情感计算的信息隐私法律风险及其应对

褚婧一*

目次

摘要 情感计算是指以人类情感为机器学习对象，具有情感识别和情感分析功能的智能科技。情感计算不以信息主体的身份识别为前提，旨在通过对输入端生物反馈信息或状态的分析处理而输出情感信息。为应对情感信息不当获取与利用所带来的情感操纵风险，需以信息隐私理论为框架展开风险应对之探讨。然而，现行信息隐私规范却存在风险应对不足的问题。情感信息因其具体类型的多样性而难以划入敏感个人信息的保护范围；生物反馈信息因其情感识别而非身份识别的目的难以作为生物识别信息而被保护；作为信息处理之合法性基础的告知同意也因情感信息的特殊性存在被架空的风险。鉴于欧盟《人工智能法案(草案)》中规制情感识别的经验，我国可尝试在算法治理中构建“生物数据”的概念并以之为媒介展开技术风险的分级分领域规制，从而构建可信赖情感智能的数字环境。

关键词 情感计算 信息隐私 情感信息 生物反馈信息 可信赖情感智能

一、问题的提出

伴随人工智能技术的发展，机器不仅能够学习人的认知能力，还可以探知并回应人的情感思绪，人机交互突破了传统的行为交互而步入情感交互，情感智能时代(Emotional AI Era)正在到来。情感

* 中国政法大学法学院博士研究生。本文系国家留学基金委(CSC)国家留学基金(留金选〔2021〕70号)的阶段性研究成果。

智能以情感计算(affective computing)技术的广泛应用实现对人工智能的升级与扩展。“情感计算”一词最初由麻省理工学院罗莎琳德·皮卡德(Rosalind Picard)教授于1995年提出,[1]是指能够感知(measure)、理解(understand)、模仿(simulate)和影响(react to)人类情感的智能科技,涉及心理学、认知科学、密码学、计算机学等多个学科。[2] 根据皮卡德教授的观点,“情感”既包括以语音、图像和视频的实时处理为基础的短时情绪分析,又包括建立在自然语言处理等基础上的长期情感与观点分析。[3] “计算”则是指上述感知、理解、模仿、影响四项功能,涉及情感识别(recognition)与情感分析(analysis)两个维度。就两者的关系而言,情感是计算的语料也是产出,计算则是对情感加工处理的中间环节。

由于通过面部表情识别开展情感分析是当下情感计算的常见形式,故而该技术与人工智能时期的人脸识别产生关联。但相较于人脸识别,情感计算具有如下四点特殊性:第一,情感识别的方式具有多样性。探测面部表情并非情感测量的唯一方式,情感识别算法还可以通过测量皮肤电反应、脑电波活动、骨骼肌活动和皮肤温度等推断情绪状态。[4] 第二,情感识别不以身份识别为必要前提。情感识别算法输入端提取的是能够反映情感状态的生物反馈信息,如抿嘴、微笑、皱眉等,而面部五官的准确位置及基于此的身份信息则并非该技术的辐射范围。因此,理论上信息主体可以以匿名的方式存在于情感计算算法之中。[5] 第三,与人脸识别的正向验证不同,情感分析需预先建立算法模型(template),通过被采集的生物反馈信息反向推理(reverse inference)情感状态。[6] 目前基于面部信息的情感计算大多依据保罗·艾克曼(Paul Ekman)的基本情感理论,将表情类型化为开心、沮丧、惊讶、恐惧、生气和憎恶六种[7]并展开建模。第四,不同于建立在信息主体基础数据库之上的人脸识别,情感分析常表现为不依赖于信息主体过往数据的实时分析,即通过对输入端的实时处理判断情感状态。

不可否认的是,当下情感计算技术面临着准确性不足的批评,其根源在于艾克曼的基本情感理论遭受到强烈的质疑。批评者认为仅基于实时情感分析归纳的六种基本情感表现形态并不能准确地反映复杂的人类情感状态。[8] 纽约大学现在人工智能研究所(AI Now Institute)提出情

---

〔1〕 See Jennifer S. Bard, *Developing a Legal Framework for Regulating Emotion AI*, 27 Boston University Journal of Science & Technology Law 271, 277 (2020).

〔2〕 See Meredith Somers, *Emotion AI, Explained*, MIT Management Sloan School (8th Mar. 2019), https://mitsloan.mit.edu/ideas-made-to-matter/emotion-ai-explained.

〔3〕 See Rosalind W. Picard, *Affective Computing: Challenges*, 59 International Journal of Human-computer Studies 55, 55-64 (2003).

〔4〕 See Goran Udovičić, Jurica Derek, Mladen Russo & Marjan Sikora, *Wearable Emotion Recognition System based on GSR and PPG Signals*, in Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care (MMHealth '17), Association for Computing Machinery, New York, 2017, p. 53-59.

〔5〕 但不可否认的是,由于情感计算与人脸识别的算法共同作用于人脸面部信息,加之技术黑箱的存在,实践中很难判断技术使用者是否将两种算法的功能合并,即识别情感的同时也确认了信息主体的身份。

〔6〕 See Andrew McStay, *Emotional AI, Soft Biometrics and the Surveillance of Emotional Life: An Unusual Consensus on Privacy*, 7 Big Data & Society 1, 2 (2020).

〔7〕 1970年,保罗·艾克曼提出可以从人们的面部表情中识别情绪状态,其认为即便存在不同的文化背景,开心、沮丧、惊讶、恐怖、生气和憎恶这六种情绪仍然具有普遍性。See Paul Picard, E. Richard Sorenson, Wallace V. Friesen, *Pan-Curtural Elements in Facial Displays of Emotion*, 164 Science 86, 86-88 (1969).

〔8〕 See Lisa Feldman Barrett, *Are Emotions Natural Kinds?*, 1 Perspectives on Psychological Science 28, 28-58 (2006); also see Robin Markwica, *Emotional Choices: How the Logic of Affect Shapes Coercive Diplomacy*, Oxford University Press, 2018, p. 72.

境、社会和文化因素会对情感表达产生重要的影响。[9] 因此，实时采集的面部表情信息是“缺乏背景(lack of background)、脱离语境(out-of-context)、不全面(incomplete)和部分污染(partially polluted)”的，[10]基于此情感识别的准确性将大打折扣。但值得思考的是，为提高算法准确性，解决途径之一则是收集更多的情境、环境与背景信息，随之而来的便是将有更多的个人信息处于风险状态之中。尽管面临种种质疑，情感计算在遍布式计算的帮助下仍具有被广泛应用于日常生活的现实性与可能性。例如，其既可以装载于智能手机、自动驾驶汽车、儿童玩具等日常物品，也可以应用于保险、教育、警务、边境管控等行业领域，还可以配置于办公室、医院、监狱、商店等公共场所。[11]

技术的应用热潮需要法律的冷静省思。情感计算也将呈现出对个人权益影响的新面貌。一方面，情感计算步入了人类最深层也是最私密的空间——大脑，这里是情绪与思想的迸发地，是决策做出的中枢机构，关系着意志自由与人的尊严。由此，该技术将个人权益的关注点从行为转向思想，进而将意志自由聚焦于情感自由，引发法律应如何保护人的情感的问题。另一方面，由于情感计算不仅可以应用于个人，还可以应用于特定或不特定的群体，所以对群体权益或群体利益(collective good)也产生了深远影响。例如，2014 年 Facebook 对近 70 万用户开展了情感计算试验，从而验证其是否具有通过操纵新闻推送来影响用户情绪的能力。[12] 于是，法律还需解决特定的个人权利能否扩展至群体权利或群体利益，以及群体利益该如何保护的问题。

目前国内外关于情感计算的社会科学研究刚刚起步。在域外，媒体与新闻传播领域学者就不同社会群体对于情感计算的隐私态度[13]以及共感媒体(empathic media)所带来的隐私忧虑[14]展开了讨论。法学领域的研究则关注情感计算所带来的宪法挑战[15]及其伦理与法律规制[16]。学者们关注到这一技术对个人权益的潜在威胁：英国学者安德鲁·麦克斯泰(Andrew McStay)指出其将损害不特定群体的隐私权益；比利时学者佩吉·瓦克尔(Peggy Valcke)提出该技术具有情感操纵的潜能，将对个人意志自由及人的尊严造成宪法挑战；美国学者詹妮弗·巴德(Jennifer Bard)认为该技术的不当运用将进一步加剧种族歧视等社会不公，损害特定群体的权益。我国法学领域目前仅有王禄生教授撰文以情感计算在医疗、教育等五个领域的应用为观察点，指出技术应用将给个人信息、隐私权、人身自由及人的尊严的平等保护带来风险，提出以“轻推”为底线、风险为基础的差异规制策略。[17] 从既有文献来看，学者们认同新技术将对个人信息与隐私保护造

---

〔9〕 See AI Now Institute, AI Now Report 2018, p. 14, https://kennisopenbaarbestuur.nl/media/257225/ai_now_2018_report.pdf.

〔10〕 See Bert-Jaap Koops, *On Decision Transparency, or How to Enhance Data Protection after the Computational Turn*, in Privacy Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology, Routledge, 2013, p. 199.

〔11〕 See McStay, *supra* note〔6〕, at 1.

〔12〕 See Adam D. I. Kramer, Jamie E. Guillory & Jeffrey T. Hancock, *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 111 Proceedings of the National Academy of Sciences of the United States of America 8788, 8788 - 8790 (2014), available at https://www.pnas.org/doi/pdf/10.1073/pnas.1320040111.

〔13〕 See McStay, *supra* note〔6〕, at 1.

〔14〕 See Andrew McStay, *Emotional AI: The Rise of Empathic Media*, SAGE Publications, 2018.

〔15〕 See Peggy Valcke, Damian Clifford & Vilté Dessers, *Constitutional Challenges in the Emotional AI Era*, in Constitutional Challenges in the Algorithmic Society, Cambridge University Press, 2021, p. 57.

〔16〕 See Bard, *supra* note〔1〕, at 271.

〔17〕 参见王禄生：《情感计算的应用困境及其法律规制》，载《东方法学》2021 年第 4 期。

成严重风险，但由于研究视角较为宏观而尚未展开具体探讨。因此，本文在既有研究的基础上以信息隐私作为讨论框架，明晰情感计算在该领域所带来的具体风险，研究其与现有信息隐私规范间的衔接与空隙，并以欧盟经验为参考对空隙之填补提出建议。

二、情感计算引发情感操纵的信息隐私风险

虽然当下法律将理性与情感二分，以“理性人”假设为前提展开相应的规则构建，[18]但不可否认的是人们做出决策时往往不仅依赖理性（rationality），情感也是影响个人决策的重要因素。1999年，乔恩·汉森（Jon Hanson）与道格拉斯·凯泽（Douglas Kysar）从行为经济学的角度提出“市场操纵”（market manipulation）的概念，其反对传统经济模型中“个人基于自身利益最大化而做出理性决策”的前提假设，认为市场中的消费者是“可预期的非理性主体”（predictably irrational），经营者企业有可能通过利用（exploit）、控制（control）或不当影响（undue interference）消费者的偏见与认知局限从而获取更多利润。[19] 在应用情感计算技术的网络平台与用户关系中，情感操纵（emotion manipulation）是指平台基于情感识别向不同用户投放特定目的的信息，以利用、控制或影响用户情绪的方式操纵其个人决策从而最终获益。在此过程中，获取情感信息是操纵的前提，影响个人决策则是操纵的后果。

（一）情感信息的不当获取将侵犯私域隐私

情感操纵往往建立在用户画像的基础之上，情感识别则为精确勾勒用户画像奠定基础，故情感信息的获取是实现情感操纵的前提。情感是具有高度私密性的“自我谈话”，情感信息的不当获取很有可能构成对隐私权的侵犯。

在情感识别出现之时，有部分学者认为基于面部信息推断情感状态的情感计算并不涉及对个人隐私的侵犯，其主张面部表情信息并不属于“私域”的范畴，主要理由有两点：第一，与思想、态度等精神活动相比，由面部表情所展示的情感是一种“公共泄露”（public leaks），从而并不具备成为私域的潜质；第二，面部表情具有社交性，情感表达也具有交流性，即使是被抑制与隐藏的“面部表情”（disguised emotion）也具有潜在的交流属性（implicitly social）。[20]

事实却是，面部信息公共性的论断会引发人们的焦虑与不安。以商品零售为例，情感计算的应用有两种方式：第一种是在商店的出入口安装配有情感识别功能的监控器，以此探测顾客对于商店整体的满意度；另一种是在某些商品的标签上安装情感探测装置，记录顾客对于特定种类商品、包装以及摆放顺序等的情绪感受，从而帮助商店优化销售行为。[21] 那么问题在于顾客在此情境下产生的情感信息属于隐私的范畴吗？对此答案是肯定的。情感具有相较于人的行为而言更高程度的私密性。如果说通讯隐私是建立在双方或多方主体之间的私密谈话，那么情感隐私则是私域范围内的自我谈话（*forum internum*）。[22] 正如美国社会学家和生物伦理学家保罗·鲁特·

---

〔18〕 See Terry A. Maroney, *Law and Emotion: A Proposed Taxonomy of an Emerging Field*, 30 Law and Human Behaviour 119, 119 (2006).

〔19〕 See Ryan Calo, *Digital Market Manipulation*, 82 The George Washington Law Review 995, 995 (2014).

〔20〕 See Andrew McStay, *Privacy and the Media*, SAGE Publications Ltd, 2017, p. 140.

〔21〕 Ibid., at 135 - 144.

〔22〕 See Karen Yeung, A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility within a Human Rights Framework (DRAFT), Council of Europe 2018, p. 79, https://rm.coe.int/draft-study-of-the-implications-of-advanced-digital-technologies-inclu/16808ef255.

沃尔普(Paul Root Wolpe)所言,“人的头脑应该被界定为绝对的私域,不应以违背主体意愿的方式探察人的思想”。[23] 若赋予一般情况下的情感探测以合法性,人们将不可避免地被深度“透明化”,由此引发监控社会 2.0 版本的信息隐私风险。正因如此,英国学者安德鲁·麦克斯泰将有关个人情感的信息命名为“亲密信息”(intimate information),认为其是敏感且值得保护的。接下来,这类亲密信息有无隐私的例外呢,即在上述例子中商店记录与商品有关的顾客情绪仍然是对隐私的侵犯吗?笔者认为,人的情感具有复杂性,商店记录的顾客高兴与否的情感状态可能并不是对商品的单一反馈,比如顾客可能因为今天的好天气而处于一个持续兴奋的状态,那么尤其是在顾客不知情情况下对此情绪的记录则侵犯了顾客的隐私权。

此外,若将信息时代中的个人视为一个信息系统,个人对自身信息享有一定控制权。一方面,控制自身产出哪些信息,即自身情感是否要以数据信息的方式呈现;另一方面,控制哪些信息可以从私域流出,进入公共领域或与其他私主体共享。情感信息的不当获取则是对个人上述两种控制权的减损,对应表现为平台往往诱导用户做出不利于自身的隐私决定,允许收集涉及情感推断的个人信息。[24]

(二)情感信息的不当使用将减损自治自决

操纵意味着对充分思考与慎重决策的排斥;[25]而自治并非没有选择余地或不行使做出选择的能力,而是含有强制成分的非自由选择的对立面。[26] 如上文所言,情感也是影响个人决策(decision-making)的重要因素,情感信息的不当使用将产生减损个人自治与自决的法律风险,从而造成操纵个人决策的后果。

在互联网平台与用户的场景中,平台通过采集情感信息形成更为精准的用户画像,从而针对不同用户投送更具个性化的信息,提高平台的服务能力。但是,当平台基于特定目的,尤其是不当目的投放信息时,先前的用户画像则成为情感操纵的基础,平台进而以潜移默化的方式逐渐改变着用户的情绪与观点。例如,在剑桥分析丑闻(Cambridge Analytica Data Scandal)事件中,剑桥分析通过 Facebook 收集用户的性格特征并以此判断特朗普的潜在支持者,随后以精准投放信息的方式影响选民的决策。比如向摇摆选民投放希拉里的负面新闻促使其转向特朗普,或者对冒险性格的选民推送枪支可以“以暴制暴”的信息促使其支持特朗普主张的持枪权。[27]

然而,不可否认“情感操纵”是一个难以界定的概念。与剑桥分析丑闻事件形成对比的是,政府要求烟草企业在商品包装上印制“吸烟有害健康”的标语或者能够引起人不适的有害健康的图片,这同样会在一定程度上通过影响情感的方式助推(nudge)个人减少对烟草的消费选择,但是后者并不被认为是一种操纵。如市场操纵理论提出者所言,市场操纵是一种以利益为目的的助推

---

〔23〕 Paul Root Wolpe, *Is My Mind Mine? Neuroscience and the State*, The Center for Medical Humanities & Ethics, https://archive.org/details/podcast_templeton-research-lectures_is-my-mind-mine-neuroscience_1000085038155, last visited at 2022-08-22.

〔24〕 例如,企业利用“暗模式”(dark pattern)在用户隐私设置中默认自动勾选同意情感信息采集的选项。See The Norwegian Consumer Council, Deceived by Design (Report 2018), p. 7, https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf, last visited at 2022-08-02.

〔25〕 “A statement or action is manipulative to the extent that it does not sufficiently engage or appeal to people's capacity for reflective and deliberative choice.” See Cass R. Sunstein, *Fifty Shades of Manipulation*, 1 Journal of Behavioral Marketing 213, 213 (2016).

〔26〕 See Joseph Raz, *The Morality of Freedom*, Clarendon Press, 1986, p. 382.

〔27〕 See Luke Stark & Kate Crawford, *The Conservatism of Emoji: Work, Affect, and Communication*, 1 Social Media + Society 1, 1-8 (2015).

(nudge for profit)。[28] 因此在界定情感操纵时,有必要厘清以用户利益为导向的"合理劝导"(permissible persuasion)和以企业经济利益为优先的"不可接受之操控"(unacceptable manipulation),[29]只有后者落入操纵的范畴才构成对个人自治与自决的不当减损。

三、情感计算风险应在信息隐私理论内得到整体回应

情感操纵既是对情感隐私的直接侵入,也是对情感信息的不当采集和处理。直观来看,个人信息保护与隐私权保护均适用于情感计算的风险应对。本文主张以信息隐私理论作为讨论框架并在其下解决情感计算的风险应对问题。但在界定信息隐私理论之前,有必要明晰隐私权与个人信息保护间的关系,从而为下文的整体回应奠定基础。具言之,隐私权与个人信息保护具有保护人的尊严的共同目标,二者也因各有一定对方无法覆盖的保护法益与保护目的而具有特殊性。此外,隐私保护是个人信息保护的目的之一,二者具有交叉重合的关系,且这一重合区域在信息时代具有扩大的趋势。[30]

(一)情感计算中信息与隐私的复杂运行样态

与人工智能时代的技术不同,情感计算中信息与隐私的呈现样态具有两点特殊性:其一,基于该技术不仅可以应用于特定个体,还可以应用于特定或不特定群体,故而隐私还将以群体隐私的面貌出现;其二,由于该技术不以身份识别为前提,可识别性并非系统内信息的必备要件,可识别性的个人信息与不具备可识别性的信息在系统内共存。最终,上述两点特殊性决定了情感计算中隐私与信息呈现更为复杂的交织状态,因此需要信息隐私理论的整体性回应。

1. 个体隐私与群体隐私

由于技术应用对象的多样性,情感计算中的情感隐私不但可以个人隐私的方式存在,还可以群体隐私的方式出现,成为一种"群体利益"(a collective good)。[31] 以英国伦敦皮卡迪利(Piccadilly)灯塔的摄像头事件为例,其能够检测到皮卡迪利广场爱神雕像周围人群中的面部信息。当摄像头识别出一张面孔时,情感计算技术会计算出大致的年龄与性别,并根据面部表情如笑容或皱眉等识别出基本的情绪状态,以此优化皮卡迪利灯塔上的广告推送。[32] 这一事件随后引发了严重的道德伦理质疑,因为经常在该灯塔附近穿行的人群在无意识情况下被记录了情感与身体特征信息,即使相关信息无法识别到特定的个人,该群体的隐私权益仍受到严重的挑战。[33]

于是,情感计算引发了对于隐私权权利主体的反思,即隐私不仅是一项个人权利,也是一项群

---

〔28〕 See Calo, *supra* note〔19〕, at 1001.

〔29〕 See Valcke, Clifford & Dessers, *supra* note〔15〕, at 62.

〔30〕 "Data protection overlaps considerably with the right to privacy, as they both ensure informational or data privacy, but data protection serves a number of purposes that privacy does not and vice versa." See Orla Lynskey, *The Foundations of EU Data Protection Law*, Oxford University Press, 2015, p. 103-104,130. 这一观点被欧洲人权法院认可而成为当下的主流观点。我国学者也不乏对隐私权与个人信息保护关系的讨论。从立法与司法实践来看,我国采取了与欧盟相似的交叉重合关系认定。《民法典》第1043条规定:"个人信息中的私密信息,适用于有关隐私权的规定;没有规定的,适用有关个人信息保护的规定。"由此可见,我国隐私权与个人信息保护各自作为权利与权益存在一定的内容重合,私密信息则在重合的区域之内。

〔31〕 See Lilian Edwards, *Privacy, Security and Data Protection in Smart Cities: A Critical EU Law Perspective*, 2 European Data Protection Law Review 28, 38-44 (2016).

〔32〕 See Valcke, Clifford & Dessers, *supra* note〔15〕, at 62.

〔33〕 See Andrew McStay, Lachlan Urquhart, "*This Time with Feeling?*" *Assessing EU Data Governance Implications of out of Home Appraisal Based Emotional AI*, 24 First Monday 1, 1-16 (2019).

体权利，群体隐私应由群体作为一个整体持有而非群体成员单独享有。[34] “群体隐私”的支持者卢西亚诺·弗洛里迪(Luciano Floridi)曾以“沙丁鱼与白鲸”的例子生动描绘了个体隐私与群体隐私的关系，[35]在情感计算的背景下故事可以这样改编：情感计算在遍布式计算时代编织成了一张巨大的渔网，个人如同沙丁鱼一般游走在浅滩，这张网试图抓住的并非单一的沙丁鱼，而是整个鱼群，因此浅滩的保护便尤为重要。

2. 个人信息与非个人信息

由于情感计算输入与输出信息都不必具备针对信息主体的可识别性(non-personally identifying)，个人信息与非个人信息在算法系统内同时存在，一般个人信息、生物识别信息、敏感个人信息、生物反馈信息与情感信息在系统内交织存在，使得算法内信息同样呈现出复杂的样态。在展开具体分析之前，有必要首先厘清该技术输入端与输出端的信息种类，以下表为例做出不完全列举。(如表1)

表1　情感计算中的输入与输出信息

输入端		输出端
通过测量： 面部表情； 皮肤汗液(皮肤肌电反应)； 心动周期(心电图)； 脑电波(脑电图)； 骨骼肌活动(肌电图)； 呼吸； 皮肤温度； ……	**数据分析**	得出： 情绪状态(快乐、愤怒)； 情感状态(压力或焦虑程度)； 注意力集中程度； 内在品质(勇敢、坚韧)； ……

我国个人信息保护立法受到欧盟立法经验的影响，可识别性标准作为不证自明的通说被采纳。[36]《个人信息保护法》第4条规定：“个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息，不包括匿名化处理后的信息。”[37]其中，已识别(identified)，是指能够识别出特定自然人；可识别(identifiable)，是指存在识别出特定自然人的可能性。[38] 至于何为“识别”，欧盟采用了“特定化(single out)”的判断标准，[39]我国虽尚未有相关案例，但在学理解释中也接纳了这一观点。具言之，“特定化”是指“将特定个体挑出”(single out a particular

〔34〕 See Luciano Floridi, *Open Data*, *Data Protection*, *and Group Privacy*, 27 Philosophy & Technology 1, 1-3 (2014).

〔35〕 “There are very few Moby-Dicks. Most of us are sardines. The individual sardine may believe that the encircling net is trying to catch it. It is not. It is trying to catch the whole shoal. It is therefore the shoal that needs to be protected, if the sardine is to be saved.” Ibid., at 3.

〔36〕 参见曹博：《个人信息可识别性解释路径的反思与重构》，载《行政法学研究》2022年第4期，第134页。

〔37〕 欧盟《通用数据保护条例》第4(1)条将“个人信息”界定为“与已识别或可识别自然人相关的信息”。

〔38〕 参见江必新、郭峰主编：《〈中华人民共和国个人信息保护法〉条文理解与适用》，人民法院出版社2021年版，第39页。

〔39〕 See Article 29 Working Party “Opinion 4/2007 on the concept of personal data” (WP 136, 20 June 2007), p. 13-14; also see Christopher Kuner, Lee A. Bygrave et al. eds., *The EU General Data Protection Regulation (GDPR): A Commentary*, Oxford University Press, 2020, p. 110.

person)或“聚焦于特定的个体”(zoom in on a flesh and bone individual)。〔40〕因此,“识别”既指在群体中将某信息主体特定化、具体化(single out a person from others)(比如公共区域视频监控中的犯罪人识别系统),也指对单个信息主体的身份确认(比如上下班打卡的人脸识别系统)。

若以特定化和非特定化二分情感计算的应用场景,结合个体与群体的应用对象则有下表所示的四种情形。(如表2)需要指出的是,“个人信息”中的“个人”既包括特定个体(natural person),也包括特定群体(one or more natural persons)。〔41〕特定化场景中的输入信息与输出信息指向特定的个体或群体,如驾驶搭载智能系统的汽车司机、课堂中的二十名学生,这些主体的身份已经被提前确认甚至建立起相关的用户档案,因而输入端与输出端的信息满足可识别标准。继而,问题在于非特定化场景中个人信息与非个人信息的判断。〔42〕此时,可以个体、群体为横坐标,输入信息与输出信息为纵坐标展开具体分析。(如表3)

表2　情感计算的应用场景

应用对象 / 应用场景	个　　体	群　　体
特定化场景	例如汽车的“智能安全系统”,当检测到司机情绪愤怒时能够自动播放欢快的音乐。〔43〕	例如智慧课堂中的“慧眼系统”,能够识别学生的注意力集中程度及课堂情绪。〔44〕
非特定化场景	例如放置在医院公共区域的压力检测仪,实时显示压力指数。	例如上文英国伦敦皮卡迪利灯塔摄像头事件。

表3　非特定化场景中不构成个人信息的理由

信息主体 / 信息端口	个　　体	群　　体
输入信息	A+B	A+B+C
输出信息	B+D	B+D+C

注:A 由于抓取技术与实时分析,输入端未生成被记录的“信息”;
B 情感识别不以身份识别为前提,信息主体处于匿名状态,难以被特定化;
C 作为应用对象的群体往往由不特定个体组成,整个群体的形成具有随机性,个体在其中的匿名化程度更高;
D 人类的情感状态具有一定的普遍性,情感信息的特殊性为特定化增加了难度。

对于非特定化场景中的输入信息而言,情感计算中识别技术的具体类型将影响其是否构成个人信息的判断。(1)以抓取(extract)为识别技术的情感计算导致输入端并未生成“信息”(表3中

〔40〕 See Article 29 Working Party “Opinion 4/2007 on the concept of personal data” (WP 136, 20 June 2007), p. 13 - 14; also see Nadezhda Purtova, *From Knowing by Name to Targeting: The Meaning of Identification under the GDPR*, 11 International Data Privacy Law 163, 173 - 177 (2021).

〔41〕 See Kuner, A. Bygrave, Docksey & Drechsler, *supra* note〔39〕, at 111.

〔42〕 需要说明的是,此处并非“一刀切”地判断非特定化场景中哪些属于个人信息,哪些不属于个人信息,而是明确其中基于可识别标准的判断思路,实践中个人信息的判断仍需结合具体场景而具体分析。

〔43〕 见前注〔17〕,王禄生文,第53页。

〔44〕 同上注,第52页。

的理由 A)。具言之,伴随技术的发展,以实时方式识别人的情感愈发普遍,其抓取体现身体特征(bodily traits)的生物反馈(bio-feedback)状态且不会以电子或者其他方式记录下来,继而在情感计算算法中直接生成情感信息,由此输入端并不存在"信息"。(2) 理想匿名化状态的情感计算并不需要识别或验证被测试者的身份(表 3 中的理由 B),被测试者始终处于匿名的状态之中,因而输入信息难以构成个人信息。以面部表情为例,理想匿名化状态的情感识别在输入端会基于算法模型需要生成表情数据,如 A 眉毛弯曲弧度为 X、B 上嘴唇扬起下嘴唇紧绷等类似信息。[45] 一方面,这类表情信息具有时空场域的限制,某一时刻或时段细碎的生物反馈信息即使与其他信息结合也难以回溯至特定的信息主体,因而该信息难以具备主体指向性;另一方面,需要注意匿名化使得这类信息并不包含如 A 眉毛形状样态等具有个人可识别性的生物数据,即面部识别信息并不会作为单独或附加的信息被分类和存储,[46]因此难以被界定为"个人信息"。(3) 在非理想的匿名状态下,输入端生物反馈信息可能会与生物识别信息同时存在,例如 A 的眉毛弯曲弧度是以 A 的眉毛样态的方式呈现的,则其具有成为间接可识别个人信息的可能。

至于非特定化场景中的输出信息,其回溯至信息主体的难度要远大于其他种类的信息(表 3 中的理由 D),难以被认定为"个人信息"。具言之,人类的情感具有普遍性,从艾克曼最初将情感归为六种类型则可见一斑。例如,在检测中输出如"A 的情感状态为开心 80%、惊讶 10%、焦虑 10%"的情感信息,虽不否认其具有成为"间接可识别个人信息"的潜质,但回溯至特定信息主体是十分困难的。此外,从群体与个体的对比来看,如果说个体在该技术下天然地处于匿名状态之中,那么由随机个体组成的群体则使得"特定化"难上加难(表 3 中的理由 C)。

(二) 信息隐私理论内情感计算的法益类型化

基于情感计算中信息与隐私的复杂运行状态,群体隐私或以群体作为对象的个人信息的方式呈现而可享有个人信息保护,非个人信息或以个人隐私的状态存在而受到隐私权保护,二者的复杂交织需要隐私权与个人信息保护理论的共同因应。立足于隐私权与个人信息保护的交叉重合关系,当下"信息隐私"的两种界定方式均难以满足共同应对之需要。第一种,有学者基于"大隐私"范畴而提出"信息隐私权",认为信息隐私是隐私在信息时代的主要表现形式,[47]信息隐私权是隐私权在数字时代的新变革。[48] 这在本质上将信息隐私视为隐私权,虽然法益保护的重心放在了个人信息与隐私保护的交集,但是法益保护的范围仍未超出作为私域的隐私。(如图 1)第二种观点则将信息隐私视为信息化的隐私或私密的个人信息,[49]即信息隐私是二者的共有法益,信息隐私=个人信息保护∩隐私权,此种界定方式同样难以满足情感智能时代的共同应对需要。(如图 2)

---

〔45〕 See Peter Lewinski, Jan Trzaskowski & Joasia Luzak, *Face and Emotion Recognition on Commercial Property under EU Data Protection Law*, 33 Psychology & Marketing 729, 729 - 746 (2016).

〔46〕 "... extracting only the features needed, and then compressing them. A particular psychological state or sociodemographic characteristic is recognized through comparison with a dataset of such modeled images using an artificial neural network." See Jan Czarnocki, *Will New Definitions of Emotion Recognition and Biometric Data Hamper the Objectives of the Proposed AI Act?*, 2021 IEEE Xplore 1, 1 - 4 (2021).

〔47〕 有学者将信息隐私权视为一个权利演化树(evolutionary tree),从而容纳不同代际隐私权。参见余成峰:《信息隐私权的宪法时刻:规范基础与体系重构》,载《中外法学》2021 年第 1 期,第 52 页。

〔48〕 "Regard data protection as the most recent phase in the evolution of the right to privacy." See Daniel Solove, *The Digital Person: Technology and Privacy in the Information Age*, NYU Press, 2004, p. 75.

〔49〕 Ibid., at 8 - 9.

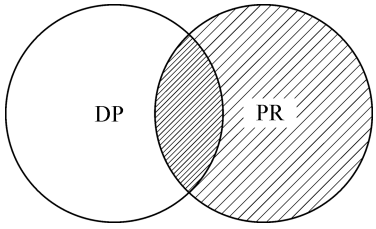

图1　信息隐私权所保护的法益

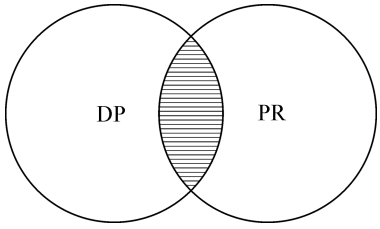

图2　隐私信息＝个人信息保护∩隐私权

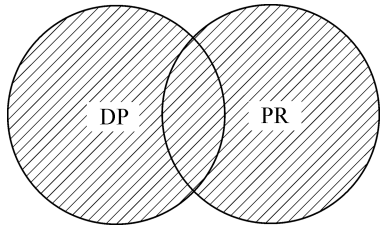

图3　信息隐私＝个人信息保护∪隐私权

注：DP为个人信息保护，PR为隐私权，阴影部分为信息隐私保护的法益。

为划定情感智能时代的信息隐私框架，可首先提出这样一种假设：既然个人信息保护与隐私权是常被并提的法律概念，那么是否存在这样一种法律风险，需要二者的共同应对，且威胁二者所要保护的法益总和。若该假设是成立的，此时便产生了本文所主张的信息隐私概念，即从法益保护的范围看，“信息隐私＝个人信息保护∪隐私权”。（如图3）若从微观的角度讲，情感计算的法律风险作为个例既使信息主体的情感信息安全受到威胁，又因闯入了人脑这一人类最私密的空间而挑战隐私权，故而从事实上看这样的法律风险是存在的。从宏观角度分析，情感计算这类技术风险并非个例，即数字时代同时威胁个人信息权益与隐私权利的技术具有一定的普遍性，正因此欧洲学者提出以“数字宪政”讨论数字技术对宪法基本价值的影响。〔50〕在我国的语境下，参考上文隐私权与个人信息保护共同旨在保护人的尊严的价值目标，那么当数字技术企图减损人的尊严时，该权利与权益则必然受到威胁，该风险可被称为“人的数字化（digitalisation of human beings）”风险。“人的数字化”风险是指人成为被测量的客体，并在此过程中逐渐丧失对自身数据的控制权，最终可能沦为技术支配客体的风险。至此，法益角度“信息隐私＝个人信息保护∪隐私权”的公式得以成立，以此厘定了作为本文讨论框架的信息隐私的范围。

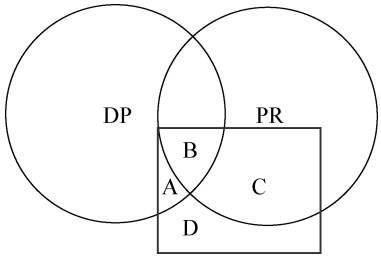

图4　情感计算应用中所涉及的法益

最后，若以情感计算所涉及的法益作为一个平面与图3相叠加，则得出如图4所示的图像，由此可以明确情感计算中不同种类信息隐私的请求权基础。其中，DP代表个人信息保护，PR则代表隐私权保护。但在厘定ABCD四项所指向的法益之前，有必要首先基于上文“信息与隐私的复杂运行样态”而界定DP与PR的排斥项，即哪些信息或状态无法为各自所涵盖。对此，DP的排斥项为：（1）不具备可识别性的信息（匿名或不特定个体或群体的信息）；（2）未以电子或其他方式记录而未生成“信息”的数据或状态。因群体隐私尚无法为隐私权理论所包含，PR的排斥项之一为不特定群体；又因情感隐私与隐私主体之间存在一种亲密状态而构成私密隐私，故PR的另一排斥项为非亲密（的信息或状态）。以此为前提，可做出如下的分类：

〔50〕欧洲学者提出“数字宪政”（digital constitutionalism）以回应数字技术对公民基本权利的挑战。［个人信息受保护权（right to the protection of personal data）与隐私权为《欧盟基本权利宪章》规定的公民基本权利。］“数字宪政”的规范价值（normative value）在于对传统宪法价值的重新调整，以适应数字时代的新挑战；“数字宪政”的程序价值（programmatic value）在于其作为一种新的路径去思考数字时代下基本权利保护与权力限制。See Giovanni De Gregorio, *The Rise of Digital Constitutionalism in the European Union*, 19 International Journal of Constitutional Law 56, 56－70 (2021).

A为“可识别+非亲密”的信息，例如智慧课堂“慧眼系统”记录的学生上课注意力集中程度的信息。〔51〕

B为“可识别+亲密”的信息，例如汽车“智能安全系统”所记录的司机情感信息。

C为“未记录为信息+亲密”的状态，即与个人信息无涉的隐私法益。这类法益在情感计算中可以两种形式存在：第一种是在输入端不生成信息的技术中，被探测主体的面部表情、皮肤温度等生物反馈状态构成该法益保护的对象；第二种则是在直接输出群体情感状态的技术中，被探测个体的情感状态也可能构成该法益保护的对象。

D在理论上为“不可识别/未记录为信息+非亲密/不特定群体”的信息或状态，实际则常常以“不可识别+不特定群体”的形式存在。又因在理想匿名状态下，不特定群体实则排除了信息主体的识别可能，故D可被认为是不特定群体的信息隐私法益，其尚未落入现有DP与PR的保护范围，而又与这两个概念有着紧密的关联。现实中这类场景广泛出现，例如英国伦敦皮卡迪利灯塔摄像头事件。

若对侵害ABCD四种法益产生的危害后果进行排序，则暴露可识别性信息的危害大于不具有可识别性的信息，侵入亲密状态的损害大于非亲密状态，由此产生B>A，C>D的排序。其中，信息主体对B享有个人信息权益与隐私权的双重请求权基础，保护力度最强。结合上文鲸鱼、沙丁鱼与浅滩的例子，ABC三类信息主体由于是相对特定的、可识别的，因此可以类比为沙丁鱼；而D类信息主体则是随机与完全匿名情况下的主体，其往往以相对不特定群体的方式存在，可类比为沙丁鱼群，因此D更加依赖于对浅滩整体环境的保护。在情感智能时代，每个人都有步入沙丁鱼群的潜在可能，故D类相关权益的保护直接关涉公众对于情感计算技术的信任，影响数字社会整体环境的安全度与可信度。

四、现有规范难以契合情感计算中的法益保护需求

（一）情感信息难以作为敏感个人信息而被保护

“情感信息应如何适用于个人信息保护规定”是情感计算带给现行法的第一个难题，由此引发对情感信息与个人信息、情感信息与敏感个人信息之间关系的思考。

情感信息与个人信息的关系问题似乎不难回答：情感信息在具备可识别性时则成为个人信息，受到个人信息保护法的保护；情感信息在不具备可识别性时可能构成情感隐私，受到民法隐私权的保护。但是，从实践来看，算法中的情感信息是否具备可识别性成为判断的难题。在英国伦敦皮卡迪利灯塔摄像头事件中，技术使用者声称其仅依据往来人群的面部信息推测大致的年龄、性别及心情，相关人脸图像未被记录，且已收集的信息没有将信息主体特定化的任何可能。其广告牌所有者Landsec称“我们并不能也永远不会利用情感计算展开身份识别”。〔52〕尽管如此，基于情感信息与面部信息的紧密关系，加之情感算法中技术黑箱的存在，使用者很可能以算法不具备身份识别为由规避情感信息的保护，进而人为将情感信息的法益由A转向C。然而，在作为情感隐私法益保护的C中，权利主体一方面需要证明该情感隐私与其存在关联，另一方面要证明本人因情感计算而受到损害。就前者而言，关联性的证立虽然易于可识别性，但是仍存在主体特定化

〔51〕 此处之所以将智慧课堂场景中的学生注意力集中程度的信息视为“非亲密信息”是因为：一方面，学生的注意力集中情况可以通过眼神、表情等具有外部性的状态显露出来，构成“公共泄露”(public leak)，即使没有该智能系统，有经验的教师也可以大致判断出学生的注意力程度；另一方面，这类信息具有互动性，作为师生间的一种沟通方式，学生在课堂中的表现恰好要与教师分享。所以将该场景的注意力信息归类于对学生而言不具有亲密性的信息。

〔52〕 See Matthew Moore, *Personalised Ads Delivered by the Billboard that's Got Its Eye on You*, The Times (London, 18 October 2017), https://www.thetimes.co.uk/article/personalised-ads-delivered-by-the-billboard-that-s-got-its-eye-on-you-hrxwrrb3z.

的难题。就后者而言，情感隐私被损害的事实要件往往也难以证明。换言之，在司法实践中判断隐私权是否受到侵犯的核心标准在于是否导致受害人社会评价的降低，〔53〕然而情感作为一种普遍的内心表达或难以与社会评价产生直接关联。所以，在个人信息保护中情感信息面临可识别性判断的难题，在隐私权保护中又面临了损害事实难以证立的困境。

因为情感信息内容的广泛性，其能否被认定为敏感个人信息也需要在具体场景中加以分析。我国《个人信息保护法》第28条对敏感个人信息的规定采用了“概括＋列举”的方式，虽然法条采用了非穷尽式列举，但情感信息并不在明确列举范围内。那么问题则在于情感信息能否因满足敏感个人信息之本质特征而被纳入特殊保护，判断标准则在于其是否与自然人的人格尊严或者人身、财产安全紧密相关。如表1中情感计算输出端信息内容所示，情感信息是一个内容广泛的概念，喜怒哀乐的情绪、爱恨情仇的情感均被纳入其中。此外，情感信息还是一个与场景有关的概念，例如图4中的A法益，课堂中学生的注意力集中度存在于师生关系之场景中，其作为师生间的沟通方式很难被认定为与人格尊严紧密相关，且不当利用也并不会对人身、财产安全造成严重损害，因此该信息难以被认定为敏感个人信息。综上，虽然不否认情感信息往往是私密的、与人格之形成有关的信息，但出于其内容的广泛性和场景关联性，不能将情感信息自动纳入敏感个人信息给予特殊保护，而需要结合具体场景加以判断，这给情感信息的保护造成困难。

（二）生物反馈信息难以认定为生物识别信息

情感计算输入端的生物反馈信息是与生物识别信息相关的概念，但二者却并不等同。我国《个人信息保护法》虽未明确生物识别信息的意涵及其所包含的具体类型，但作为我国立法重要参考的欧盟《通用数据保护条例》第9条第1款明确指出，生物识别信息以信息主体的有效识别（uniquely identifying a natural person）为目的。换言之，生物识别信息主要存在于人工智能时代的技术之中，以个人身份的有效识别与验证为目的。〔54〕我国《信息安全技术个人信息安全规范》（GB/T 35273－2020）虽然作为软法不具有强制执行力，但其附录B中提到“个人生物识别信息包括个人基因、指纹、声纹、掌纹、耳廓、虹膜、面部识别特征等”，以此可推断出我国生物识别信息同样指具有个人专属性且能够辨识个人身份的信息。〔55〕

正如本文伊始对情感计算特殊性的说明，情感识别不以身份识别为前提，生物反馈信息的收集与处理是为了识别信息主体之情感状态，而非确认信息主体的身份。所以，生物反馈信息与生物识别信息在目的上存在差异，前者难以为后者直接包含。需要说明的是，虽然无法作为生物识别信息享有特别保护，但生物反馈信息仍有可能作为医疗健康信息而被纳入敏感个人信息的范畴，例如心动周期、骨骼肌活动、脑电波信息等。

（三）情感信息的特殊性减弱告知同意之实效

《互联网信息服务算法推荐管理规定》第16条与《个人信息保护法》第13至17条规定了情感计算技术使用者的告知义务，其中前者的告知事项包括算法的基本原理、目的意图和主要运行机制等，后者则涵盖了个人信息的处理目的、方式和个人信息的种类、保存期限。告知同意作为个人信息处理的核心法律规则意在保护信息主体的知情权与选择权，〔56〕实践中可以隐私政策或设置显著标识的方

〔53〕 参见周汉华：《平行还是交叉——个人信息保护与隐私权的关系》，载《中外法学》2021年第5期，第1176页。

〔54〕 See Andreas Häuselmann, *Fit for Purpose? Affective Computing Meets EU Data Protection Law*, 11 International Data Privacy Law 245, 250 (2021).

〔55〕 见前注〔38〕，江必新、郭峰书，第268页。

〔56〕 参见刘俊臣（全国人大常委会法制工作委员会副主任）：《关于〈中华人民共和国个人信息保护法（草案）〉的说明》，2020年10月13日在第十三届全国人民代表大会常务委员会第二十二次会议上。

式展开。例如上文英国伦敦皮卡迪利灯塔摄像头事件，即使面对的信息主体是不特定群体，仍然可以在街边醒目位置设立告知牌，说明该地摄像头具有探知情感的功能并以提升广告投放效果为目的。当人们不愿情感被探知时，便可以选择不经过这一区域或绕道而行，以此保护公众的选择自由。

然而，隐私政策中网络平台作为技术使用者可能会以“该技术处理的个人信息包括情感信息”的方式履行告知义务，但基于情感信息的特殊性，上述告知方式难以达到有效告知的效果。其一，如上文提到情感信息的内涵宽泛，告知处理情感信息的具体类型及其处理目的方能达到保护知情权的效果。例如，情感计算可能被应用于招聘面试的场景中，此时面试者有权提前知晓相关算法开展信息处理的目的，如通过观测皮肤温度的变化测试其抗压能力、焦虑程度抑或是情绪起伏等。在面试者充分知晓并同意的前提下应用情感计算，方能符合面试者的隐私期待并赋予技术使用者信息处理的合法性基础。其二，如果是情感信息在已采集的生物反馈信息的基础上分析生成的场景中，如果说生物反馈信息是“第一代数据（first generation）”，那么此时情感信息则为“第二代数据（second generation）”或被推测的个人信息（inferred personal data）。该情境下，情感计算的使用者应在采集第一代数据之时便充分告知该数据将被二次处理及其情感探测的具体目的，针对处理生物反馈信息的单独同意不应成为处理情感信息的概括授权。

五、构建以生物数据为媒介的差异化风险规制模式

从现行法检视可以看出，不管是需证明可识别性的个人信息保护，还是以存在个人相关性和损害事实为前提的隐私权保护，图 4 中 ABC 三个区域的法益保护皆存在现实难题。因此，为构建“可信赖情感智能（trustworthy EAI）”的浅滩环境，展开对这一技术的整体规制便显得尤为重要。在这一方面，欧盟走在了对情感计算算法治理的前列，其于 2021 年 4 月颁布了《人工智能法案（草案）》（以下简称《草案》）并提出对情感识别（emotion recognition）算法规制的整体设想。因此，我国可以欧盟经验为参考构建情感计算规制的“中国模式”。

（一）以差异化风险规制建构可信赖情感智能

2021 年 4 月，欧盟委员会颁布的《草案》中提出依据对社会与个人的安全、健康与基本权利造成损害的风险，对人工智能系统进行分类风险治理。风险类别可以具体分为“由于不可接受风险禁止的 AI 应用（prohibited AI applications due to the unacceptable risks）”“高风险 AI 应用（high-risk AI applications）”“有限风险的 AI 应用（limited risk AI applications）”和“极小风险的 AI 应用（minimal risk AI applications）”。若将情感计算技术置于该评判体系之中，[57]可以得出如下结论：（1）整体而言，该技术被归类于有限风险的 AI 应用之中，《草案》对其施加了更高的透明度要求。需要注意的是，该归类并非一成不变，需要考虑技术应用的相关情境而具体分析。（2）当该技术具有操纵的潜在目的时，该技术被禁止应用。（3）当该技术应用于关键基础设施、教育、就业、重要的私人或公共服务、法律实施、边境控制以及司法协助系统等领域之时，其被认定为高风险应用，在投入市场之前需要经过更为严格的检验。例如，进入市场前的合规性评估（pre-market conformity assessment）、标准和认证（standards and certification）、风险管理评估（risk management assessment）、数据治理（data governance）、透明度评估（transparency）和人工监督（human oversight）。《草案》对于情感计算技术风险的认定有助于动态、具体地评定特定情感计算应用的风险，有利于弥补现行法对于 ABC 区域法益

〔57〕《人工智能法案（草案）》中使用的是“情感识别”（emotion recognition）一词，笔者认为是立法者考虑到该技术发展目前仍主要在情绪感知阶段，对于后续分析与影响涉及较少，从而使用“识别”而非“计算”。

保护的不足，并为图4中D区域的法益保护创造了条件。

为充分应对情感计算中的操纵风险和保护信息隐私，法律有必要在技术发展的初期确立风险应对的整体策略，以促进技术向善。我国可参照欧盟《人工智能法案（草案）》的治理逻辑，不仅在算法分级分类安全管理制度中设立整体的风险等级门槛，还可按照行业领域予以分级管理。此外，为划定技术发展之边界，笔者建议禁止减损自治与自决的“不可接受之操纵”，允许以用户利益为导向的“合理劝导”。而“合理劝导”的意义在于促使用户做出于己而言更佳的选择，如司机可以选择在汽车内安装智能安全系统，进而可以在情绪愤怒时听到舒缓的音乐以帮助其安全驾驶。然而，基于人类固有的认知与思维局限，做出最佳决策并非人们的道德义务（moral obligation），因此为合理劝导而非强制，是告知同意这一前提下的自我选择。

（二）从生物识别信息保护转向生物数据规制

如上文所述，情感计算中的生物反馈信息以情感识别而非身份识别为目的，故难以作为生物识别信息进行保护。考虑到生物反馈信息既是情感计算的输入数据，也可能是推断出情感信息的基础数据，贯穿了该技术应用的全过程，如何在个人信息保护法尚未覆盖之时保护生物反馈信息所承载的信息隐私法益，则是法律需要面对的问题。

《草案》在发布之初采取了与《通用数据保护条例》相同的对“生物识别数据”的界定方法，即“基于特别技术处理自然人的相关身体、生理或行为特征而得出的个人数据，这种个人数据能够识别或确定自然人的独特标识，例如面部形象或指纹数据”。另外，《草案》规定除应用于执法领域的实时远程生物识别外，该类数据的其他应用须遵守《通用数据保护条例》第9条第1款的规定，即禁止以识别特定自然人为目的的生物识别数据处理。

随后，《草案》中“生物识别数据”的概念遭受了广泛的批评，〔58〕其原因在于技术规制法案直接搬用《通用数据保护条例》中建立在可识别性基础上的“生物识别数据”会严重缩小规制对象的范围，将不以识别为目的的数据排除在保护范围之外。一方面，当《草案》针对的是以识别为基础的生物识别数据时，那么上文图4所示的四种法益中只有AB落在了规制的范围之内。此外，当情感计算应用于执法之外的领域时，这类数据被“一刀切”地禁止收集和使用。事实却是，即使在可识别的情境中，情感计算仍然可以帮助自闭症儿童的情绪学习、舒缓司机情绪减少事故发生等，具有一定的正向功能，因此一概禁止AB两类数据的使用不利于技术的创新与发展。另一方面，CD两类情境所涉及的数据被遗漏于法律规则之外，而这类数据正在逐渐影响着社会对技术的信赖程度。因此，《草案》作为对人工智能技术建立基础规则的立法，其当时的规定是不全面的。

基于此，欧盟理事会于2021年11月发布了对于《人工智能法案（草案）》的修正版（presidency compromise text），提出删除前一版定义中的“能够识别或确定自然人的独特标识”，〔59〕这使得情感计算等技术中不以识别为目的但基于身体或行为特征采集的“生物反馈数据”被囊括在内。最终，《草案》中“生物数据”的范围要大于《通用数据保护条例》中“生物识别数据”的范围。对于如何解释这一不同，笔者认为《草案》旨在通过“权力限制”实现对技术的整体规制，因此即便是难以确定权利主体但存在不特定风险的D类法益仍应在该法案中得到保护。与之相对，《通用数据保护条例》意在保护以个人信息自决为

〔58〕 See Catherine Jasserand, Jan Czarnocki et al., *The EU'S Approach to AI: A Short Discussion of The Draft Artificial Intelligence Act from A Biometrics' Law Perspective*, EAB-Note on proposed EU AI Act (March 2022).

〔59〕 Council of the European Union, Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts-Presidency compromise text [2021/0106(COD)], published in November 2021, p. 36, https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf.

基础的个人信息受保护权，故可识别性是相关生物信息受到保护的前提条件，也是行使权利的必备要件。

我国同样会面临算法相关立法与个人信息保护法之间的衔接问题，欧盟《草案》的立法经验也为我国解释算法相关立法中“数据”的范围提供了参考。因为前者的立法目的是“权力规制”，而后者是“权利制衡”，所以可以在前者中建立不以识别为目的的“生物数据”之概念，后者则可坚持以身份识别为目的的“生物识别数据”的内涵，最终“生物识别数据”所无法涵盖的数据将在“生物数据”的概念下得以规范。如此，在情感计算的法律规制中，“生物数据”可作为规制之媒介在保护数据安全的同时辅助差异化风险规制模式的建构。

六、结　　语

情感计算作为科学技术，其始终应以工具性的方式作用于人们的日常生活且旨在让人们生活得更好。通过在信息隐私框架下的探讨可知，情感计算对个人信息权益与隐私权造成了不成比例的风险，尤其是情感操纵存在将人客体化的风险。法律需要保护人的情感隐私与自由，但法律在面对新技术时往往具有一定滞后性。从既有信息隐私理论与规范来看，情感计算引发对部分传统法律概念内涵的反思，如敏感个人信息是否应在明确列举的项目中增加情感信息，生物识别信息是否应在情感智能时代继续保持以身份识别为目的的内涵，隐私权能否由个体隐私扩展至群体隐私进而由群体享有。本文虽以信息隐私理论作为问题研究的框架，但关注点在于情感计算的输入与输出端信息能否以及在何种程度上能够在现有理论与立法下受到保护，因此本文的研究是不全面的。情感计算与信息隐私理论相遇，仍有一些问题值得深入研究，比如如何处理情感计算准确性批判与个人信息准确性原则、信息修正权之间的矛盾等。

Abstract　Affective computing is an intelligent technology that uses human emotions as the object of machine learning, embodying emotion recognition and analysis functions. Affective computing is not intended to identify the information subjects and aims to output emotional information by analyzing and processing the bio-feedback information or state of the input. In order to deal with the risk of emotional manipulation brought about by the improper acquisition and utilization of emotional information, it is necessary to discuss the risk response within the framework of information privacy theory. However, the current information privacy norms have deficiencies in the risk response. Emotional information is difficult to be regarded as sensitive information because of the diversity of specific types. Bio-feedback information is difficult to be protected as biometric information because of emotional recognition rather than identification. Informed consent as the basis for information processing is also challenged due to the particularity of emotional information. Given the EU's experience in regulating emotion recognition in AI Act Proposal, it is suggested to construct the concept of "biometric-based data" in algorithmic governance and use it as a medium to carry out hierarchical and sub-domain regulation of technical risks, so as to enhance trustworthy EAI.

Keywords　Affective Computing, Information Privacy, Emotional Data, Bio-Feedback Data, Trustworthy EAI

（责任编辑：雷槟硕）