智媒时代公民的身份确认与信息性隐私的保护

——基于情景化识别的视角

顾理平 俞立根[1]

【摘 要】：智媒时代，技术对人的识别日渐成为一种人机交互的前提，但随着技术方在识别能力上占据主动，作为权利主体的公民开始失去了互动中的平等和参与权利。在这一时代背景下，我国的立法实践通过明确隐私权和个人信息权益赋予了公民拒绝被识别和选择性被识别的权利。然而，正确的立法初衷却遇上了实践难题——对技术识别行为的辨析和举证困难，导致了隐私权利的无力感，并催促着赋权行为从现实情境中寻求解释。基于此，本文以“识别”为研究核心，尝试从人与技术的互动行为中归纳识别的类型，并基于不同情景的特质来理解信息性隐私保护中的识别认定困境。研究认为，识别关系是复合多样的，单以个人识别的身份视角为依据，难以回应智媒时代的隐私保护诉求，尤其平台间共享合作形成的“多对一”识别，和群体识别情景中平台“一对多”的识别形式已显示出挑战现有保护模式的可能。为了实现法律文本和客观实践的统一，本文提出一种“情景化识别”的方式，在全面理解技术识别行为的基础上，以“情景认定”和“识别检验”为核心，建立从“身份对应”到“收集行为与伤害性结果”的侵权认定方式。

【关键词】：情景化识别 信息性隐私 群体识别 隐私侵权

【中图分类号】G206．3**【文献标识码】**A**【文章编号】**1001－8263(2022)05－0099－12

自互联网时代开始，公民的隐私信息开始从完整变得碎片化。当碎片的数据可以经由技术整合形成用户画像时，隐私侵权就进入新一轮的失控状态。论及失控的原因，既因为碎片的信息是在用户许可或不知情的场景中被无感地收集，又因为这些碎片数据无法凭借个人的理性完成侵权指证，于是信息性隐私的侵权问题就不得不面对碎片信息的识别困境。具体而言，文中的信息性隐私指的是处在智能技术系统中具有个人隐私性质的信息，多以数据的形式存在于用户和平台的交互中，构成个人的私密信息。与传统意义上的隐私信息有所不同，信息性隐私以数据化的形式处于隐私和个人信息之间，既具有智媒时代个人信息的技术复杂性，又象征着公民的人格尊严并要求对其进行尊重和保障，因此理解技术逻辑是保护信息性隐私的先决条件。基于此，本文以信息的识别问题为研究对象，关注技术于细微之处识别用户的逻辑和方式，以及侵权认定方面，探讨用户如何举证平台的信息行为具备识别和针对性，以此解释信息性隐私中的技术行为和现实困境，最终探寻智媒时代公民隐私保护的微观路径。

随着《网络安全法》《民法典》《个人信息保护法》的相继实施，目前对个人信息识别的理解有了更清晰的方向。依照《民法典》第1034条第2款的定义：“个人信息是以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人的各种信息，包括自然人的姓名、出生日期、身份证件号码、生物识别信息、住址、电话号码、电子邮箱、健康信息、行踪信息等”。其中，能够单独或结合“识别”特定自然人成为“个人信息”的核心特征。联系我国《网络安全法》第76条第5项和《个人信息保护法》第4条的解释，并参照欧盟《通用数据保护条例》(General Data Protection Regulation, GDPR)第4条，美国《加利福尼亚州消费者隐私法案》(California Consumer Privacy Act, CCPA)第1798.140条。可以判断，虽然对“识别”的理解略有不同，但能否实现信息和个人的对应(即识别或关联)是世界范围内确认个体身份的共同标准。此时，信息化社会逐渐打通了虚

[1]**作者简介**：顾理平，南京师范大学新闻与传播学院教授、博导南京210024;俞立根，南京师范大学新闻与传播学院博士生南京210024

基金项目：国家社科基金重大项目“智媒时代公民隐私保护问题研究”(21&ZD324)的阶段性成果

拟和现实之间的边界，流动在虚拟世界中的零散数据经由“识别”与现实空间中的主体紧密相连，形成了肉身个体和数字化个体在“个人信息”层面的统一。另外，将《民法典》第 1034 条与 1038 条结合发现[1]，“识别”行为在法律意义上存在“单独识别”“结合识别”“无法识别”三种区分标准。其中，《民法典》将“单独识别”和“结合识别”的信息类型纳入到个人信息保护的范围内，而对“经加工无法识别”的信息设立豁免原则。于是，能否被识别既是“个人信息”成立的条件，也进一步影响到信息性隐私的判断。

一、传播行为中“识别”何以成为问题

从媒介技术介入民事行为的纵向演变来看，早期借助语言的人际交往中，双方即使叫不出对方的姓名，也能确定侵权者和受害者是谁，形成默认的指认状态。到了明显的中介化交流时期，报刊的风靡使通信脱离了个体的身体并开始在交往中占有一席之地时，受害者需要从特定的新闻话语中确认流言蜚语指向的对象就是自己，形成直接辨认的状态。进入信息化社会，多元的符号不满足于脱离这样的附加形式，而是成为形成“我们是谁”的组成部分[2]，于是，“识别”行为不需要从信息洪流中找出完整的“我”，任何“零碎”和“流动”的数据都与“我们是谁”有关，一种模糊的间接识别方式成为主流。

(一)传统民事行为中的默认识别

以语言为媒介的人际交往被视为文明社会的传统形态，这种民事行为中当然可能出现侵权行为，为了维持人际交往的公平与正义，传统的民事行为找到了一条“矫正主义”侵权救济道路，从中我们可以理解早期民事侵权中隐含的识别逻辑。在亚里士多德的伦理学中，“矫正”一词内在地含有某种不公正的行为，以及对不公正行为的恢复。当人与人之间存在某种破坏算术比例的不公正行为时，一种追求公平的思想就会在侵权者和受害者中寻求算数上的平衡。[3]在此，亚里士多德并没有在侵权者和受害者的辨认上诉诸笔墨，而是关注特定的关系行为。科尔曼(Jules L. Coleman)意识到了这一点，他对行为矫正观念的解读发掘了“矫正主义”之中的“关联性”，即矫正正义的两端联结的是特定指向的个人，且仅适用于双方之间的关系。[4]可见，矫正虽然针对的是不公正的行为，但行为指向的责任人和受害者应当是清晰可见的。我国的民法理论有着德国法的渊源，与“矫正主义”的理念具有连贯性的是，加害行为所侵害的对象虽然是“民事权益”，但“民事权益”背后的受害人同样可以清晰指认。[5]

由是可以说，在侵权法的理念中，确定且具体的侵权人与受害人是一种默认的条件。正如伯克斯(Peter Birks)所言，民事不法行为的概念并不内在地要求有受害人，但“原告必定以某种方式遭受了不利的影响，法律认为该方式足以使其被确认为义务的违反的受害人，并赋予他以为了自己的利益起诉的资格”。[4]因此，“识别”并不是民事行为的一般问题，尤其是在“礼俗社会”中，涉及隐私信息的口头流传多以关系网络中熟知的信息代指受害者。由于每个人的过往信息皆是社会中共同知晓的内容，即使不出现人名，仅是流言蜚语中的特征都具有清晰的指认性，都能与特定的人对号入座。

(二)传统媒体时代新闻传播行为中的直接识别

大众媒介的兴起带来了一种经由报纸扩散，因而带来了影响和伤害更大的侵犯方式。当沃伦(Samuel Warren)家族的私生活细节径直地出现在《星期六晚报》(The Saturday Evening Gazette)上，并成为大众消费的对象时，为此恼怒的沃伦联手布兰代斯(Louis Brandeis)发起了对新闻媒体侵犯隐私权的声讨[6]，由此写就了隐私权的开山之作《隐私权》(The Right to Privacy)。[7]与普通民事侵权行为所不同的是，沃伦确认自己作为受害人的依据不再是侵权行为所直接指向的肉身“我”，而是姓名、地址、身份等信息要素的集合，这些信息成为“我”的意义和指代。基于新闻传播在社会进步中的重要意义，新闻媒体也获得了一种豁免的可能——公开的信息不可指认就不构成人格侵权。如多年前《北京青年报》的记者安顿曾在“口述实录”专栏中公开了大量涉及情感隐私的个案调查，但由于可指认的信息被隐去，她记录私人情感隐私的行为就获得了豁免。

鉴于新闻侵权具有这样的特殊性，早期新闻传播法学的学者们对新闻报道中的指认问题投以关注。在大陆较早出版的两本新闻传播法学类书籍中可以寻见，“受害人”是否可以被“指认”或“辨别”，在新闻侵权中占有独特的位置。《新闻法学》中，

笔者认为“与一般的民事侵权相比较，新闻侵害隐私权具有自身的特征，不宜简单照搬民事侵权的构成要件”,转而提出“只有发表的新闻作品中提及隐私并使其当事人可以被指认时，才算真正构成了新闻侵害隐私权”。[8]《新闻传播法教程》中，魏永征在侵害隐私权的排除类型中，将“使不可辨认”作为一种可行的抗辩事由。[9]

如此可知，在以新闻出版物为代表的传播行为中，个人的身份信息通常会散落于媒介的内容之中。虽然零碎的信息要素增加了识别的难度，但借助特定的时空关系，我们仍然可以将姓名、身份、地址等信息要素对应到特定的个人。因此，此时的侵权行为虽然遇上了新闻报道带来的挑战，但通过能否被“直接指认”缓冲了侵权的困境。

(三)新媒体时代信息传播行为中的结合识别

大众化报刊时代之后，信息技术以更快的速度推动着社会生活的大变革，置身其中的个人身份也呈现出更丰富和复杂的面向。一方面，信息技术将个人身份进一步打碎，“我是谁”在不同的情境中不断变化着，“自我的一贯性和连续性也因此变得要根据背景灵活地构建了，自我的稳定性不再与实质性的身份定义有关系”。[10]另一方面，信息技术在主体之外培育着另一个可操控的主体形式，这种由数字痕迹绘制、凝结成的新主体隐含了更多深层的个人特质，“它既可以让我们依附于无形的数字网络，失去内在的灵魂，也可以让我们在离散的剩余数据中发现异质性的我，一个流溢的我”。[11]这就意味着，一个拥有数字化身份的“我”正式形成。如此，个人隐私信息的可识别要素变得超出理性人的认知范畴，由此而来的不确定性被极大地增强。

对此问题的回应形成了两种实践路径。美国偏重商业发展的取向在早期形成了以个人信息等同于“可(直接)识别个人身份的信息”为核心的“个人可识别信息”(personally identifiable information)标准，简称“PII”,而欧盟则在注重人格保护的取向下吸纳了个人信息“可间接识别”的内涵，二者的区别可以理解为欧盟在 PII 标准之上，保护利用碎片信息识别公民身份的行为。针对这一分歧，施瓦茨(Paul M. Schwartz)和沙勒夫(Daniel J. Solove)在美国学术界提出了个人可识别信息的第二代版本，简称“PII2.0”模式。依据风险程度，将直接识别信息(identified data)、结合识别信息(identifiable data)和不可识别信息(non-identified data)视为连续的体系[12],对于可结合识别的信息提出透明度、数据安全和数据质量(准确、相关)等要求，这也成为国内学者们的解释论基础。[13]

正因如此，我国关于隐私信息保护的法案逐渐认同了“PII2.0”模式的连续体系，造成了《民法典》中的“单独识别”“结合识别”“无法识别”的思路。然而，由于“结合识别”和“无法识别”的概念边界十分模糊，形成了规制层面的一系列难题。其一，面对技术黑箱，个人如何举证“结合识别”的过程和结果具有针对性?其二，当技术行为不透明时，个人如何辨析平台方“无法识别”的抗辩话语?这两重关于识别的认识难题指向了进一步的探讨：一方面，在持续且流动的数据面前“无法识别”(匿名)是否存在或可能?另一方面，由于“无法识别”为侵权纠纷行为提供了豁免，那么需要追问，技术平台不采取识别行为但却对特定个人造成实际伤害的情形是否可以诉诸信息性隐私侵权?综上而言，媒介的居中调节对于“识别”具有很大的影响，而技术的催化剂作用也让“识别”逐渐成为问题。因此，媒介技术成为本文的核心视角，支撑着对所提出问题的理解。

二、识别的情景：智媒平台中的信息识别类型

对以上问题的回应需要回到具体的识别行为中，在智媒生活的场景里，五花八门的信息处理方式虽然都被归入识别范畴之下，但具体的行为模式需要深入讨论，即需要区分识别情景的思路，承认识别行为的内部具有差异性，分析用户如何使用平台提供的技术，平台技术又如何分析用户，进而在人与技术方的互动中发现具有侵犯性的识别行为。

(一)基于“收集—存储”的直接识别情景

在用户知情且授权的条件下，通过将及时提取到的个人敏感信息与事前采集的身份信息进行比对是最常规的身份识别方式。大至机场、银行等公共场所，小到门锁、手机等私人物件，“收集—存储”模式都能让用户体验到识别带来的便捷、高效、稳定。

然而，“收集—存储”背后的行为逻辑包含了对身份特征的提取和对提取信息的建模存储，由此带来两类和现实相关且具有法律风险的识别情景，这两种情景也正成为技术发展和应用的趋势。

其一为泛在识别情景。泛在识别依托无处不在的智能装置，并以非侵入的方式消解用户对数据收集行为的感知和排斥。在卡彭特(Timothy Carpenter)诉美国政府侵犯隐私案中，运营商为政府提供了卡彭特 127 天内的共 12898 个位置信息，每当卡彭特的手机连接到蜂窝站点时，它都会生成一个带有时间戳的记录，称为蜂窝站点位置信息(CSLI)。[14]结合万物互联的技术背景会发现，曾经习以为常的手机摇身一变成为了配备智能装置的技术产物，并源源不断地将个人数据向外传递。如斯坦福团队的智能马桶、亚马逊的智能音箱、特斯拉自动驾驶的汽车、威斯敏斯特的智能路灯，诸如此类的智能设备连同贴身的手机、可穿戴设备已经形成了跨越时空的数据采集系统，其中不乏进行身份验证的直接识别系统。美国白宫前首席信息官特蕾莎·M.佩顿告诫我们：“家庭越来越多地集成了智能系统和智能电网”。“当你由各种应用程序控制的互联网家居改善了你的生活，它也向营销公司、不法分子甚至政府的密探广播关于你的数据”。[15]

其二为生物识别情景。智能生物识别技术日渐变成一种普遍的身份识别技术，成为开启数字化生存之门的钥匙，但从隐私保护的视角来看，生物特征亦是公民的核心隐私。[16]由此展开的应用带来了更多福祸相依的可能性，仅针对“强生物特征”[17]而言，其独特的识别性一旦被采集且存储成功就会有近乎永久的效果 [18],使信息性隐私侵权面临难以补救的风险。然而，如此重要的生物特征又容易在不易感知的情景中被采集。日前，在美国伊利诺伊州就发生了一起餐饮场景中私自采集用户声纹信息的侵权案件，被告麦当劳公司为了简化餐厅的运营，使用人工智能语音助手与原告互动并采集声纹信息，原告卡彭特(Shannon Carpenter)在事后觉察不对，以违反州生物信息隐私法向法院提起千人的集体诉讼。[19]在我国涉及生物特征的案件也开始逐渐增多，如“郭兵诉杭州野生动物园”和湖南“林某某诉中欣物业管理”两起案子都指向了“人脸识别”行为对生物特征信息的不当应用。与此同时，生物信息的交易乱象更加昭示着此情景的现实迫切性。新华网的调查文章《0.5 元一份!谁在出卖我们的人脸信息?》揭示了个人信息批量倒卖的黑产，人脸数据经过“照片活化”将会关系到个人的生命财产安全。[20]

最近，在 ITIF(Information Technology & Innovation Foundation)关于 AR、VR 和隐私的报告中，一种通过 AR/VR 设备大量收集生物特征数据，进而识别个体并推断其他信息的技术形式正在形成。[21]可以说，现实生活中的泛在识别与生物识别正在趋于一致，二者共同作用于对个体的直接识别。

(二)基于“处理—加工”的整合识别情景

整合识别情景常见于通过技术的处理、加工能力将收集到的零散用户数据绘制成可以识别的数字画像，这种画像与用户形成一一对应关系，并且以数字传记的形式持续更新。在《数字人》(The Digital Person)中，沙勒夫(Daniel J. Solove)指出了网络空间的“整合效应”:“个人数据被整合在一起，创造了一个关于我们的数字传记。”[22]这种基于“处理—加工”的识别模式可以从细微、零散的数据中提升辨别个人的能力，如“庞某某诉东航和趣拿信息技术公司隐私权纠纷案”中，庞某某通过他人在“去哪儿网”订购机票，在未预留自己手机号码的情况下收到手机诈骗短信，庞某某以隐私权被侵害为由将“去哪儿网”公司和东航诉至法院。对此，法院认定：“即使单纯的庞理鹏的姓名和手机号不构成隐私信息，但当姓名、手机号和庞理鹏的行程信息(隐私信息)结合在一起时，结合之后的整体信息也因包含了隐私信息而整体上成为隐私信息。”[23]如今，整合识别在数据时代扶摇直上，“整合型隐私作为一种新类型隐私，伴随大数据时代的到来而产生，并将长期存在下去”,给公民的个人生活带来重大影响。[24]

另一类整合识别的情景借助对关联信息的分析推测用户身份。如“凌某某诉抖音”案中，针对原告未授权自己的通讯录信息却被推荐通讯好友的行为，被告辩称是因为其他抖音用户授权访问的通讯录中有原告的姓名和号码。[25]类似的情况也发生在 Facebook 的好友推荐中,Facebook 的发言人在电子邮件中说：“‘你可能认识的人’的评选取决于多种因素,包括共同的朋友、工作和教育信息、你所处的社交网络、你引入的联系方式以及许多其他因素。”[26]于是乎，一种新型的针对社交关系的处理方式正在弥漫开来，让用户逐渐失去了对隐私信息的控制。

由此可见，整合识别情景中受害人所面临的最大挑战是如何确认自己被识别，以及被谁识别。一方面，算法黑箱的不透明性使得识别的过程和结果模糊不清；另一方面，当以一种开放的视角看待增长的数据和发展的算法时，技术具有了朝向未来的识别可能。

(三)基于“提供—公开”的扩散识别情景

在扩散识别情景中，平台对用户的识别能力已经从单一场景放大至整个信息系统。此时公民的数据虽然以合目的、合场景的方式被收集，但由于整个系统处于联通状态，当中介平台或多个行为主体尝试共谋时，公民的数据形态就会如同拼贴中的马赛克网络，逐步细化为可识别的图像。

一类扩散识别情景常见于具有纵向合作关系的平台之间，由主导方向依附方提供用户信息。如 Brightest Flashlight Free 将收集到的信息提供给第三方广告投放的商家，以此寻求获利的行为。在研究关联平台之间的信息流动状况时我们发现，大型平台普遍会在隐私政策中声明与关联方(The affiliated party)平台进行信息共享，另有部分平台会与外部第三方共享。[27]如此，使得用户可识别的信息场景发生迁移，从而扩大了被识别的范围。

另一类扩散识别情景形成于较平等的、多方合作下的信息共享。扎拉奇(Ariel Ezrachi)发现，平台间正合力追踪和解读我们的个人信息：“在提取用户数据的阶段，互联网企业之间精诚合作，致力于用户数据追踪、数据库的打造与行为定向广告投放”。[28] 多方在合作中能够拥有多个用户浏览习惯的观看视角 [29],并且可以通过戮力合作将各自领域的信息汇集到更大的数据库中，完美地绕开“合目的收集使用”的信息规制原则。

(四)基于“重组—分类”的群体识别情景

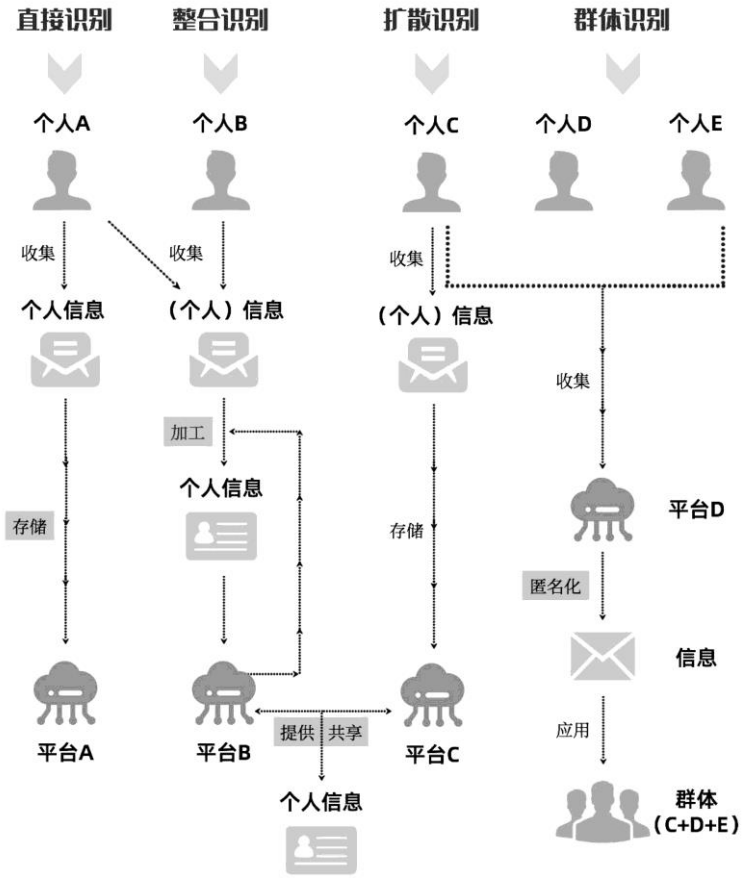
在对用户“行为歧视”的路上，无论是数据掮客还是平台算法正在尝试将有“共同之处”的用户分为一组，并借助分组后的模型来类推其他可能具有此特征的用户 [30],如此演绎着基于“重组—分类”的群体识别情景。在信息技术初出茅庐的时候，甘地(Oscar H. Gandy, Jr)就提出，社会正在兴起一种“全景分类”(Panoptic Sort)的监视方式，大量从官方问卷、市场调查收集来的个人信息正被分入社会阶级、市场经历、消费习惯等分组中，这些经过分类的信息正用于评估公民的消费行为和态度，最终形成一种反民主的控制体系。[31]

随着数据技术日渐炉火纯青，“群体不再基于某些观察者的感知进行分类，而是通过不透明的算法过程进行分类”,它往往会以模糊分类的方式，根据人们的行为倾向来“观察”他们，而不是把他们作为个体。在 50 厘米的分辨率下，没有人能拥有个人身份，但却能感受到真切的影响。[32]具体的行为逻辑为，跳出识别个人的思维惯性，在现有数据基础上直指“活跃”群体，依据设定的参数或新的分析方法“发现新群体”,最后针对“新群体”确定分类并应用。依照群体分类的方式，匿名成为了次要问题，个人身份在该情景中并未暴露，但针对该个体的行为已经产生。此外，群体识别方式并不只停留在分类阶段，如若想要识别个体，可以通过不断地添加组的属性来完成，每当新的属性被添加，组的分类范围就会变小，直至最后可以确定个人。[32]

如今群体分类识别的方式已成主流。不久前，字节跳动旗下的数据分析公司发布了一份《十大新消费人群》的报告，如其所言，“后续完整版报告将通过捕捉新消费人群的消费趋势以及潜在的增长机遇，进行深入分析。面对这些不断崛起的新消费态度与消费行为，对企业来说，其实细分领域机会多多……”[33]。令人触目惊心的是，这仅仅是分类分析的开始。

综上所述，如“智媒平台识别情景中的技术形式图”所示(见图 1),不同情景的区别在于识别方式的变化。直接识别情景是对个人私密信息的“存储”,可单独识别。整合识别情景则将碎片的信息置于循环的“加工”中，生成可识别的个人信息。对于扩散识别情景，平台与平台间的信息“共享”成为关键，即使经过“去标识化”的碎片信息也能够在交换中重新具备了识别性。最具颠覆性的情景出现在群体识别行为中，虽然经过“去识别化”(即“去标识化”或“匿名化”),但对个人特征的应用具有超

越个人隐私权益的破坏性。当然，智媒平台与用户的互动可能仅涉及部分情景，也可能涉及全部情景。对情景的类型化区分是为了从技术与人互动的视角理解识别的具体形式，并析出其中隐秘的关键步骤，这些步骤将是侵权认定的焦点和难点。



智媒平台识别情景中的技术形式图

三、识别的认定困境：从难以匿名到无须匿名

通过对识别情景的区分可以发现，“识别”在形式上具有不同的认定取向。一种以个体身份的确认为标准，能够将以标识符号为主的零散数据对应到特定的个人，如个体识别情景和整合识别情景，即使进行了去标识化处理，但借助额外信息和重标识方法进行关联和推断，仍然具有再识别的可能，如部分整合识别情景和扩散识别情景；另一种为绕开对个体身份的认证，转而以相似的群体行为设立组属性，实现对用户的区分，如群体识别情景。这两种识别方式虽然取向殊异，但都可能使特定个体因为被凸显出来而受到伤害。当然，由于现阶段对识别的理解指向了身份的确证，使得“无法识别”成为法律结构中的豁免条款，由此形成了新的认定困境。

(一)身份视角下的个体识别与匿名条件的实现困境

身份视角的认定方式建立在“确定”作为核心特征的基础上[34],作为“识别”的前身，指认意在强调个人和身份的直接对应。如今，大多数国家均以身份视角建立了保护个人信息的法律，如欧盟《通用数据保护条例》第 4 条，日本《个人信息保护法》第 2 条，《德国联邦数据保护法》(2015 修订版)第 3 条等。我国现行的个人信息保护框架中，虽然都以“识别”作为构成个人信息的标准，但在判定时引入了“关联”概念。参照《个人信息安全规范》的国家标准 GB/T35273-2020,某项信息是否属于个人信息

的路径不仅是对应的识别，还可以依据“如已知特定自然人，由该特定自然人在其活动中产生的信息(如个人位置信息、个人通话记录、个人浏览记录等)即为个人信息”。如“黄某诉腾讯隐私权、个人信息权益网络侵权责任纠纷案”中，法院引用 2017 年颁布的《个人信息安全规范》(GB/T35273-2017)从“识别”和“关联”两个路径对微信好友关系、读书信息等是否属于个人信息和隐私进行认定。[35]与个人信息保护制度中的“识别”略有不同，“关联”之所以可以成立于个人身份识别中，关键在于已确定的个人身份推及的相关信息，与之相关的信息均在明确的个人之下。因此，无论是由信息推至个人，还是由个人遮盖下的信息，都是建立在对特定个体的对应和认定基础上的，最终目的都是要有一个清晰的个体身份出现。

身份视角下的个体识别方式清晰易行。从逻辑上看，直接识别、间接识别到无法识别的连续体系具备完整性，然而这种连贯仅仅满足形式上的合理。以去识别信息的社会实践为例，前有美国在线网站(AOL)第 4417749 号匿名搜索者被准确识别[36],后有《纽约时报》隐私项目组在规模巨大的手机数据中准确地还原出用户本人及其具体行为[37],去识别和识别就好像既对立又统一，相互缠绕又无法各自剥离。因而，无法识别或称“匿名”面临一个根本性的实现困境。对此困境，《个人信息保护法》《个人信息安全规范》《信息安全技术个人信息去标识化指南》等共同将“无法识别”确定为有条件限制的概念，其一，通过“匿名”和“匿名化”的调整使“无法识别”从一个结果状态到过程状态；其二，通过“去标识化”和“匿名化”的区分，实现无法识别的分层，对于“去标识化”接纳其借助额外信息进行重识别的风险。然而，有目共睹的是，“个人信息去识别化之后，无论是去标识化还是匿名化，仍可以通过再识别化技术加以复原，都具有个人信息可识别性的本质特征，仍属于个人信息的范畴”。[38]因此，对于“无法识别”的理解需要充分考虑其实现的条件性，并依据识别情景中的技术特性来厘定。

当“信息”在流动的数据社会中无法实现绝对“匿名”时，识别的连续体系就只能在特定的静态场景中实现，所以可以看见“情景脉络完整性”或称“场景原则”在学术论文中成为寻求解决方案中的热点。[39]于是，身份视角下的识别正在变的无所适从，一方面因为经由碎片信息识别个人身份的过程不易阐明，所以形成关于“可识别”与“无法识别”的争议；另一方面，从群体识别情景中可以发现，身份视角下的识别不再是必经之路，平台可以绕过识别来直接应用，形成对匿名条款的降维打击。

(二)行为视角下的群体识别与匿名条件的存在困境

平台方为何要识别用户?按照身份视角的逻辑，平台方因为需要了解用户而采取识别行为，通常来看，一旦可以实现用户和信息的对应就可以将用户的其他信息和行为尽收眼底。如此，对应式的识别仅是实现管理和认知的惯用方式，其逻辑和原型可以从福柯所指的利用社会编号进行的规训及调节生命的技术中找到，这是前数据社会中最高效的应用方式。与身份视角不同的逻辑出现在对行为的识别中，由于持续进行数据监测的可能，平台通过对特征的连续把握就可以直接了解用户而不需要知道具体是谁，曾经作为中介的个体身份就渐渐淡出了整个流程。如今，在智媒平台的商业化应用中，通过身份识别建立的关系是不经济的，除了技术成本、运营成本外，日渐收紧的隐私保护法律也增加了平台的违规成本。于是依托于数据基础和算法技术的优势，平台仅通过对重要特征和特征群体的创建，就能取得规模性效益。然而，从用户的角度来看，这种群体识别的应用方式，不仅会运用到带有私密性质的个人敏感特征，以“推论隐私”(inferential privacy)[40]的方式实现预测，还会得到不透明又不接受质疑的结果，从而给消费者带来不公正和种种危害。

行为视角下，平台方与用户之间是一种互相生成的动态关系，以用户的行为偏好、特征、趋势等源源不断的个人数据为中介，平台方和用户形成稳固的合作关系。此时，由于数据是持续流入的，识别用户的身份就不再是了解的最优方式。2018 年颁行的美国加州《消费者隐私法案》(CCPA)注意到了消费者信息的这一特征，其中提出了“整合消费者信息”(Aggregate consumer information)一词，用于指代与消费者群体或类别相关的信息。[41]参照 CCPA 的相关定义可以获知，由于“整合消费者信息”中的个人身份已被删除，所以被排除在“个人信息”之外，然而“整合消费者信息”又并非是无法识别的单一或多个个体消费者的记录，因此，“整合消费者信息”和“匿名信息”在“个人信息”定义中处于并列位置。

那么，从行为特征建立的群体分类方式为何可以认定为识别?一方面，如群体识别情景中所示，当个人信息汇聚成仅保留单一或少数特征的群组时，虽然个体被淹没于群体之中，但实际的受力者是特定的个人，形成了“个体—(无数)对象—个体”的应

用模式。另一方面，即使收集到的个人信息经过了去身份的流程，即在(无数)对象阶段的个体确实处于不可识别的状态，然而行为视角的动态关系特征决定了，当新的数据流入时，可以实现对特定对象的再识别。

综上所述，信息识别的困境面临“难以匿名”和“无须匿名”的双重困境，尤其在行为视角下，匿名与否对于信息的利用已经影响甚微了。数据生态社会才刚刚开始，对于刚触及数字技术的我们来说，搭建起适合数字社会的保护体系并非是一蹴而就的。若是将《民法典》和《个人信息保护法》视作一个发展过程来看，其中对匿名的理解就发生了变化。《民法典》第 1038 条中的匿名为无法识别且不能复原的信息结果。《个人信息保护法》(草案)则有不同，一审草案第 24 条第 2 款指出：“个人信息处理者向第三方提供匿名化信息的，第三方不得利用技术等手段重新识别个人身份。”立法者显然注意到了匿名化信息可被重新识别的可能，以及现实社会中存在的实践行为。虽然二审稿出于对“匿名”词义的忠实去掉了该款，但匿名在《个人信息保护法》(草案)中偏重了“匿名化”的过程(详见草案对“匿名化”的定义)。最终，《个人信息保护法》采用了“匿名化处理后的信息”这一调和性的说法。因此，对识别的认定应当在“直接识别—间接识别—无法识别”的连贯形式中保有实际的开放性。

四、情景化路径：侵权纠纷中的行为认定方式

既然要保持身份识别标准在实际应用中的开放性，就需要回到具体的行为方式中去，本文提倡以情景化识别的方式作为对现有识别标准的补充。虽然以“情景”命名有罗列类型之嫌，但秉持的是从技术与人的互动关系出发形成关于识别的复合认知，实际想要呈现的也是平台与用户间“一对一”“多对一”“一对多”的识别面向。基于以上对技术识别行为的理解，可以尝试一种认定法律责任的情景化视角。

具体而言，情景化识别可以视为一种依据实际情景类型来认定侵权纠纷的方式。首先确定侵权纠纷的双方是普通民事法律关系中的个人对个人，个人对公权力机构，个人对私权力平台；其次，依照纠纷行为中的技术性因素划分“传统隐私权、个人信息权益纠纷”与“信息技术主导的隐私权、个人信息权益纠纷”。对后者需厘清纠纷涉及的信息是个人直接识别情景(直接识别)、整合识别情景(间接识别)、扩散识别情景(多方参与的结合识别)还是群体识别情景(行为特征)。最后依照不同情景中的行为方式对识别进行判定，通过技术情景厘清双方责任，依据如下。

针对个人识别情景应以“身份对应”为判断中心，主要关注是否具有可以直接识别用户身份的行为，如个人基本信息(姓名、电话、住址),虚拟信息(账号、IP 地址、唯一设备识别码),以及由已知自然人生成的各类关联信息(浏览记录、行踪记录、社交关系)。属于此情景中的信息处理行为能够较为简单、清晰地还原到用户的个人，且无太大争议。对此，被告方多可能以“知情同意”作为抗辩事由。

整合识别情景中的用户信息趋于碎片化，识别成立与否的关键在于碎片信息结合的结果是否具有个人的指向性。此时，被告方多以“匿名化”为抗辩事由，因而对“匿名”的判定成为关键，提倡以“技术评估”为核心方式。首先是操作层面，是否匿名可以通过技术调查官进行技术上的确定。参考《最高人民检察院关于指派、聘请有专门知识的人参与办案若干问题的规定(试行)》第三条规定，涉及专业性较强的案件，人民检察院可以指派、聘请“有专门知识的人”参与办案。[42] 在“凌某某诉抖音隐私权、个人信息权益侵权纠纷案”中，北京互联网法院就曾指派技术调查官参与诉讼，技术调查官对案件中的 Cookies 记录、同源策略、IP 地址的精确度等技术问题发表意见。[43] 其次是原则层面，技术鉴定会以“就本案所涉及的技术问题”给出意见，其中案件具体的情节所分析的是该平台的识别行为。一者，不同平台有不同的技术识别能力；二者，匿名化是一个过程，没有绝对的匿名，只有相对场景的匿名，如此会产生平台匿名能力和社会匿名标准之间的差异。鉴于信息系统时刻处于联通状态，各平台也有共享行为，有必要从原则层面树立社会整体标准为主、平台具体能力为辅的识别认定原则。

对于扩散识别情景，由于会涉及多方的信息识别行为，需要厘清多方间的关系和侵权纠纷行为中的信息扩散路径，所以将“关系与目的核查”视为目标。以“黄某诉腾讯案”为例，案件涉及“微信读书”通过“微信”获取用户信息的行为。原告腾讯公司以关联方内部共享为抗辩理由，意在凸显所共享的信息属于腾讯公司已经取得控制权的信息，但一审法院认定微信读书与

微信在应用软件市场为独立的两款应用软件。对于被告提出共享 OPEN_ID 等信息无法知道对应的用户身份的理由，一审法院以“均由腾讯计算机公司运营”，“获取 OPEN_ID 即可以识别用户身份”对此否定。[44]厘清多方关系的同时，出于利用知情同意的抗辩行为，或对于难以认定的识别行为，可以考虑回避追究是否匿名，转而对扩散行为的合法性、正当性和必要性进行考量，以此判断扩散行为的目的和基础是否合法。

最困难的情景当属群体识别，群体识别情景中最大的挑战在于对“匿名”的悬置，因此，需要结合“收集行为与伤害性结果”来共同判定。由于从“个体—(无数)对象—个体”的过程中确实可以做到难以识别或无法举证是否可以识别，于是对此情景可以从行为方式和结果上加以判定。首先，确定该平台针对受害人有切实的信息收集行为，即用户有过知情使用；其次，通过与使用有因果关系的伤害性结果来检验是否具有违法行为。若有信息收集，又有违法的伤害性结果，如大数据杀熟造成的财产和精神损失，即可绕过识别认定直接诉诸信息侵权。采用如此方式需要回答一个问题：对“识别”的悬置是否可以诉诸个人信息侵权?依照《个人信息保护法》的立法初衷(参见总则第 1、2 条),保护个人信息权益不受侵犯是重要原则之一。当平台收集了个人信息，经过匿名化处理，最终作用于群体，却导致信息被收集的个人受到实际的伤害。可以认为，该信息处理方式不合《个人信息保护法》的原则，以欺诈、误导等方式处理个人信息(参见第 5 条),且侵害了个人的人格权益。

五、结语

智媒时代，技术识别是开启数字化生存之门的钥匙，这把钥匙是用个人敏感信息打造，并且通向无限开放的私密空间，如此，钥匙就不应当是各个平台随意可用的，而对于钥匙的保护更是关系着权利的保护。基于这样的理解，本文重点关注个人信息中的识别问题，其中，对于识别的立场是从以下三个层面展开的，这也是思索此问题的初衷和开启智能时代隐私研究的立足点。

其一，在智能化的社会中，识别可以视作人与技术交互的根本方式。正如人和人之间的交流，技术对人的识别与人对技术的理解日渐成为智能时代的生存方式——人与机器的共存共生。对隐私信息侵权的强调并非是对技术识别人的行为和指令的反对，而是以更好的共生方式为出发点，对违背个人意志且伤害人格尊严的识别行为进行抵制，意在化解人与技术、人与技术背后的人之间的紧张关系，维护社会正义秩序的发展。

其二，智能时代，暗流与明流涌动的数据仿佛给公民的隐私信息加入了时间的向度——无限记录的过去和无限开放的未来。若以一种动态的视角在数据流中寻找，任何特定的个体都会有迹可循，进而显露出自己可识别的数字身份，这同样是技术带来的生存方式。如此，对于“无法识别”的理解不应当是绝对的，因为没有一个清晰的标准可以成为区分识别的边界，但“识别”和“匿名”是相辅相成的，这又决定了“无法识别”不能被否定，因为失去了“无法识别”的参照，识别就在实践中失去了方向。如此，划分识别的情景并用一种情景化的思路去理解就显得格外重要。

其三，基于对时代发展和具体实践的调和，本文提出情景化识别的视角，将识别到匿名的过程还原到具有特定条件的框架中。对情景的分类不仅仅是让识别的行为类型能够显现出来，更是将高维的识别方式拆解到平面中，这样可以找到识别与匿名的边界。情景化识别可以视作从媒介调节方式入手的一种认定思路，将法律与代码的逻辑思路进行类比，当“If”出现某种技术与人的识别方式，“Then”可归属于某种识别情景，最后依此作出判定。

综上所述，识别在智媒时代的语境中的重要性日渐彰显，因为我们生活在一个可以不被人识别，但不能也无法不被机器所识别的社会。正因如此，我们对“识别”的理解应当置于媒介技术与人的互动关系中，跳出“一对一”的身份确认，形成“多对一”和“一对多”的多维认识。可以预见的趋势是，随着技术和人之间的关系发展，还会有新的情景涌现出来，而技术对个人身份的辨别正在让位于对动态特征的把握，如此传统的识别认知只是一种可能但不必要的技术形式，识别的过程隐而不见(不识别)也可以有侵害。对此，情景化识别尝试建立的是以尊严感而非识别结果为根基的保护形式。

注释：

1《民法典》第1038条提出了个人信息识别标准的豁免条款，“未经自然人同意，不得向他人非法提供其个人信息，但是经过加工无法识别特定个人且不能复原的除外。”

2 Koopman,C.,How We Become Our Data:A Genealogy of the Informational Person,Chicago:The University of Chicago Press,2019,p.8.

3【古希腊】亚里士多德:《尼各马可伦理学》,廖申白译，商务印书馆2019年版，第150页。

4【美】戴维·G.欧文:《侵权法的哲学基础》,张金海等译，北京大学出版社2016年版，第68、42页。

5张新宝:《侵权责任法》,中国人民大学出版社2016年版，第26页。

6 Glancy,D.J.,The Invention of the Right to Privacy,Arizona law review,21(1),1979,p.6.

7 Warren,S.D.,and Brandeis,L.D.,The Right to Privacy,Harvard law Review,4(5),1890,p.193.

8顾理平:《新闻法学》,中国广播电视出版社2005年版，第376、378页。

9魏永征:《新闻传播法教程》,中国人民大学出版社2002年版，第179—180页。

10(11)【德】哈尔特穆特·罗萨:《加速：现代社会中时间结构的改变》,董璐译，北京大学出版社2015年版，第278页。

11(12)蓝江:《外主体的诞生——数字时代下主体形态的流变》,《求索》2021年第3期。

12(13)Schwartz,P.M.,and Solove,D.J.,The PII Problem:Privacy and a New Concept of Personally Identifiable Information,New York University Law Review,86(6),2011,p.1877.

13(14)宋亚辉:《个人信息的私法保护模式研究——〈民法总则〉第111条的解释论》,《比较法研究》2019年第2期；丁晓东:《用户画像、个性化推荐与个人信息保护》,《环球法律评论》2019年第5期。

14(15)See “Carpenter V.United States,585 U.S.”,available at https://www.supremecourt.gov/opinions/17pdf/16-402_h315.pdf.

15(16)【美】特蕾莎·M.佩顿、西奥多·克莱普尔:《大数据时代的隐私》,郑淑红译，上海科学技术出版社2017年版，第98—99、189—191页。

16(17)顾理平:《智能生物识别技术：从身份识别到身体操控——公民隐私保护的视角》,《上海师范大学学报》(哲学社会科学版)2021年第5期。

17(18)广义的生物特征分为强生物特征、弱生物特征和软生物特征，覆盖了个人可感知和不可感知的生理数据与行为数据。强生物特征主要指指纹、DNA、面部ID等可直接识别且长期有效的个人信息。详见Mordini,E。,and Tzovaras,D.(Ed.),Second Generation Biometrics:The Ethical,Legal and Social Context,Berlin:Springer,2012.

18(19)Mordini,E。,and Tzovaras,D.(Ed.),Second Generation Biometrics:The Ethical,Legal and Social Context,Berlin:Springer,2012,pp.8-9.

19(20)See “Shannon Carpenter v.McDonald’s Corporation,Case No.2021CH02014”,available at https://www.classaction.org/media/carpenter-v-mcdonalds-corporation.pdf.

20(21)新华网：《0.5 元一份!谁在出卖我们的人脸信息》,2020-07-13,http://www.xinhuanet.com/politics/2020-07/13/c_1126232239.htm。

21(22)Dick,E.,Balancing User Privacy and Innovation in Augmented and Virtual Reality,Information Technology & Innovation Foundation,March 2021,available at https://itif.org/publications/2021/03/04/balancing-user-privacy-and-innovation-augmented-and-virtual-reality.

22(23)Solove,D.J.,The Digital Person:Technology and Privacy in the Information Age,New York:New York University Press,2004,p.44.

23(24)详见“庞理鹏与北京趣拿信息技术有限公司等隐私权纠纷二审民事判决书”,(2017)京01民终509号。

24(25)顾理平：《整合型隐私：大数据时代隐私的新类型》,《南京社会科学》2020年第4期。

25(26)详见“凌某某与北京微播视界科技有限公司隐私权、个人信息权益网络侵权责任纠纷一案民事判决书”,(2019)京0491民初6694号。

26(27)Hill,K.,Facebook recommended that this psychiatrist’s patients friend each other,Splinter,available at https://splinternews.com/facebook-recommended-that-this-psychiatrists-patients-f-1793861472.

27(28)顾理平、俞立根：《关联方信息共享与公民的隐私保护——基于手机 App 的研究》,《现代传播》(中国传媒大学学报)2019年第9期。

28(29)【英】阿里尔·扎拉奇、【美】莫里斯·E.斯图克：《算法的陷阱：超级平台、算法垄断与场景欺骗》,余潇译，中信出版社2018年版，第209、229页。

29(30)Krishnamurthy,B.,et al.,Privacy Diffusion on the Web:A Longitudinal Perspective,WWW 2009,April 20-24,2009,Madrid,Spain.

30(31)【英】阿里尔·扎拉奇、【美】莫里斯·E.斯图克：《算法的陷阱：超级平台、算法垄断与场景欺骗》,余潇译，中信出版社2018年版，第138、164页。

31(32)Gandy,O.H.,The Panoptic Sort:A Political Economy of Personal Information,Review by:Gilmore,S.,Contemporary Sociology,(1),1995,p.122.

32(33)(34)Taylor,L.,et al.(Ed.),Group Privacy:New Challenges of Data Technologies,Berlin:Springer,2017,pp.43,31,p.47.

33(35)巨量引擎营销观察：“巨量算数首发年终盘点报告，解锁 10 个新消费族群的新力量”,2020-12-16,https://mp.weixin.qq.com/s/mQb8teOZglD2ZihO5LWvYA。

34(36)早在 1980 年《OECD 委员会关于管理隐私保护和个人数据跨疆界流动的指导原则的建议书》就把“个人数据”界定为“与确定的和可以确定的个人相关的任何信息”。

35(37)详见“腾讯科技(深圳)有限公司等网络侵权责任纠纷一审民事判决书”,(2019)京 0491 民初 16142 号。

36(38)Barbaro,M.,and Zeller T.,Jr.,A Face Is Exposed for AOL Searcher No.4417749,New York Times,August 9,2006,available at https://www.nytimes.com/2006/08/09/technology/09aol.html?ex=1171771200&en=fc3fb33.

37(39)Thompson,S.A.,and Warzel,C.,Twelve Million Phones,One Dataset,Zero Privacy,New York Times,December 19,2019.Available at https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html.

38(40)张勇：《个人信息去识别化的刑法应对》,《国家检察官学院学报》2018 年第 4 期。

39(41)“情景脉络完整性理论”由尼森鲍姆(Helen Nissenbaum)提出，其理论框架需要以动态的思维来考察三个要素：信息流动过程中的行为主体(actor)都是谁；所涉及的信息类型(information type)是怎样的；以及该流动情景里的传播原则(transmission principle)是怎样的。

40(42)Loi,M.,and Christen,M.,Two Concepts of Group Privacy,Philosophy & Technology,33(1),2020,p.218.

41(43)See “California Consumer Privacy Act of 2018,1798.140”,available at https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5.

42(44)最高人民检察院：“最高人民检察院关于指派、聘请有专门知识的人参与办案若干问题的规定(试行)”,2018-04-04,https://www.spp.gov.cn/spp/zdgz/201804/t20180404_373479.shtml。

43(45)详见“凌某某与北京微播视界科技有限公司隐私权、个人信息权益网络侵权责任纠纷一案民事判决书”,(2019)京 0491 民初 6694 号。

44(46)详见“腾讯科技(深圳)有限公司等网络侵权责任纠纷一审民事判决书”,(2019)京 0491 民初 16142 号。