现的任何材料，以免过度收集，遵守数据最少化原则和重点调查原则，并确保材料收集工作不会侵犯个人隐私权。开源调查人员在确定是否应从互联网上收集某个数字物项时应考虑以下因素。

1. 相关性

148. 开源调查应确定某一数字物项是否与特定调查有初步关联。任何物项是否相关取决于其内容和来源，还取决于调查目标和对情况的了解。在调查早期阶段，可能很难知晓哪些内容相关，而这可能导致调查人员易犯过度收集的错误。尽管如此，开源调查人员应能阐明为什么认为某一物项可能相关，而且这一评估应记录在案（例如，通过对用户友好的简单标签或存储系统将收集的信息与正在调查的地点、日期、事件、人物或违法行为类型等联系起来）。

2. 可靠性

149. 开源调查人员应确定数字内容所含信息或主张是否看上去可靠，通过审查和评价文档内容以及其中的背景信息。其中包括检查嵌入的元数据、关联信息和来源。[140] 这一过程应包括努力查明材料的原始来源，而这可能需要追踪数据的在线出处、上传者或作者。

3. 删除

150. 开源调查人员应评估某一数字物项是否可能从互联网或公众访问中删除。当内容很可能被删除时，应收集最可靠的已知版本内容，尽管更早或更好的版本已经被进一步核实和调查了。可以根据若干因素评估内容被删除的可能性，包括推定的来源身份、内容的位置以及内容与服务提供方服务条款的兼容性。例如，最有可能被删除的内容包括可能对确定犯罪或违法行为具有较高证据价值的图片或攻击性内容。

4. 安全性

151. 开源调查人员应确定数字物项是否可以安全地收集，或者是否可以而且应该采取更多预防措施。如果从可能包含使内部系统遭到破坏的受损物项的网站收集数字物项则更可能引起顾虑。

5. 嗣后职责

152. 开源调查人员应确定保管数字物项可能产生的职责，如遵守数据保护法以安全方式进行保存的职责。[141]

[140] 见下文第六章 E 节关于核证的内容。
[141] 见下文第六章 D 节关于保存的内容。

C. 收集

153. 收集是指采取截屏、转换为 PDF、取证下载或其他获取形式取得在线信息的行为。数字内容一经发现并认定与调查有关，而且初步证明与目标相关而且可靠，调查人员必须确定适当的收集方法。收集方法可能有所不同，具体取决于在线内容是否在司法程序中具有潜在的证据价值，是否用于决策目的或作为决策依据，或者是否仅有助于内部工作成果。在仅涉及工作成果的情况下，截图或转换为 PDF 可能就已足够，而具有潜在证据价值的内容可能需要更全面和完善的获取方法（例如见下文所述分配散列值的方法）。

154. 可以按照标准操作程序手动收集在线内容，也可使用各种工具或脚本自动收集。无论过程如何，最好是在收集时就取得下文所列各项信息。这些信息可能有助于确定数字物项的真实性。在数字物项作为证据出示的法律诉讼中，特别是当作者或创建者无法确定、无法找到或无法作证时，这一点可能尤为重要。开源调查人员应以原始格式或尽可能接近原始格式的状态收集在线内容。由收集过程引起的任何改动、变化或转换都应记录在案。

155. 下文为收集什么以及如何收集提供了指导。可以利用一些工具协助收集以下信息，也可以使用人工方式收集。尽管收集以下全部信息被视为最佳做法，但前三项（统一资源定位符 (URL)、超文本标记语言 (HTML) 源代码和整页获取）是在法庭上提供证据的最低标准。当然，这种标准因情况而异，但获取下列全部要素将为应对任何情况提供坚实的基础：

(a) 目标网址：应记录所收集内容的网址，也称统一资源定位符 (URL) 或标识符 (URI)；

(b) 源代码：如适用，调查人员必须获取网页的 HTML 源代码。HTML 源代码包含比网站可见部分更多的信息。HTML 源代码将有助于对收集的材料进行认证；

(c) 整页获取：调查人员应首先对显示日期和时间的目标网页进行截屏。这一过程是为了尽可能体现收集时看到的内容；

(d) 嵌入的媒体文档：例如，如果下载带有视频或图像的网页，也应从网页中提取和收集这些特定的物项；

(e) 嵌入的元数据：调查人员应收集数字物项的附加元数据（如果有且适用）。元数据可能因来源而异，但常见元数据包括上传者的用户标识符；帖子、图片或视频标识符；上传日期和时间；地理标签；话题标签；评论；以及注释；

(f) 背景数据：如果背景内容与理解数字物项有关，也应收集。背景数据可能包括对视频、图像或帖子的评论；上传的信息；和（或）上传者／用户信息，如用户名、真实姓名或个人介绍。是否应收集周边的信息，需要根据案件和数字物项的具体情况而定；

(g) 收集数据：开源调查人员必须记录与收集工作有关的所有相关数据，如收集者姓名、用于收集信息的机器的 IP 地址、所用虚拟身份（如有）以及时间戳。调查人员应确保系统时钟准确，最好与网络时间协议服

务器同步。采取这一步的原因是确保在所收集的文档中准确体现与时间有关的元数据。如果用虚拟身份访问所收集的信息，应指出这一点；

(h) 散列值：散列值是一种独特的数字识别形式，通过使用加密技术确认收集的内容独一无二，而且自收集时起未经修改。在收集时，应由开源调查人员手动添加散列值，或由收集工具自动添加。有许多不同类型的散列可供选择，其标准也随时间推移而演变。调查人员应根据当前公认标准评价应使用哪种散列。[142]

156. 在自动收集的情况下，所述一些过程可由旨在收集相关内容和元数据的工具执行。对于收集到的每个物项，应编写一份包含上述信息的技术报告，以便日后确定物项的真实性。如下节所述，背景信息和所有类型的元数据应始终与数字物项共同储存和保存。

D. 保存

157. 在线信息的持久性和可得性经常变动。社交媒体平台可能根据使用条款从其平台上删除内容，用户也可能选择删除或编辑自己上传的内容。此外，在线信息很容易脱离背景、丢失、删除或损坏。[143] 要使数字材料保持可访问和可使用状态，以确保法律问责，就需要在短期和长期对其进行保存。[144] 通常，数字保存的目的是使材料保持可访问状态。[145] 然而，当为确保法律问责而进行数字保存时，其目的是以有助于确保数字材料可访问、真实可靠和可供问责机制使用（包括在法律诉讼中可被采用）的方式管理和维护。因此，在开展调查的情况下，数字保存要求长期对信息进行维护，以便所收集的物项在真实性得到充分确认的情况下始终能被目标用户独立理解。

158. 为实现长期保存的目的，存储硬件和格式可能需要更新，以确保始终能用当代设备对材料进行访问。

1. 必须长期保护和保存的数字物项的属性

159. 根据档案学家的观点，必须长期保护和保存的数字物项的属性包括：真实性、可得性、可识别性、持久性、可呈现性和可理解性。各项属性简述如下。

(a) 真实性

160. 真实性是证明数字物项自收集时起保持不变的能力。真实性要求数字物项在存档期间保持不变，或对其所作任何修改均记录在案。[146]

[142] 美国国家标准和技术研究所是可就现行标准提供指导的一个组织。见 www.nist.gov.

[143] Ng, “How to preserve open source information effectively”.

[144] 同上，第 143 页。见联合国教育、科学及文化组织，“数字保存的概念”。可查阅 https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-preservation#:~:text=Digital%20preservation%20consists%20of%20the,hardware%20tools%20acting%20on%20data.

[145] Ng, “How to preserve open source information effectively”.

[146] 同上。请注意，此处“真实性”一词的用法有别于其在法律上的用法。

(b) 可得性

161. 可得性是指数字物项可以获取，既指简单意义上持续存在和可以检索，又指法律意义上确保获取和使用物项所需的适当知识产权。[147]

(c) 可识别性

162. 可识别性是指数字物项可被参考引用的能力。数字物项必须可识别，并可与其他数字物项相区别，例如，可用唯一识别码这样的标识符进行记录。[148]

(d) 持久性

163. 持久性用技术术语表述是指数字物项的完整性和可用性。数字物项的位序列必须完好无损、可处理和可检索。[149]

(e) 可呈现性

164. 可呈现性是指人类或机器通过适当的硬件和软件使用数字物项或与数字物项互动的能力。[150]

(f) 可理解性

165. 可理解性是指目标用户解释和理解数字物项的能力。[151]

2. 特定调查问题

166. 调查人员还应考虑保全过程中可能或者将要出现的特定调查问题并制定相应计划。

(a) 保管链

167. 保管链是指按时间顺序记录一项信息或证据的历任保管员，并记录任何此类证据的控制措施、日期和时间、移交、分析和处置情况。数字物项一经收集，应通过建立适当的数字保全制度来维护其保管链。

(b) 证据副本

168. 证据副本是调查人员收集的原始形式的数字物项，不得改动或变动。数字物项应以原始形式存储。这意味着以数字物项被收集时的全部格式保全一份干净的原件。

(c) 工作副本

169. 出于分析目的，应创建一个或多个数字物项副本，并单独存储，以便调查人员可以在工作中使用副本而不是原件。这样可以尽量减少对原件的处理，减少其遭破坏或改动的风险。对物项所作的任何和一切变动，包括制作副本，都应记

[147] 同上。
[148] 同上。
[149] 同上。
[150] 同上。
[151] 同上。

录在案。如有可能，证据副本和工作副本应使用单独的存储系统。

(d) 存储

170. 存储有助于确保数字物项的持久性以及查找和提取它们的能力。存储不应被视为被动过程，而应视为涉及持续的、受管理的任务和责任的主动过程。它包括永久存储（使用存储介质），还包括管理存储层次、更换介质、检查错误、检查固定性（确保物项未经修改）、灾后恢复以及定位和返回存储对象。[152] 数字信息可存储在现场（在线或离线）或异地（在线或离线）。[153] 数字内容的存储选项既包括本地硬盘或本地可移动介质，也包括作为局域网、远程服务器或云存储系统的一部分的联网驱动器。与存储选择相关的考虑因素包括：存储容量（空间）；访问和控制；备份；相关法律；信息安全和数据保护。存储选择还应考虑速度、可得性、成本、可持续性、存储管理和检索系统。[154]

(一) 备份

171. 如果发生数据丢失或错误，档案管理员或技术人员可以尝试恢复数据。理想情况下，这些数据先前已在另一位置进行了备份或复制。信息技术专家建议在至少两种不同类型的存储设备上保存至少三份数据副本，其中至少有一份副本在地理位置上与其他副本分开。

(二) 退化

172. 存储面临的一项挑战是介质会随着时间的推移而退化。档案管理员可以通过使用特别耐用的介质类型来降低发生存储故障的风险；然而，任何存储设备最终都会有缺陷或产生缺陷、出现磨损或者发生随机故障。即使没有完全出故障，随着存储介质的衰退，也会发生数据错误或文件损坏。因此，有必要维护备份副本并定期监测存储基础设施和所存储文件的持久性，例如定期检查随机样本的散列值，以确保没有发生退化。

(三) 过时

173. 一旦访问数据所需的硬件不再能够合理获取或维护时，数字化文件就过时了。任何存储介质无论多耐用，都有过时的风险，造成所存储数据提取困难或无法提取。因此，调查工作应确保维护并视需要更新存储介质，以保持数据的可呈现性和可用性。

(四) 恢复

174. 数字化文件可能被意外或刻意删除。当用户在计算机上“删除”一个文件时，被删除文件的内容将保留在存储介质上，

[152] 同上，第 154 页。

[153] Shira Scheindlin and Daniel J. Capra, *Electronic Discovery and Digital Evidence in a Nutshell* (Saint Paul, West Academic Publishing, 2009), pp. 21–22.

[154] Ng, “How to preserve open source information effectively”, p. 156.

直到被另一个文件覆盖。[155] 因此，计算机或其他存储介质上的活动越多，被删除文件遭到覆盖和变得不可恢复的速度就越快。大多数计算机的操作系统都内置了软件实用程序，可以恢复被删除的文件。此外，还可购买数据恢复软件，不时用于“撤销删除”文件。为了访问被删除的数据，开源调查人员可能需要寻求信息技术专家的帮助。

(五) 刷新

175. 刷新是指将内容从一个存储介质复制到另一存储介质。刷新只针对媒介过时的问题，并非全面的保全策略，但应被视为更大的保留策略的组成部分。[156]

E. 核证

176. 核证是指确定在线收集的信息的准确性或有效性的过程。开源信息的核证可以纳入全来源分析（包括非公开来源和保密来源信息）进行，也可以完全基于开源。核证细分为三项独立考虑因素：来源；数字物项或文件；内容。这些因素应放在一起来看，并进行一致性比对。

1. 来源分析

177. 来源分析是评估来源可信度和可靠性的过程。在线环境给来源分析造成挑战，因为许多来源是匿名的或使用假名。为了正确分析信息来源，开源调查人员必须首先识别要分析的正确来源，这意味着将信息归因于原始来源。归因分析是指确定数字信息的来源，来源可能是某一特定网站，特定账户或平台的订阅者或用户，或者制作、创建或上传特定内容的人的身份。归因分析并不总是可行，可能需要采取额外的在线和现实世界调查步骤或借助先进的搜索和分析技术。识别制作者的身份固然有帮助，但制作者身份不明对于确定在线物项的真实性通常并不关键，因为还有其他认证开源信息真实性的方法。

(a) 出处

178. 出处涉及某一事物的起源或最早的已知存在。就在线内容而言，出处可以指最早在网上出现的版本或上传到互联网之前的原始物项。对于在线内容，出处最好是指“在线找到的第一个副本”而不

[155] Scheindlin and Capra, *Electronic Discovery and Digital Evidence in a Nutshell*, p. 24.

[156] Cornell University Library, “Digital imaging tutorial”. 可查阅 http://preservationtutorial.library.cornell.edu/tutorial/preservation/preservation-03.html.

是“第一个在线副本”，因为原件可能已被删除。即使调查人员确信他们已经从在线公开来源找到了诸如视频或其他信息的第一版，也无法确定其出处，因为存在电子邮件和私人消息群组等非公开渠道，这些渠道可能在该物项在网上公开出现之前已被用来分享该物项。[157]

(b) 可信度

179. 来源的发布历史、在线活动和互联网存在可能包含损害或支持来源可信度的相关信息。开源调查人员应审查来源的在线存在和发布历史，这甚至可能有助于发现故意欺骗的企图。例如，如果来源发布了关于某个特定国家的事件的帖子，那么此人前后的发帖是否表明其确实身处该国？

(c) 独立性和公正性

180. 调查应审查来源的公正性。为此可以研究与个人有关联的任何团体、组织或附属机构，以及个人的营生方式和资金来源。是否与正在调查的案件或事件中的任何当事方有联系或关系？在考虑来源独立性时，应审查他们是否可能与相关实体（如冲突当事方）存在关联。来源的意识形态以及是否从属于任何团体可能也很重要。调查人员应研究并揭示所有来源的深层动机、利益或图谋，以及这些动机、利益或图谋对来源真实性的可能影响程度。

(d) 具体性

181. 信息和说法越精确，就越容易被证明或证伪。宽泛模糊的说法往往更难严格评估。

(e) 衰减

182. 在事件发生当下起草的相关文本往往被认为比事件发生后很久才制作的文本更可靠。[158] 在数字文本创建时间不明的情况时，这一因素给开源调查人员带来挑战。

2. 技术分析

183. 技术分析是指对数字物项本身进行分析无论是文档、图像或是视频形式。为了检验文件的完整性，即文件是否已被数字化改动、操纵或修改，开源调查人员不妨对文件进行数字取证检查，有时也称为数字调查分析。这种分析由以下几部分构成。

(a) 元数据

184. 元数据是描述并提供关于其他数据信息的数据。它们可以由生成物项的用户、其他用户、通信服务提供商或任何在其上创建、传输、接收或查看数据的设备创建。元数据与对物项及其生成、传播或改动情况的描述有关。元数据可能包

[157] 例如，一个用户可能将照片通过电子邮件发送给另一个用户，后者之后将照片上传到社交媒体。因此，这张照片的源头是电子邮件发送者，而不是发帖者。

[158] Institute for International Criminal Investigations, *Investigators Manual,* 5th ed. (The Hague, 2012), p. 88.

括文件的创建者、创建日期、上传数据、修改情况、文件大小和地理数据。元数据可以嵌入到文件中、在网页上显示，或存在于源代码中。一些元数据可能在上传前或上传过程中被剥离，或因使用社交媒体应用程序而被剥离，但如果有元数据可用，则应进行审查，查明是否有助于确定真实性。原始元数据可能会丢失，因为平台通常会对上传的媒体进行转码优化，便利在线查看、分享或回放。在这种情况下，元数据将反映新的文件，而不是原始文件。在元数据被剥离的情况下，开源研究人员应寻找其他方法来核实物项。

(b) 可交换图像文件格式数据

185. 可交换图像文件格式是一种元数据，它说明了数码相机、扫描仪以及处理数码相机记录的图像和声音文件的其他系统所使用的图像、声音和辅助标签的格式。

(c) 源代码

186. 源代码是任何网页或软件背后的编程。就网站而言，任何人都可以使用各种工具查看这些代码，甚至可使用网络浏览器本身查看。使用一些免费提供的工具就可以很容易地查看网站的源代码。它可能包含元内容或被隐藏或篡改的内容，并会显示链接结构和失效链接。

3. 内容分析

187. 内容分析是对视频、图像、文档或报表所包含信息的真实性和准确性进行评估的过程。内容分析同样是多方面的，包括分析视觉线索，例如利用元数据证实图像之类的方法。在线环境的特点引发了许多问题，这些问题会影响在线开源信息实际的或感知到的有效性或真实性。这些问题包括循环报道、断章取义和曲解。内容数据是包含在数字物项中的数据，如视频、图像、音频记录、文档或非结构化文本。

(a) 独有标识

188. 调查人员在接到核实视觉内容的任务后，应首先寻找独有或标识特征。这些特征可能包括建筑物、动植物、人、符号和徽章。在分析人类特征以识别特定人员时，应尤其谨慎。[159] 识别工作通常需要专门技能，例如那些由法证专家通过长期积累和专门培训获得的技能。由未经培训的专业人员进行的非专业分析可能不准确、有偏见和(或)存在其他问题。

[159] 使用工具或人体分析(例如面部识别、步态分析等)对人体特征进行法证分析和鉴定，需要法证专家。见Nina M. van Mastrigt and others, "Critical review of the use and scientific basis of forensic gait analysis", *Forensic Sciences Research*, vol. 3, No. 3 (2018), pp. 183–193 (可查阅 www.tandfonline.com/doi/full/10.1080/20961790.2018.1503579); Royal Society and Royal Society of Edinburgh, "Forensic gait analysis: a primer for courts" (London, 2017) (可查阅：https://royalsociety.org/-/media/about-us/programmes/science-and-law/royal-society-forensic-gait-analysis-primer-for-courts.pdf)。另见European Network of Forensic Science, *Best Practice Manual for Facial Image Comparison* (2018) (可查阅 http://enfsi.eu/wp-content/uploads/2017/06/ENFSI-BPM-DI-01.pdf); National Center for Audio and Video Forensics, "Height analysis of surveillance video" (可查阅 https://ncavf.com/what-we-do/forensic-height-analysis)。

(b) 可客观核实的信息

189. 通常，从识别哪些信息可视为“可客观核实的信息”入手，会有所帮助。例如，某天的天气、指挥官的姓名和军衔或者建筑物的位置都可以客观地加以核实。对开源材料的评估应包括对照此类可客观核实的信息，审视开源材料的内容。

(c) 地理定位

190. 地理定位是指对物体位置、活动位置或物项生成位置进行识别或估计。例如，可以利用地理定位技术确定从互联网上下载的视频或照片的拍摄地点。这种技术包含了利用照片中的独特地理特征识别这些特征在地图上的实际位置。

(d) 时间定位

191. 时间定位是对一条信息（通常是视觉图像）所描述事件的日期和时间进行印证。例如，可以通过检查日照形成的阴影长度以及其他指标，确定照片拍摄的时间。

(e) 完整性

192. 不完整的文档或视频片段仍然可以作为证据，但其中的空白可能会影响物项的证明力。因此，在收集开源信息时，必须完整地获取目标文件，并酌情捕获周围环境。

(f) 内部一致性

193. 可以对来自在线公开来源的单条信息或来自特定来源（和（或）拥有相同出处或制作者的来源）的一组信息进行内部一致性评估。对单条在线信息内部一致性的评估旨在确定信息本身是否连贯一致。一条或一组内部一致的信息不应自相矛盾。

(g) 外部印证

194. 外部印证由不属于数字物项本身、但与物项内容吻合从而支持物项内容真实性的信息提供。

F. 调查分析

195. 调查分析是审查和解读事实信息从而得出与裁决或立案有关的实质性结论的做法。开源信息数量庞大，质量参差不齐，因此有必要采取结构合理的分析办法。

196. 在进行某些类型的分析之前，可能需要对开源信息进行预处理。处理工作可能涉及外语翻译或不同数据集的汇总，以便利分析个人行为、地点和物体，以及关系或网络、动向、活动或交易；还可能涉及改变数字物项的性质或格式，使其与特定软件兼容。常见的数据处理类型包括：

(a) 翻译：如果数据使用的语言不是调查人员使用的语言，或者没有经过必要的材料审查软件处理，在采取进一步措施之前，可能需要翻译数据；

(b) 汇总：调查人员可能需要将不同的数据集汇总成一个更大的数据集，以便进行分析；

(c) 调整格式：为了使数据更容易被搜索或提取，调查人员可能需要改变数字物项的格式。

197. 建议只处理数字物项的工作副本，而不是原件或证据副本。对数字物项的任何处理均应记录在案。如果调查人员利用数字技术处理数据，例如，利用包括自然语言处理和深度学习在内的算法分析数据，那么他们必须了解处理此类数据存在的偏倚风险。

198. 处理后的信息可供分析。开源信息的分析工作成果将因目的、基础来源信息的类型和范围、制作时间表和受众而有所不同。这些成果将视调查需要编制，可包括图表、摘要、词汇表、词典和视觉辅助材料，包括地图和映射工作。[160]

199. 调查人员应采用严格标准，确保分析成果所载数据和结论的客观性、及时性、相关性和准确性，并保护隐私和顾及其他人权考虑，特别是在处理个人身份信息时。这类信息只有在调查人员征得所涉人员同意的情况下才能纳入成果，并且直接服务于调查目的。还应顾及针对这类信息使用的法律和伦理限制。[161]

200. 以下各节列出了可用于借助开源信息推进调查目标的常见分析类型。

1. 图像 / 视频比对分析

201. 比对分析或比对科学是将物体、人和 / 或地点的特征同其他未知项和 / 或已知项进行比对的过程，其中至少有一个有关物项是图像。比对分析是对图像和视频内容的分析，包括分析不同物项和特征之间的比对要素，以及它们的图像质量和视觉环境 (光线、视角等)。虽然现在许多非专业人士了解图像比对分析的基本知识，但拥有视频取证分析和 / 或数字取证合格资质的专家的协助可有助于作出科学分析，包括出具专家意见。这种专门知识还有助于人权和其他类型的调查，使调查结论更有分量。

2. 图像 / 视频解析

202. 与图像 / 视频比对相关联的是图像 / 视频解析，涉及分析数字物项，以理解其中的视觉内容。例如，对枪击、伤口、血液、车辆、武器和军事资产的分析，或者对移动车辆的速度或个人年龄的分析，都属于图像 / 视频解析的一部分。分析人员出于调查目的可进行图像 / 视频解析，法证或专题专家为了在法律诉讼或人权调查结果中确定事实也可进行相关解析。

3. 空间分析

203. 空间分析或地理空间分析可包括对提供地理坐标或地名的物项进行视觉内容分析和元数据分析。空间分析包括以适当的分辨率检查不同物体和景观特征，并对照卫星或其他图像、地理数据和地图、

[160] 见下文第七章 (报告调查结果)。
[161] 见上文第三章 (法律框架)。

适当的案例和背景知识以及地理信息系统[162]工具进行检查。

4. 行为人映射

204. 行为人映射是一种了解关键行为人并确定权力关系和影响渠道的技术。[163] 因此，首先是要识别关键行为人，然后绘制出他们之间的关系。

5. 社交网络分析

205. 与行为人映射类似，社交网络分析是对人员、团体、组织、计算机、统一资源定位符和其他相联系的信息 / 知识实体之间的关系进行映射和衡量。[164] 人员和团体通常被称为节点，而链接则显示节点之间的关系。社交网络分析利用社交媒体和其他移动平台或网络平台上的联系来建立和了解个人之间的关系。调查人员可以手动或使用分析软件来分析联系或链接数据。

6. 事件映射

206. 事件映射是一种分析技术，用于确定不同事件之间的时间和地理关系，在国际犯罪和国际侵犯人权行为的背景下，这可能是指此类侵权或犯罪行为（包括该事件前后的事件）的地点。事件映射还可能包括对其他相关事件进行映射，例如被控施害者在何时何地发表了声明。

7. 犯罪 / 侵权行为模式分析

207. 在国家执法的背景下，犯罪模式是指执法部门接报或发现的一组两起或两起以上的犯罪，这些犯罪具有独特性，因为它们在犯罪类型上至少有一个共性：罪犯或受害人的行为；罪犯、受害人或目标的特点；取走的财产；发生地点。[165] 同样，可以根据开源信息确定在国际刑事和人权案件中犯罪和侵权行为的模式。

[162] 地理信息系统是管理和分析空间数据的计算机数据库。

[163] 人权高专办，《人权监督手册》，第 8 章（分析），第 24 页。

[164] Orgnet，“Social network analysis: an introduction”。可查阅 www.orgnet.com/sna.html.

[165] International Association of Crime Analysts，“Crime pattern definitions for tactical analysis”, Standards, Methods and Technology Committee White Paper 2011-01, p. 1.

七

报告调查结果

本章摘要

- 开源调查的结果，即收集的数据或从这些数据中得出的结论，可以口头、可视化或书面形式报告。
- 调查人员在决定 (a) 所要采用的格式以及 (b) 所要列入的数据时，应考虑何种格式最适合其任务规定和目标受众，同时考虑到受众技术素养以及无障碍性、客观性、透明度和安全性等因素。

208. 本章介绍了开源调查（包括方法、原始数据和分析结果）的呈现或报告方式。许多情况下，开源信息会与通过其他调查方法收集的其他信息一并呈现。呈现形式可以多种多样，包括书面报告、口头报告或可视化报告，或这些形式的任意组合。报告可供内部使用或对外公布，并可依据若干因素划分为专家报告和非专家报告。报告应确保以下要素：

(a) 准确性：报告应准确反映收集的数据。[166] 应列入可开脱罪责的信息，并对一切编辑或空白作出解释；

(b) 归属性：报告应明确区分公共域内容或一般非保密信息；机密信息或受到其他限制的信息；反映调查人员和/或其他专业人员的判断或意见的内容。调查人员或其他报告开源信息的人员还应进行尽职调查，在使用可能属于他人的内容时获得适当许可，例如保障一切必要的知识产权；

(c) 完整性：调查结果应表明基础数据的完整性，特别是在有数据被刻意排除的情况下；

(d) 保密性：尽管材料是在开源环境中找到的，但报告应考虑哪些材料应排除或编辑，以保护机密性，或尽量减少风险，特别是对与开源信息有关的消息来源人士、证人、受害人和社区成员的潜在风险；

(e) 语言：报告应使用中性语言，避免使用煽动情绪或带有感情色彩的语言。报告应清楚陈述事实，不过多使用形容词或强调手法。报告需要以对性别问题有敏感认识的语言编写。理想情况下，除了调查人员或调查机构使用的任何官方语言之外，还应以受影响社区的语言提供公开报告；

(f) 透明度：报告应清楚说明调查人员如何开展工作，以及他们的目标、过程和方法。通常，这方面内容会列入报告的方法部分，但透明度还应指导整个文本的描述。描述应在不造成安全漏洞（例如泄露机密信息）的情况下力求透明。

A. 书面报告

209. 开源调查可以书面形式呈现，既可以是内部报告和提交客户的报告，也可以是公开报告。书面报告是通报分析结果的一种方式，包括非政府组织、调查委员会、实况调查团和联合国的报告，以及提交法院或法庭的专家报告等。[167] 数字开源信息往往会与其他形式的公开和非公开来源数据和分析进行整合。书面报告应分析所收集的信息，以便得出合乎逻辑的结论、估计和预测。报告应体现合理的方法并能够向目标受众解释该方法。报告中基础信息的真实性和完整性至关重要。不良数据会导致错误结论。[168]

[166] 见上文第二.B 章（方法原则）。

[167] 有关数字开源调查书面报告的例子，例如见 Human Rights Investigations Lab, "Chemical strikes on Al-Lataminah: March 25 & 30, 2017-a student-led open source investigation" (Berkeley, Human Rights Center, University of California, Berkeley, School of Law, 2018)。

[168] 建议视具体情况和保密要求开展同行审议，确保数据以及从中得出的分析和结论的准确性和质量。

210. 除非有正当和明确理由，例如，需要对某些在线调查技术、方法和来源保密，否则书面报告应包含以下部分：

(a) 调查目标：报告应包含调查目标、基本任务或客户指示，包括界定明确、表述清晰的研究问题；

(b) 方法论：报告应包含研究方法，以实现可复制性，并使受众能够了解和评估调查信息和结果的可信度，包括调查所涵盖的内容；

(c) 开展的活动：报告应概述所开展的对调查结果或分析质量评估具有重要意义的活动，包括概述为确定基础数据所开展的活动、收集了什么、分析了什么；

(d) 基础数据和来源：报告应对基础数据、包括数据来源和质量进行说明；

(e) 空白或不确定性：报告应指出基础数据或分析中可能对调查结果有重大影响的任何空白或不确定性；

(f) 结果和建议：报告应包含调查人员对数据的解读或基于数据分析的结论，并指出限制条件和新线索。

B. 口头报告

211. 如果开源调查结果呈交法院，调查人员可能须以证人身份作证，为此，须以口头证词形式介绍其调查。其他形式的口头报告可包括在真相委员会、非政府组织论坛、人民法庭或媒体活动上进行介绍。

212. 任何被要求口头介绍其开源调查结果的人都必须能够清楚准确地说明这项工作，包括所采用的方法和所使用的工具。这有助于确保口头证词和所述调查结果得到应有重视。

213. 在诉讼程序中，通常要求调查负责人作证，负责人应当能够介绍本团队的工作。当然，这要求他们了解团队做了什么，并且能够就团队承担的职责以及作出的调查范围、方法、所用工具等方面的任何决定背后的理由回答相关问题。调查人员可以是专家证人，也可以是普通证人。专家证人，即因其经验、知识、技能、所受培训、学历或相关证书而被视为专家的证人，可以就其得出的结论和其他分析工作成果作证。普通证人通常只限于就事实作证，具体而言，就他们亲自观察到的事实作证。

C. 可视化报告

214. 数据可视化是以图表、图示、表格、地图和信息图等形式对信息进行图形化表示，是查看和理解数据趋势、离群值和特征的便捷办法。[169] 它可以包括图表和其他时空数据的图形化表示、图示（包

[169] 不同背景下的可视化报告实例包括：在国际刑事法院的检察官诉艾哈迈德·法基·马赫迪案和黎巴嫩问题特别法庭检察官诉萨利姆·贾米勒·阿亚什等人案中利用数字平台作为展示证据；巴勒斯坦被占领土抗议活动独立国际调查委员会的详细调查结果报告（可查阅 www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session40/Documents/A_HRC_40_74_CRP2.pdf）；BBC Africa Eye, "Cameroon atrocity: what happened after Africa Eye found who killed this woman", BBC News, 30 May 2019（可查阅 www.bbc.com/news/av/world-africa-48432122/cameroon-atrocity-what-happened-after-africa-eye-found-who-killed-this-woman）。另见，"法证建筑"组织和"SITU 研究"开展的总体工作。

括展示数学联系、趋势或关系的图示)、展示不同人之间关系的网络图、统计图表和示意图。以时空维度显示物体的二维和三维地图，以及各种现场(包括犯罪现场)的三维重建，也属于数据可视化。[170] 这些工具有助于理解数量庞大的数据(常见于开源调查)，或更好地理解复杂的事实情况。

215. 其他数据可视化类型包括：

(a) 思维导图：思维导图是一种表示想法和概念以及它们之间关系的图形化手段。思维导图的信息组织方式使得信息更容易被分析、整合和理解。思维导图通常会说明基础数据是如何被发现的；

(b) 流程图：流程图是一系列事件(如嵌入算法、工作流程或类似流程的各步骤)的一种图形化表示；

(c) 信息图：信息图是想法或概念的一种图形化表示，可用来表示统计信息。

216. 开源信息的呈现方式多种多样，从单个视频或网站的视听展示，到交互式、数字化、聚合式的多媒体呈现，不一而足。[171] 可视化演示和说明，或数字平台都可以用来展示信息，使目标受众更容易理解基本事实。这方面的例子包括时间轴、合成照片(如犯罪现场的360度视图)和经过编辑的视频。

217. 在法庭上或在向其他公众展示数据可视化和多媒体证据时，调查人员应了解可能会出现哪些技术问题，包括了解律师可能需要哪些平台才能使证据展示对事实认定者尽可能有帮助。在决定基础数据的最佳表示形式时，应考虑一系列因素。这些因素包括目标受众、他们对潜在格式的适应程度以及他们理解所传达信息的能力。[172] 归根结底，所有呈现都应推进以有证明力且不带偏见的方式阐明与案件相关事实的目标，并应遵守信息呈现地司法管辖区的法律和道德要求。

[170] 例如见国际刑事法院数字平台：马里通布图(由“SITU研究”开发，作为国际刑事法院马赫迪案的一项资产)。可查阅 http://icc-mali.situplatform.com. 另见“法证建筑”组织网站上的各种在线开源调查及其可视化报告。可查阅 https://forensic-architecture.org/methodology/osint.

[171] 虽然不提供给法庭，但《纽约时报》可视化调查小组已经制作了一些说明性的可视化材料，旨在汇总在线开源信息、支持分析复杂事件并报告这些调查结果。例如见 Nicholas Casey, Christoph Koettl and Deborah Acosta, “Footage contradicts U.S. claim that Nicolás Maduro burned aid convoy”, *New York Times,* 10 March 2019 (可查阅 www.nytimes.com/2019/03/10/world/americas/venezuela-aid-fire-video.html)；Malachy Browne 等人，“10 minutes. 12 gunfire bursts. 30 videos. Mapping the Las Vegas massacre”, *New York Times,* 21 October 2017(可查阅 www.nytimes.com/video/us/100000005473328/las-vegas-shooting-timeline-12-bursts.html)。

[172] 见 Alexa Koenig, “Open source evidence and human rights cases: a modern social history”, in *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation and Accountability,* Sam Dubberley, Alexa Koenig and Daragh Murray, eds. (Oxford, Oxford University Press, 2020), pp. 38-40。

八

词 汇 表

摘 要

■ 开源调查中使用的术语和定义，或在相关或有关资源中可能出现的术语和定义。

218. 本章包含可能对开源调查人员有帮助的术语和定义。并非所有术语都在《规程》中用到，但由于它们可能在相关或有关资源中出现，因此予以列入。

气隙：数字设备不直接连接到互联网或任何网络，从而保障该设备所存储信息的安全。

算法：定义明确、使计算机能够解决问题或对预设场景作出反应的一个程序或一组指令。

匿名化：使特定个人的身份无法被识别的过程。

应用程序接口 (API)：使软件计算机程序可以相互通信的代码。

人工智能 (AI)：计算机科学的一个分支，致力于为机器开发程序，使其学习如何对未知变量作出反应和适应新环境。

信标：一种跟踪用户活动和行为的机制。信标由网页中的一个不起眼（通常不可见）的小元素（小到单个透明像素）构成，当浏览器显示信标时，信标将向第三方传送有关正在使用的浏览器和计算机的详细信息。

大数据：可通过分析来检测数据点之间的相关性并揭示可能有助于预测能力的特征的庞大数据集。大数据的关键特点是数量巨大和复杂性。

区块链技术：一种基于密码学的技术，通过这种技术，一个由“区块”组成的开放的分布式账本可以用来高效、可验证、永久地记录双方或实体之间的交易。

布尔搜索：一种互联网搜索技术，允许用户将关键词与运算符或修饰符（即 AND、NOT、OR）结合起来，缩小搜索结果范围，从而提供更相关、更具体的搜索结果。

Captcha：英文首字母缩写，即“全自动区分计算机和人类的图灵测试”，是一种在计算中使用的挑战—响应测试，用于确定用户是否是人类。

聊天室：一种可供用户进行实时在线对话的互联网网站。

云计算：一种能够通过内联网或互联网进行数据存储、处理和分析的运算模式。云有三种类型：私有、公有和混合。

Cookie：由网站发送并存储在用户计算机存储器中或写入计算机磁盘以供浏览器使用的一小段数据。Cookie 通常是网站高效运行所必需的，能够存储用户的网站偏好和详细身份信息，使用户在以后访问时不再需要不断输入数据。

加密签名：一种验证数字物项真实性的数学过程。使用算法可以生成两个在数学上相关联的密钥：一个私钥和一个公钥。为了创建数字签名，须使用软件创建电子数据的散列值。然后使用私钥对散列值进行加密。

密码技术：对信息进行数字编码或解码的做法。

暗网：互联网中只有通过特殊软件才能访问的部分，从而使用户和网站运营商可以保持匿名和不可追踪。

数据挖掘：从数据库中检验和提取数据以生成知识或新信息的做法。

数字档案馆：一种文档、网页或电子记录的集合。这个术语也可以指承担保全信息并将其提供给授权用户的责任的正式或非正式组织。

数字保全：管理和维护具有长期价值的数字信息、以便该数字信息在未来可供目标用户访问和使用的必要政策和战略。

域名：一种标识网络域的标签。在互联网中，域名是由域名系统 (DNS) 的规则和程序形成的。一般而言，一个域名代表一个互联网协议 (IP) 资源，如一台用于访问互联网的个人计算机、一个托管网站的服务器、网站本身或通过互联网通信的任何其他服务。

域名注册者：拥有或持有域名的个人、公司或其他实体。

域名系统 (DNS)：管理域名分配的系统。

搜索网：在网络语境下指一种广泛的自动收集或监视系统。

嵌入式数据：存储在源文件或网页中的数据。

加密：使数据在没有解密密钥的情况下无法访问的过程。

哈希或散列值：可以在任何类型的数字化文件上运行以生成一个可证明数字化文件未被修改的固定长度的字母数字字符串的计算。只要文件没有变动，该字符串在每次运行计算时都会保持不变。

超文本标记语言 (HTML)：一种编程语言，用于设计通过浏览器访问的网页。

超文本传输协议 (HTTP)：一种定义数据传输和接收方式的互联网基础协议。

互联网号码分配局 (IANA)：一个监督 IP 地址、自治系统编号和域名系统的全球分配的组织。

互联网名称与数字地址分配机构 (ICANN)：一个负责通过协调多个与互联网名称和数字空间有关的数据库的维护和流程来确保互联网稳定安全运行的组织。

互联网论坛 (也称为讨论板)：一种可供用户发布消息和进行对话的网站。论坛上的消息通常比聊天室中看到的消息长，并且更有可能对内容进行归档。

互联网协议 (IP) 地址：任何连入互联网的数字设备都有一个 IP 地址。IP 地址有两种类型：IPv4(32 位数字) 和 IPv6(128 位数字)。IP 地址可用于识别互联网上的计算机和其他设备。

互联网服务提供商 (ISP)：一种为互联网用户提供互联网访问和使用服务的实体。

内联网：一种使用互联网协议和网络连接建立内部版本互联网的专用计算机网络。

局域网 (LAN)：在一个限定物理地点连接到同一网络的数字设备的集合。

机器学习：一种人工智能，利用统计技术使计算机具备从数据中“学习”的能力，而不对其进行明确编程。

恶意软件：旨在对数字设备、网络、服务器或用户造成损害的具有险恶意图的软件。恶意软件有多种不同类型，包括病毒、特洛伊木马、勒索软件、广告软件和间谍软件。

元数据：关于数据的数据。它们包含有关电子文件的信息，这些信息嵌入在文件中或与文件相关联。元数据通常包括一个文件的特征和历史，如文件名称、大小以及创建和修改日期。元数据可以描述数字化文件是如何、何时以及由谁收集、创建、访问、修改和格式化的。

原生文件：原始格式的文件。

可携式文档格式 (PDF)：一种固定布局的文件格式，无论使用何种软件、硬件和操作系统打开和查看这类文档，文档格式 (包括字体、间距和图像) 均保留不变。将一个文件从原始格式转换为 PDF，会剥离元数据，生成文档的静态图像。

预测软件：使用预测算法和机器学习分析数据以预测未来或未知事件或行为的软件。

假名化：对个人数据进行处理，使得在不借助额外信息的情况下，无法再将信息归于特定的数据主体。

抓取：一种从网站提取大量数据的方法。

社会工程：为访问未经授权的信息而对人进行的心理操纵。它类似于黑客攻击，但涉及利用人的漏洞而不是技术漏洞。社会工程的类型多种多样，包括网络钓鱼和鱼叉式网络钓鱼。

剥离：一种从文件中删除元数据但不转换文件格式的技术过程。

结构化数据：存储库 (通常是数据库，但也可以是一组被填写的表单) 中遵循严格格式、以便随时处理和分析元素的数据或信息。

表层网：互联网中可通过任意浏览器访问并可使用传统搜索引擎进行搜索的部分。

追踪器：一种 cookie，利用浏览器的能力来记录访问过哪些网页、输入过哪些搜索条件等。追踪器通常是记录某一特定访问者行为日志的持久性 cookie。

流量数据：为在电子通信网上传递信息或为计算该项通信的费用而处理的任何数据。这类数据包括与通信线路、时间或时长有关的数据。

统一资源定位 (URL)：一个网页在互联网上的位置，即网址。

非结构化数据：形式多种多样、不按严格格式组织、因此不容易处理和分析的信息和数据。它们通常是文本，但也可以包括图像、音频和视频文件。

虚拟机：模拟计算机系统的软件。

虚拟专用网络 (VPN)：一种安全网络或安全节点系统，使用加密和其他安全过程确保只有授权用户才能访问网络。VPN 会隐藏 IP 地址，防止数据被拦截。

基于网络的服务提供商：在互联网上提供服务和产品的实体，如社交媒体公司。

网络爬虫 (也称为网络蜘蛛或蜘蛛机器人)：一种根据自动化脚本系统性地浏览互联网以下载访问过的网站并建立索引的程序。

WHOIS：一种可根据注册某一特定域名的实体识别该域名所有者的记录。开源调查人员可将 WHOIS 查找工具纳入来源分析和核实流程。

万维网 (WWW)：一种信息空间，使用 URL 标识其中的文档和其他网络资源，URL 可通过超文本相链接并可通过互联网进行访问。用户可以使用称为网络浏览器的软件应用程序访问万维网资源。

附　件

摘 要

- 在线调查计划模板
- 数字威胁和风险评估模板
- 数字格局评估模板
- 在线数据收集表
- 验证新工具的考虑因素

附件一

在线调查计划模板

调查编号：	
评估日期：	
调查概况：调查的主题事项以及地域和时间范围	

1. 目标和计划开展的活动

此处填写在线调查的目标和策略，以及具体活动和实施时间表。

2. 数据格局评估概况

此处对所调查地理区域数据格局进行评估，例如有哪些热门的社交媒体、移动应用程序和其他技术，以及谁能获取和使用这些技术。

3. 减轻风险战略和保护措施

此处填写数字威胁和风险评估的主要结论，以及识别、管理和应对此类威胁的战略。

4. 相关行为体的映射

此处列出名单，可能已经收集了后来消失的潜在相关在线内容的最初响应人员、数字档案馆以及互联网服务提供商和基于网络的服务提供商，这些行为体可能拥有在线内容的原始版本或额外元数据，可通过提出援助请求获取。虽然非法律调查人员可能没有法律权力要求提供非公开来源信息，但互联网服务提供商的内部联系人在解答问题和协助用户使用其平台方面可能发挥重要作用。

5. 作用和责任

此处确定团队成员的作用和责任，并应指定一名在线活动协调人。此处还可评估哪些人员可能要负责接受法庭传唤，出庭作证。

6. 资源

此处评估人员配置需求（调查员人数、工作人员的多样性和包容性），以及开展在线调查活动所需的任何专门培训和设备。

7. 文件工作

此处填写具体指示，说明团队成员应如何以及在何处记录其在线调查活动。

附件二

数字威胁和风险评估模板

调查编号：	
评估日期：	
调查概况：调查的主题事项以及地域和时间范围	
调查目标：	

1. 你有哪些资产？

人员（按性别分列）：

有形资产：

无形财产（如数据）：

2. 你有哪些弱点？

3. 哪些类型的威胁可以利用这些弱点并损害你的资产？

4. 谁是潜在的威胁行为体？

A. 他们对什么感兴趣？

B. 他们具备哪些能力？

C. 攻击的可能性有多大？

5. 哪些减轻风险措施是可能的 / 适当的？是否需要应对不同性别所面临的不同风险？

应考虑以下几方面：

- 物理伤害
- 数字伤害
- 心理社会伤害

附件三

数据格局评估模板	
调查编号：	
评估日期：	
调查概况：调查的主题事项以及地域和时间范围	
调查目标：	

星号 (*) 表示调查人员应考虑各种因素，例如年龄、性别、地点和其他相关人口统计信息。

1.	有关各方 (即特定社区、武装团体等)。说明各方在技术使用或在线代表性方面是否存在性别、年龄或残疾方面的差异。
2.	相关语言 (包括俚语和其他行话)*
3.	常用搜索引擎 *
4.	热门社交媒体平台 *
5.	热门网站 *
6.	互联网使用 / 普及情况 (按性别、年龄等分列)
7.	移动电话 / 操作系统偏好 (按性别、年龄等分列)
8.	热门移动应用程序 (按性别、年龄等分列)
9.	电信服务提供商
10.	连通性：Wi-Fi/ 手机信号塔位置
11.	相关法律 (表达自由、信息获取、隐私)
12.	媒体机构和记者 (在线存在)
13.	开放数据库 (例如政府数据、非政府组织 / 研究人员数据)
14.	付费数据库 (例如政府数据、私营公司 / 研究人员数据)
15.	在线内容的代表性 (包括与排除的群体)

附件四

在线数据收集表

1. 收集者信息

调查：

收集者：

收集者的 IP 地址：

收集开始 (日期戳 / 时间戳)：

收集结束 (日期戳 / 时间戳)：

2. 对象信息

网址 (URL)：

HTML 源代码：

截图：

捕获的数据：

IP 地址：

3. 收集包信息

收集包文件名：

收集包哈希列表：

收集包哈希列表文件的散列值：

4. 所使用的服务

软件产品：

授时服务：

IP 服务：

WHOIS 服务：

附件五

验证新工具的考虑因素

特征

- 开源代码还是闭源代码
- 付费还是免费
- 所有者（个人或公司）的身份、隶属关系或利益关系
- 资金（工具的资金来源和供资情况如何？产品的可能寿命是多长？）

安全问题

- 谁拥有工具或底层代码？
- 底层代码是开源还是闭源？
- 工具是否接受过独立审计？
- 收集的数据将存储在何处？
- 谁有权限访问任何收集的数据？
- 工具的安全基础架构是什么？
- 哪些法律义务可能影响工具使用的安全性？
- 如遇违法行为，是否有权获得补救？

运行问题

- 工具的功能是什么？
- 工具的可用性如何？
- 所有者、提供商或工具的用户支持能力如何？
- 工具多久更新一次？
- 工具与其他系统的兼容性如何？

HUMAN RIGHTS CENTER
UC Berkeley School of Law

University of California
Human Rights Center (HRC)
2224 Piedmont Avenue
Berkeley, CA 94720
Email: hrc@berkeley.edu
Website: https://humanrights.berkeley.edu/

联合国人权高级专员办事处

Office of the United Nations
High Commissioner for Human Rights
Palais des Nations
CH 1211 Geneva 10, Switzerland
Email: ohchr-infodesk@un.org
Website: www.ohchr.org/zh

本作品由联合国代表联合国人权事务高级专员办事处(人权高专办)
与加州大学伯克利分校法学院人权中心联合出版。

ISBN: 978-92-1-154248-6

Printed at United Nations, Geneva – 2012695 **(C)** – January 2024 – 830 – **HR/PUB/20/2**