



## 揭开情感理解的新篇章 ——ACII 2023 会议报道

Tomas Ariel D'Amelio<sup>1 2</sup> & Hongxia Xie<sup>3</sup>

<sup>1</sup> 布宜诺斯艾利斯大学物理系和跨学科及应用物理研究所 (INFINA-CONICET)

<sup>2</sup> 布宜诺斯艾利斯大学心理学系

<sup>3</sup> 台湾阳明交通大学电子学研究所

ACII 2023 (情感计算与智能交互) 会议于 2023 年 9 月 10-13 日在麻省理工学院媒体实验室举行, 该实验室也被公认是情感计算学科的诞生地。大会期间, 来自情感计算、人工智能 (AI) 和情感科学领域的研究人员和创新者汇聚一堂。本报告旨在提炼和阐明会议期间的关键对话、新兴技术以及伦理和方法论层面所面临的紧迫挑战。

### 跨学科协作促进情感理解

ACII 2023 会议生动地展示了人工智能方法与情感科学模型之间日益增强的相互依存关系。例如, 在为期半天的 EPiC 研讨会上, 来自多个学科的专家 (包括心理学家、计算机科学家、工程师和生理学家) 进行了跨学科对话, 他们的共同目标是利用生物识别数据促进对情感的理解。Stanislaw Saganowski 博士 (计算机科学家) 指出, 尽管所使用的理论很复杂, 但与研讨会相关的情绪预测挑战却没有给出明确的解决方案。挑战的结果进一步证明了利用生物识别数据预测情感的复杂性, 以及相关预测的内在挑战。Saganowski 提出了应对这一挑战的几种方法, 包括提高数据集质量和重新审视该领域, 以确定可能限制模型有效性的系统性因素。Maciej Behnke 博士 (心理学家) 对此表示赞同, 并强调计算机科学家需要与心理学家合作, 这样才能更全面地了解情感状态。Behnke 指出, 情感计算和情感科学领域的重大进展在很大程度上有赖于跨学科合作, 将技术知识与对人类情感的深刻理解融为一体。

在研讨会上, 心理学家 Lisa Barret 博士分享了自己对情感体验复杂性的独特观点。她介绍了两种截然不同的范式: 传统的类型学方法 (topological approach) 和创新的群体研究法 (population approach)。类型学范式是整个会议期间反复提及的研究法, 它将情感划分为固定的分类标准, 将每种“类型”视为标准, 将差异视为偏差。这种方法虽然广为人知, 但往往忽略了情感的细微差别。Barret 对这一范式提出了质疑, 并主张采用群体范式。这种研究法考虑到了情感体验的流动性, 并将各种差异视为情感谱系中不可分割的组成部分, 而不是异常现象。Barret 以

“愤怒 (anger)” 为例，论述了任何表达方式（从尖叫到隐秘的报复行为）都不是异常现象，而是定义这种情感的基本要素。Barret 强调以群体范式来理解自然、多样状态下的情感，这为我们提供了一个全新的视角：即情感不是僵化的、自然的类别，而是动态的、蕴含文化内涵的体验。Wu 等人<sup>[1]</sup>在其获奖论文《信念错配系数 (BMC)》(Belief Mismatch Coefficient) 中提到了相关概念。他们提出的 BMC 指标解决了情感体验和情感报告中的不确定性这一特殊问题。因此，该指标有助于指导情感计算和心理学模型对情感的动态性、模糊性和复杂性的探索。

## 英文大语言模型与情感理解

ACII 2023 会议最令人印象深刻的议题就是大语言模型 (large language models, LLMs) 在情感计算中所扮演的革命性角色。例如，会议上介绍的[2]中的研究表明，像 CLIP 这样的视觉语言模型为视觉情感分析领域打开了新局面。值得注意的是，在 FI 数据集（该数据集包含 23,308 张标有 8 个情感类别的图像）中，CLIP 零样本法 (zero-shot approach) 在情感分类的准确性方面优于目前最先进的模型。[3] 中提出的研究进一步强调，LLMs (即 GPT-4) 不仅在预测多样情境下的情感评估方面展现出了极高的准确性，同时还能与既定的评估理论保持一致。这表明，语言模型有助于我们理解目标相关性、情感强度和情感差异之间错综复杂的关系。[3] 中的研究还强调了 GPT-4 (著名的 GPT 语言模型的最新迭代) 在情感推理方面的强大功能。作者的研究表明，与它的前身 GPT-3.5 不同，最新迭代版本在预测情感评估方面表现出了令人印象深刻的准确性，例如 Milton Bradley 的战舰棋盘游戏场景，该游戏的基本评估结构类似抛硬币游戏。除此之外，GPT-4 还与既有的评估理论相吻合，凸显了其揭示人类情感内在原因的能力。除了情感识别之外，会议还强调了 LLMs 在共情回复 (empathic response) 方面的技能。在专项研究 [4] 中，研究者们对大语言模型，尤其是 ChatGPT 进行了严格评估。结果令人鼓舞：该模型在 91.7% 的情况下都能准确识别情感并做出适当的反应。尽管这一结果低于在没有阿斯伯格综合症/自闭症的参与者中观察到的平均水平，但这一发现对于在日常和临床环境中实现人工智能与人类在情感层面的智能交互来说大有可为。

上文所提及的研究凸显了语言模型在情感智能方面取得的重大进展。然而，这些研究和会议上的其他讨论也传递了一个清晰的信号：这一议题仍然需要认真考虑和细致研究，因为挑战依然存在，并且新问题也在不断涌现。在情感识别方面，LLM 仍然面临的挑战之一是将其他带有情感信号的模态融入大语言模型。Mohammed Soleymani 博士在会议上强调，与单独使用 LLM 的效果相比，多模态技术在现实世界中的改进往往微不足道。另一个问题是混杂因素 (confounding factor)。例如，会上分享的研究报告[5]指出，耳麦和麦克风对视频观察和理解有很大影响。研究还强调了偏见等伦理问题。研究[6]采用了一种新颖的研究方法，利用 LLMs 生成文本扰动 (text perturbation)，模拟基于欧洲语言共同参考框架 (CEFR) 特征的语言能力水平。研究结果表明，随着语言能力的下降，LLMs 识别对话中情感的能力也会下降。这说明 LLMs 可能存在语言能力偏见，表现为对非母语口音的误解和对母语非英语的情感暗示的误读。除了 LLMs 这一主题之外，会议还强调了公平和偏见问题。例如，[6][7] 的研究强调需要更具包容性的数据集，并呼吁通过多样性来确保公平的结果。这与 LLMs 和情感理解的情况不谋而

合，特别是因为大多数研究都是基于有限的数据库，例如来自大学环境的人口代表性对在具有不同或多样化文化和人口群体的人群中部署这些模型有影响。其他道德风险包括人类在不知情的情况下与聊天机器人或其他基于 LLMs 的计算机程序进行交互。因此，研究者们需要对伦理问题进行评估，以确保负责任和合乎伦理地使用 LLMs，并将潜在危害降至最低。随着法规的不断收紧（如欧盟提出的《人工智能法案》），这一点变得尤为重要，这可以说是由 LLMs 的普遍发展和技术飞跃所引发的。情感计算相关与会者在本次会议的一个教程中特别讨论了这一法案，包括其对该领域研究的潜在影响。

## 总结

在 ACII 2023 闭幕后的特别会议上，我们荣幸在麻省理工学院媒体实验室情感计算研究小组 Rosalind Picard 教授的办公室对她进行了独家专访。Picard 教授在采访中分享了独到的个人见解，呼应了有关情感计算领域现状与未来发展方向的反复出现的议题。最后，我们想援引她的一段话作为结尾：“情感智能并不仅仅是指学习如何识别面部表情或是设立如何回应特定人群的规则，而是要深刻理解语言所处的情境。”

## 参考文献

[1] Wu, J., Dang, T., Sethu, V., Ambikairajah, E. (2023). Belief Mismatch Coefficient (BMC): A Novel Interpretable Measure of Prediction Accuracy for Ambiguous Emotion States. In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

[2] Bustos, C., Civit, C., Du, B., Sole-Ribalta, A., & Lapedriza, A. (2023). On the use of Vision-Language models for Visual Sentiment Analysis: a study on CLIP. In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

[3] Tak, A. N., & Gratch, J. (2023). Is GPT a Computational Model of Emotion? In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

[4] Schaaff, K., Reinig, C., & Schlippe, T. (2023). Exploring ChatGPT's Empathic Abilities. In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

[5] Moreno, F., Alghowinem, S., Park, H. W., & Breazeal, C. (2023). Espresso-AI: An Explainable Video-Based Deep Learning Models for Depression Diagnosis. In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

[6] Tran, Q., Shpileuskaya, K., Zaunseder, E., Salg, J., Putzar, L., & Blankenburg, S. (2023).

Robustness Analysis uncovers Language Proficiency Bias in Emotion Recognition Systems. In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

[7] Verhoef, T., Fosch-Villaronga, E. (2023). Towards affective computing that works for everyone. In 2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE.

