



多语言多模态环境下事件意义建构与知识表示：面向东南亚事件分析大模型的计算语言学框架

侯诗濂¹

中山大学 / 广州南方学院外国语学院

【摘要】随着大语言模型和多模态人工智能技术的快速发展，事件分析研究正逐渐由传统的信息抽取任务向融合语义理解、知识表示和智能推理的事件智能方向演进。面对多语言、多模态和多源异构信息环境，现有研究主要关注事件识别、模态融合以及模型性能优化，对于事件意义形成机制及其统一知识表示问题关注不足。以东南亚多语言多模态事件分析需求为应用背景，本文从计算语言学视角出发，在系统功能语言学、框架语义学和多模态话语分析等理论基础上，提出事件意义建构（Event Meaning Construction）概念，探讨多语言多模态环境下事件意义的形成机制，并进一步构建多语言多模态事件知识表示框架（Multilingual Multimodal Event Knowledge Representation Framework, MME-KR Framework）。在此基础上，结合知识增强大模型的发展趋势，讨论事件智能的演化路径及其在复杂区域环境中的潜在应用。本文认为，事件分析研究应从结构化信息获取逐步转向意义建构、知识组织与智能推理，为知识增强大模型、复杂事件理解以及区域智能分析提供统一的理论基础和新的研究思路。

【关键词】事件意义建构；知识表示；多语言；多模态；计算语言学；事件智能；东南亚事件分析

1 侯诗濂，中山大学 / 广州南方学院外国语学院。研究方向为计算语言学。邮箱 shilianhou0224@gmail.com



Event Meaning Construction and Knowledge Representation in Multilingual and Multimodal Environments: A Computational Linguistics Framework for Large Models in Southeast Asian Event Analysis

Shilian Hou¹

Sun Yat-sen University/School of Foreign Languages, Nanfang College · Guangzhou

【 Abstract 】 With the rapid development of large language models and multimodal artificial intelligence, event analysis has gradually evolved from traditional information extraction toward event intelligence, which integrates semantic understanding, knowledge representation, and intelligent reasoning. In multilingual and multimodal environments characterized by heterogeneous information sources, existing studies mainly focus on event extraction, modality fusion, and model performance optimization, while paying insufficient attention to the mechanisms of event meaning construction and unified knowledge representation. Taking Southeast Asian multilingual and multimodal event analysis as an application context, this paper adopts a computational linguistics perspective and draws upon Systemic Functional Linguistics, Frame Semantics, and Multimodal Discourse Analysis to propose the concept of Event Meaning Construction. It further explores the mechanisms underlying meaning construction in multilingual and multimodal environments and introduces a Multilingual Multimodal Event Knowledge Representation Framework (MME-KR Framework). Building upon this framework, the paper discusses the evolutionary path of event intelligence and its potential implications for knowledge-enhanced large models in complex regional contexts. It is argued that event analysis should move beyond structural information extraction toward meaning construction, knowledge organization, and intelligent reasoning, thereby providing a unified theoretical foundation for knowledge-enhanced large models, complex event understanding, and regional intelligent analysis.

【 Keywords 】 Event Meaning Construction; Knowledge Representation; Multilinguality; Multimodality; Computational Linguistics; Event Intelligence; Southeast Asian Event Analysis

¹ Shilian Hou, Sun Yat-sen University/School of Foreign Languages, Nanfang College · Guangzhou
Research interest: Computational Linguistics. Email: shilianhou0224@gmail.com



引言

随着生成式人工智能和大语言模型技术的快速发展，事件分析逐渐从传统的信息抽取任务扩展为融合语义理解、知识表示和推理能力的智能分析任务。面对多语言、多模态和多源异构数据环境，如何实现事件的统一语义表示与知识组织，已成为自然语言处理、知识工程和人工智能领域的重要研究问题。

近年来，多语言预训练模型、多模态大模型以及知识增强技术的发展显著推动了事件分析研究的进步。然而，现有研究主要关注事件识别、信息融合和模型性能优化，普遍将事件视为触发词、论元或实体关系的组合，对于事件意义如何形成、不同语言 and 不同模态如何共同参与意义建构以及事件知识如何支持复杂推理等问题关注不足。这在一定程度上限制了大模型在复杂事件理解、态势感知和自动报告生成等任务中的能力提升。

从计算语言学视角来看，事件不仅是客观世界中发生的行为或状态，更是语言、图像、视频及社会文化语境共同作用下形成的意义建构过程。系统功能语言学、框架语义学以及多模态话语分析等理论为理解事件语义和构建统一知识表示体系提供了新的研究思路。

基于此，本文以东南亚多语言多模态事件分析为应用背景，在梳理相关研究进展的基础上，从计算语言学视角探讨多语言多模态环境下事件意义建构问题，进一步提出多语言多模态事件知识表示框架，以期为知识增强大模型、事件推理和区域智能分析研究提供新的理论基础和发展路径。

一、多语言多模态事件分析的发展背景与研究需求

1.1 大模型时代事件智能的发展

近年来，以大语言模型（Large Language Models, LLMs）和多模态大模型（Multimodal Large Language Models, MLLMs）为代表的生成式人工智能技术快速发展，推动自然语言处理研究从传统的信息处理阶段逐步迈向知识理解与智能推理阶段。事件分析（Event Analysis）作为自然语言处理的重要研究方向，也正在经历从事件抽取（Event Extraction）向事件智能（Event Intelligence）的转变。

传统事件分析主要关注事件触发词识别、论元抽取以及事件分类等任务，其目标在于从非结构化文本中获取结构化事件信息。随着知识图谱、预训练模型以及大语言模型的发展，事件分析逐渐扩展到事件知识表示、事件关系发现、事件演化分析以及复杂事件推理等更高层次的任务。与此同时，多模态大模型的发展使得事件信息处理对象不再局限于文



本,而是逐步覆盖图像、视频、语音等多种信息载体,从而推动事件分析研究进入多语言、多模态和知识增强的新阶段 (Brown et al., 2020)。

在这一背景下,事件分析系统所承担的任务已经不再只是识别“发生了什么”,而是进一步理解“事件意味着什么”“不同信息之间存在何种关联”以及“事件未来可能如何发展”。事件分析逐渐由单纯的信息抽取问题转变为融合语义理解、知识组织和推理能力的事件智能问题。因此,如何构建统一的事件表示体系,实现复杂环境下事件知识的组织与推理,已经成为当前人工智能和计算语言学领域的重要研究方向。

1.2 东南亚区域多语言多模态事件分析需求

东南亚地区具有复杂的语言生态和多元文化特征,是典型的多语言、多模态信息传播环境。该地区涉及泰语、越南语、印尼语、马来语、菲律宾语、缅甸语等多种语言,同时广泛存在英语与本地语言混合使用现象,不同国家媒体对于同一事件的叙事方式和表达策略也存在明显差异。

随着社交媒体平台和移动互联网的快速发展,东南亚地区事件信息传播呈现出明显的多模态特征。新闻报道、图片、短视频、直播以及社交媒体评论共同构成复杂的信息传播网络,使得事件意义的表达不再依赖单一文本,而是通过不同模态资源共同完成。图像提供视觉证据,视频记录事件过程,文本负责叙事组织,音频则承载情绪和态度信息,不同模态之间相互补充,共同参与事件意义的生成过程。

然而,相较于英语和汉语等高资源语言,东南亚语言普遍面临标注资源不足、平行语料稀缺以及跨语言语义表示不统一等问题。同时,混合语种表达、口语化表达以及社交媒体非规范语言现象进一步增加了事件分析的复杂性。这些特点使得东南亚地区成为研究多语言多模态事件分析的重要场景,也对跨语言语义对齐、多模态语义融合以及知识增强推理提出了新的挑战。

从应用层面来看,区域研究、国际传播、风险预警以及态势研判等领域均对东南亚多语言多模态事件分析能力提出了迫切需求。如何实现复杂信息环境下事件知识的统一组织和智能分析,已经成为推动区域智能研究发展的关键问题。

1.3 从事件抽取到事件意义建构

长期以来,事件分析研究主要建立在事件抽取框架基础之上,将事件表示为触发词 (Trigger) 及其论元 (Argument) 的组合,通过识别事件参与者、时间、地点以及相关关系实现事件结构化表示。这种方法在信息抽取和知识获取方面取得了显著成果,但其核心仍然是基于形式特征和统计关联的结构表示。



随着多语言、多模态环境下事件分析需求的不断增长，传统事件抽取框架逐渐暴露出一定局限性。首先，不同语言对同一事件往往采用不同的表达方式和叙事结构，仅依赖词汇和句法特征难以实现深层语义一致性建模；其次，不同模态之间承担着不同的意义功能，简单的信息融合难以揭示事件意义的形成机制；再次，现有事件表示方法更多关注事件组成要素，而对于事件背后的态度、立场、评价以及社会文化语境等深层意义关注不足。

随着事件分析任务逐渐从结构化信息获取扩展到复杂语义理解，事件研究的关注重点也由事件要素识别逐步转向事件意义形成机制。事件不再被简单视为若干实体和关系的集合，而是不同符号资源和社会语境共同作用下形成的动态语义单元。事件意义既包含事件本身的经验内容，也涉及态度表达、话语组织以及文化背景等多个层面。因此，多语言多模态环境下事件分析的核心问题不仅在于识别事件，更在于理解事件意义如何形成以及不同语言 and 不同模态如何共同参与意义建构。

基于这一认识，本文认为，事件分析研究需要从传统的“事件抽取范式”逐步转向“事件意义建构范式”，在系统功能语言学、框架语义学以及多模态话语分析等理论基础上，探索多语言多模态环境下事件意义形成机制，并进一步建立统一的事件知识表示体系，为知识增强大模型、复杂事件推理以及区域智能分析提供新的理论基础。

二、多语言多模态事件分析研究现状

2.1 多语言事件分析研究

事件分析一直是自然语言处理领域的重要研究方向。早期研究主要基于 ACE、ERE 等数据集，通过事件触发词识别和论元抽取实现结构化事件表示（Doddington et al., 2004）。随着深度学习技术的发展，基于预训练模型的事件抽取方法逐渐成为主流，并在事件分类、事件关系识别和事件演化分析等任务中取得了显著进展（Wang et al., 2020）。

近年来，多语言预训练模型的发展进一步推动了跨语言事件分析研究。mBERT、XLM-R 以及 NLLB 等模型在跨语言表示学习方面取得了较好的效果，使得不同语言之间的知识迁移和语义对齐成为可能。同时，跨语言事件抽取和低资源语言处理逐渐成为研究热点，研究者尝试利用多语言共享表示、迁移学习和知识蒸馏等方法提高低资源语言环境下的事件分析能力。

然而，多语言事件分析研究仍然主要建立在词汇、句法以及统计特征基础之上，其核心目标在于实现不同语言之间的信息映射和知识迁移，而对于不同语言如何构建事件意义

以及不同文化语境下事件表达差异的关注相对有限。特别是在复杂区域环境下，同一事件往往会因语言、文化和媒体立场的不同而呈现出不同的话语形式，这使得跨语言事件分析不仅是语言转换问题，更涉及深层语义和意义建构问题。

2.2 多模态事件理解研究

随着社交媒体和数字媒体的发展，事件信息逐渐由单一文本扩展到图像、视频和语音等多种模态。多模态事件理解研究因此成为近年来人工智能领域的重要方向。

早期研究主要关注图文匹配和视觉语义表示，通过联合嵌入空间实现文本与图像之间的信息关联。随后，CLIP、BLIP 以及视觉语言模型的发展进一步提升了跨模态语义理解能力，多模态大模型的出现则推动了图像、视频和文本的统一建模，使得复杂场景下的事件识别和描述成为可能。

现有多模态事件分析研究主要围绕信息融合和模态对齐展开，通过视觉特征、文本特征以及语音特征之间的联合学习实现事件识别和事件分类。然而，大多数研究将不同模态视为信息来源的简单组合，更关注模态间的一致性匹配，而对于不同模态在事件意义建构过程中所承担的不同功能关注不足。

事实上，文本、图像和视频并非简单地重复表达同一信息，而是以不同方式共同参与事件意义的生成过程。因此，多模态事件分析不仅需要解决模态对齐问题，更需要解释不同模态如何协同构建统一的事件语义空间。

2.3 事件知识表示与大模型研究

为了提升事件分析系统的推理能力，研究者逐渐将知识表示方法引入事件分析领域。事件知识图谱（Event Knowledge Graph）和时序知识图谱（Temporal Knowledge Graph）通过构建事件实体、关系以及时间演化链，实现了复杂事件的结构化组织（Gottschalk & Demidova, 2018）。事件关系发现、因果推理以及事件预测等研究进一步扩展了事件知识表示的应用范围。

与此同时，大语言模型的发展为事件分析带来了新的机遇。基于检索增强生成（Retrieval-Augmented Generation）和知识增强技术的大模型能够利用外部知识资源完成复杂事件理解和推理任务（Lewis et al., 2020）。GraphRAG、Agent 等方法的提出，使得知识图谱和大模型之间的结合成为当前研究的重要方向（Edge et al., 2024）。

然而，无论是事件知识图谱还是知识增强大模型，其核心仍然依赖已有知识结构和统计关联模式。当前事件知识表示方法主要关注事件之间的实体关系和逻辑关系，而对于事件意义的组织机制缺乏统一描述。大模型虽然具备较强的生成能力和推理能力，但其知识



表示过程仍然缺乏明确的语义层支撑，这在一定程度上限制了复杂事件理解和可解释推理能力的发展。

2.4 现有研究的局限与挑战

总体来看，多语言事件分析、多模态事件理解、事件知识图谱以及知识增强大模型等研究方向已经取得了显著进展，但现有研究仍然存在若干共同问题。

首先，现有事件表示方法主要关注结构特征和统计关联，对于事件意义本身缺乏统一描述。无论是触发词—论元框架、事件知识图谱还是大模型表示，其核心仍然建立在形式层和知识层基础之上，而对于意义层的研究相对不足。

其次，多语言研究更多关注语义迁移和跨语言映射，而缺少对不同语言事件表达机制及其话语差异的深入探讨。不同文化背景和媒体环境下的事件叙事方式往往存在显著差异，仅依赖表层语义对齐难以实现真正意义上的跨语言事件理解。

再次，多模态研究主要关注模态融合和特征匹配，而缺乏关于不同模态如何共同参与意义建构的理论解释。图像、视频、文本和语音在事件表达过程中承担不同功能，但现有研究尚未形成统一的多模态意义表示框架。

最后，大模型虽然显著提升了事件分析能力，但其知识组织和推理过程仍然缺乏可解释的语义基础。如何建立统一的事件意义空间，实现多语言、多模态和知识推理之间的有机融合，仍然是当前事件分析研究面临的重要挑战。

因此，有必要从计算语言学视角重新审视事件分析问题，探索多语言多模态环境下事件意义的形成机制，并进一步构建统一的事件知识表示框架，从而为复杂事件理解和知识增强大模型研究提供新的理论基础。

三、多语言多模态环境下事件意义建构机制

3.1 事件意义建构的理论基础

从传统自然语言处理视角来看，事件通常被定义为发生于特定时间和空间中的行为、状态或变化，并通过事件触发词及其论元实现结构化表示。然而，这种定义更多强调事件的形式特征和结构关系，而对于事件意义的形成机制关注相对有限。

从意义生成的角度来看，事件并非预先存在的静态对象，而是在语言表达、多模态资源以及认知语境的共同作用下不断被建构和重构的语义实体。不同语言、不同模态以及不同文化背景对于同一事件可能采用不同的表达方式，但人类仍然能够识别其背后的统一意



义，这表明事件理解不仅依赖形式匹配，还涉及更深层次的意义建构过程。

系统功能语言学认为，语言的本质是意义潜势（meaning potential），语言通过经验意义、人际意义和语篇意义共同实现对客观世界的建构（Halliday & Matthiessen, 2014）。其中，经验意义关注事件过程、参与者以及环境成分，人际意义体现说话者的态度和立场，语篇意义则负责信息组织和意义连贯。这一理论为理解事件的多层次意义提供了重要基础。

框架语义学进一步强调，词汇意义来源于其所激活的概念框架，事件理解依赖于参与者、行为、时间和空间等要素之间形成的认知结构（Fillmore, 1982）。与此同时，多模态话语分析认为，图像、视频、语音和文本等不同符号资源共同参与意义生成过程，不同模态承担不同的语义功能，并通过协同作用实现完整意义表达（Kress & van Leeuwen, 2021）。

因此，事件意义不应被简单理解为词汇和句法结构的组合，而应被视为语言符号、多模态资源和社会文化语境共同作用下形成的动态意义建构过程。基于这一认识，本文提出事件意义建构（Event Meaning Construction）概念，将事件理解从传统结构表示扩展到多层次意义表示，为多语言多模态事件分析提供新的理论视角。

3.2 多语言多模态环境下事件意义的构成

在多语言多模态环境下，事件意义具有复杂的层次结构。不同语言、不同模态以及不同语境共同决定事件意义的形成方式。基于系统功能语言学和多模态话语分析理论，本文认为事件意义主要由以下五个层面构成。

（1）经验意义（Experiential Meaning）

经验意义描述事件本身所包含的行为、状态和变化过程，是事件意义建构的基础。它主要涉及事件过程、参与者以及时间、地点等环境信息，对应传统事件分析中的事件结构表示。

（2）人际意义（Interpersonal Meaning）

事件不仅描述客观事实，还包含说话者、媒体或社会群体的态度、立场和评价（Martin & White, 2005）。同一事件在不同国家媒体中的报道往往体现不同的价值取向和叙事策略。因此，人际意义构成事件理解的重要组成部分，也是复杂事件分析和立场分析的重要基础。

（3）语篇意义（Textual Meaning）

事件信息并非孤立存在，而是通过特定的信息组织方式形成连贯的叙事结构。不



同语言 and 不同媒体对于事件的叙述顺序、信息焦点和话语结构存在差异，从而影响事件意义的表达和传播。

(4) 多模态意义 (Multimodal Meaning)

在数字媒体环境下，事件意义不再由文本单独承担，而是由图像、视频、语音等多种模态共同完成。图像提供视觉证据，视频记录事件过程，语音体现情绪和态度，文本负责叙事组织，不同模态之间相互补充，共同构成完整事件意义 (O'Halloran, 2004)。

(5) 语境意义 (Contextual Meaning)

事件意义的形成还受到社会文化背景、历史经验以及传播环境的影响。同一事件在不同国家、不同文化和不同媒体环境下可能具有不同解释方式。因此，语境意义决定了事件理解的深层认知基础，也是实现跨语言语义一致性的重要条件。

上述五个层面相互关联、共同作用，构成多语言多模态环境下事件意义的整体结构。

3.3 多语言多模态事件意义空间

基于上述分析，本文进一步提出多语言多模态事件意义空间 (Multilingual Multimodal Event Meaning Space) 概念。

与传统事件表示方法不同，事件意义空间不再将事件视为触发词、实体关系或模态特征的简单组合，而是强调不同语言、不同模态和不同语境之间的协同作用，将事件理解建立在统一意义层基础之上。

在这一意义空间中，语言层负责事件叙事和语义表达，多模态层提供视觉和听觉证据，知识层负责事件关系组织，语境层提供文化背景和认知基础，各层之间通过动态交互共同完成事件意义建构。

事件意义空间不仅能够实现跨语言语义对齐和多模态语义融合，还能够为事件知识表示、知识增强模型以及复杂事件推理提供统一的语义基础。

因此，从事件抽取到事件意义建构的转变，本质上体现了事件分析研究范式的演进。未来事件分析的发展不应局限于结构化信息获取，而应进一步关注事件意义形成机制及其知识组织过程，从而推动事件分析逐步走向事件智能 (Event Intelligence)。

四、多语言多模态事件知识表示框架

4.1 从事件意义空间到事件知识表示

事件意义空间为多语言、多模态环境下事件理解提供了统一的语义基础，但意义本身仍然具有抽象性和动态性。为了进一步支持复杂事件推理、知识组织和智能分析，需要将事件意义转化为可计算、可组织和可扩展的知识表示形式。

如果说事件意义空间回答的是“事件意义如何形成”的问题，那么事件知识表示则进一步关注“事件意义如何被组织、存储和利用”的问题。通过构建事件图谱实现事件知识组织。然而，这类表示方法更多关注事件之间的逻辑关系和结构关系，而对于事件所包含的深层语义信息、话语特征以及多模态信息之间的协同关系缺乏统一描述。

在多语言多模态环境下，事件知识不仅来源于文本，还来源于图像、视频、语音以及社会文化语境等多种信息资源。因此，事件知识表示不应局限于实体和关系的静态描述，而应建立在事件意义空间基础之上，实现不同语言、不同模态和不同知识之间的统一组织。

基于这一认识，本文认为，多语言多模态事件知识表示的核心任务在于建立从意义层到知识层的映射机制，使事件意义能够转化为结构化知识，并进一步支持知识推理和智能分析。

4.2 多语言多模态事件知识表示框架（MME-KR Framework）

基于事件意义建构机制，本文提出多语言多模态事件知识表示框架（Multilingual Multimodal Event Knowledge Representation Framework, MME-KR Framework）。

该框架由数据层、语义层、模态层、知识层、推理层和应用层六个部分组成：

（1）数据层（Data Layer）

数据层是事件知识表示的基础，主要包括新闻报道、社交媒体文本、图片、视频、音频以及政府公告、研究报告等多源异构数据。

对于东南亚区域而言，数据层不仅涉及泰语、越南语、印尼语、马来语、缅甸语等多种语言，还广泛存在混合语种表达和非规范语言现象，因此需要构建统一的数据组织体系，为事件分析提供数据基础。

（2）语义层（Semantic Layer）

语义层负责实现事件意义建构，包括经验意义、人际意义、语篇意义、多模态意义和语境意义等多个层面。

在这一层中，不同语言表达被映射到统一事件意义空间，实现跨语言语义对齐和意义一致性建模，从而解决不同语言表达方式差异带来的语义鸿沟问题。



（3）模态层（Multimodal Layer）

模态层主要负责不同模态资源之间的协同表示。文本提供叙事结构和语义内容，图像提供视觉证据，视频记录事件过程，语音反映情绪和态度，不同模态共同参与事件意义建构。

模态层的目标并非简单的信息融合，而是建立统一的多模态语义空间，实现不同模态之间的意义互补和协同表达。

（4）知识层（Knowledge Layer）

知识层负责事件知识组织和结构化表示。在统一事件意义空间基础上，事件参与者、时间、地点、组织、关系以及事件演化过程被进一步组织为事件知识网络，实现复杂事件的结构化表达。

知识层不仅包含实体关系，还包括事件之间的时间关系、因果关系和演化关系，从而形成动态事件知识体系。

（5）推理层（Reasoning Layer）

推理层负责利用已有知识完成复杂事件理解和知识推理。

通过知识增强、事件关联分析以及动态知识更新，实现事件关系发现、事件链构建、趋势分析以及复杂事件推理，为事件智能提供知识支撑。

（6）应用层（Application Layer）

应用层面向区域研究、态势感知、风险预警和自动报告生成等实际需求，实现事件知识的智能利用。

通过统一事件知识表示，可以进一步支持知识增强大模型、智能问答、事件预测以及区域态势分析等任务，推动事件分析从信息获取走向智能决策支持。

4.3 面向知识增强大模型的事件智能演化路径

大语言模型的发展为事件分析提供了新的技术基础，但事件智能的发展并不仅仅依赖模型规模的扩大，而更加依赖知识表示能力和推理能力的提升。

从事件分析的发展过程来看，其研究范式经历了从规则方法到深度学习、从事件抽取到知识表示、再到知识增强大模型的发展过程。

未来事件智能的发展将呈现如下演化路径：

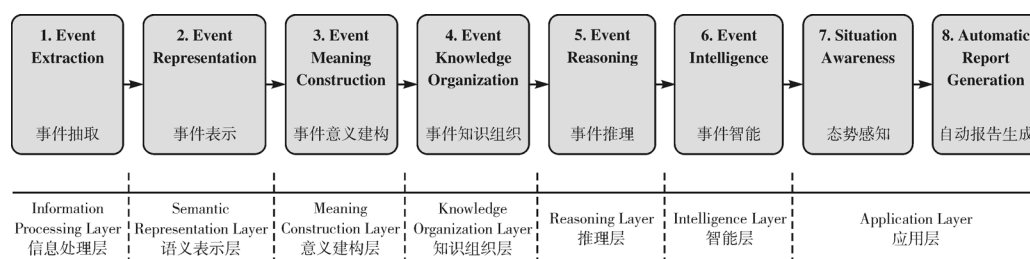


图 1. 多语言多模态环境下事件智能的发展演化路径

在这一过程中，多语言多模态事件知识表示框架将成为连接事件意义建构与知识增强大模型的重要桥梁，为复杂事件理解、区域态势分析以及智能决策支持提供统一的理论基础。

因此，从计算语言学视角出发建立统一事件知识表示体系，不仅有助于提升大模型的知识组织能力和可解释性，也将推动事件分析研究从统计关联走向知识驱动，从信息抽取走向事件智能。

五、面向东南亚事件分析大模型的发展路径

5.1 多语言多模态事件知识资源建设

统一的事件知识资源是事件智能发展的基础。对于东南亚区域而言，复杂的语言生态、多样化的信息来源以及快速变化的信息传播环境，使得传统单语言、单模态数据资源难以满足复杂事件分析需求。

未来研究需要面向泰语、越南语、印尼语、马来语、缅甸语等语言构建统一的多语言事件资源体系，实现新闻文本、社交媒体、图片、视频和音频等多模态数据的协同组织。在此基础上，进一步建立事件实体、事件关系、事件演化链以及跨语言事件映射机制，形成动态更新的多语言多模态事件知识资源。

与此同时，东南亚地区广泛存在英语与本地语言混合使用、社交媒体非规范表达以及不同文化背景下的话语差异等现象，因此需要进一步关注混合语种环境下的事件语义一致性问题，探索面向复杂语言环境的事件知识组织方法。

从长期发展来看，多语言多模态事件知识资源将成为知识增强大模型和区域事件智能的重要基础，也是推动低资源语言智能发展的关键支撑。

5.2 知识增强事件智能模型

大语言模型的发展为复杂事件理解提供了新的技术基础，但模型规模的增加并不意味



着事件智能能力的自然提升。复杂事件分析不仅依赖语言生成能力，更依赖知识组织能力和推理能力。

未来事件智能模型的发展需要建立在统一事件知识表示基础之上，通过外部知识资源、事件知识图谱以及动态知识更新机制，实现知识增强和推理能力提升。

在多语言多模态环境下，事件知识表示、知识检索、知识推理以及生成能力之间将逐渐形成协同关系，使大模型能够在复杂语境中实现事件关联分析、趋势预测和深层语义理解。

同时，随着知识图谱、检索增强生成以及智能体技术的发展，事件分析系统将逐渐由单一模型驱动转向知识驱动和推理驱动。未来事件智能模型不仅需要理解“发生了什么”，还需要进一步回答“为什么发生”“可能产生什么影响”以及“未来如何演化”等问题，从而实现从信息获取到知识发现的转变。

5.3 从事件分析走向事件智能

从事件抽取到事件智能的发展过程，本质上反映了人工智能研究范式从信息处理向知识理解和认知推理的演进。

传统事件分析主要关注事件识别和结构化表示，其目标是从非结构化数据中提取事件信息。随着知识表示和大模型技术的发展，事件分析逐渐扩展到事件关系发现、事件演化分析和复杂推理等更高层次任务。未来事件分析将进一步突破单一任务边界，形成面向复杂场景的事件智能体系。

在这一过程中，事件意义建构将成为连接语言理解和知识推理的重要纽带，统一事件知识表示则为知识增强大模型提供语义基础。通过多语言、多模态和知识推理能力的深度融合，事件分析有望逐渐发展为具备态势感知、风险评估、趋势预测和自动报告生成能力的综合智能系统。

对于东南亚区域研究而言，多语言多模态事件分析大模型不仅是技术发展的具体应用场景，也为复杂环境下事件智能研究提供了重要实验平台。通过区域场景驱动，可以进一步推动事件意义建构理论、多语言多模态知识表示以及知识增强推理方法的发展，从而促进事件分析研究从传统的信息抽取范式逐步走向事件智能范式。

因此，从计算语言学视角出发，探索多语言多模态环境下事件意义建构与知识表示问题，不仅能够为知识增强大模型提供新的理论基础，也有望推动事件智能研究形成新的发展方向，为区域研究、态势感知以及智能决策支持提供更加可靠的知识支撑。



结语

随着大语言模型和多模态人工智能技术的快速发展，事件分析研究正在经历由信息抽取向知识理解和智能推理的深刻转变。面对复杂的多语言、多模态信息环境，传统基于触发词—论元结构或实体关系表示的事件分析方法已经难以满足复杂事件理解和知识推理的需求。如何建立统一的事件语义基础，实现不同语言、不同模态以及不同知识之间的协同组织，成为当前事件分析研究亟待解决的重要问题。

从计算语言学视角来看，事件分析的核心问题并非单纯的信息提取，而是复杂环境下意义生成、知识组织和智能推理的统一过程。事件智能的发展需要建立在统一的语义基础之上，从而实现多语言、多模态和知识推理之间的深度融合。基于这一认识，本文在系统功能语言学、框架语义学以及多模态话语分析等理论基础上，提出事件意义建构（Event Meaning Construction）概念，探讨多语言多模态环境下事件意义的形成机制，并进一步构建多语言多模态事件知识表示框架（MME-KR Framework），尝试建立连接语言理解、知识表示和智能推理的统一理论基础。

与现有研究主要关注事件识别、模态融合以及模型性能优化不同，本文强调事件分析研究应当从结构表示逐步走向意义建构，从单一模态处理走向多模态协同，从统计关联学习走向知识组织与推理。在此基础上，结合东南亚区域多语言多模态事件分析需求，进一步讨论知识增强大模型的发展路径，为复杂事件理解、态势感知以及自动报告生成提供新的研究思路。

本文提出的事件意义建构和多语言多模态事件知识表示框架仍然属于初步理论探索，相关概念体系、表示方法以及推理机制尚有待进一步完善。未来研究可在多语言事件语料资源建设、多模态事件知识图谱构建、知识增强大模型、复杂事件推理以及事件智能体等方向开展深入研究，逐步形成面向复杂信息环境的事件智能理论体系。

总体而言，从“事件抽取（Event Extraction）”走向“事件意义建构（Event Meaning Construction）”，再进一步迈向“事件智能（Event Intelligence）”，不仅反映了事件分析研究范式的演进，也体现了人工智能研究从信息处理向知识理解和认知推理发展的总体趋势。如何在多语言、多模态和复杂知识环境下实现事件意义的统一表达与智能利用，将成为未来计算语言学、自然语言处理以及人工智能交叉研究的重要方向。



【参考文献】

- 胡壮麟. (2017). 《系统功能语言学概论》(第三版) [*An introduction to systemic functional linguistics*]. 北京大学出版社 [Peking University Press].
- 冯志伟. (2016). 《应用语言学综论》 [*A comprehensive introduction to applied linguistics*]. 商务印书馆 [The Commercial Press].
- 刘海涛. (2023). 大语言模型背景下语言学研究的新时代 [Opportunities for linguistic research in the era of large language models]. 当代语言学 [Contemporary Linguistics], 25(4), 481–492.
- 梁茂成. (2022). 数字人文视域下语料库语言学的发展 [The development of corpus linguistics from the perspective of digital humanities]. 现代外语 [Modern Foreign Languages], 45(5), 577–588.
- 肖仰华, 张俊林, 等. (2020). 知识图谱: 概念与技术 [Knowledge graph: Concepts and technologies]. 中国科学: 信息科学 [Scientia Sinica Informationis], 50(2), 293–316. <https://doi.org/10.1360/N112019-00249>
- 张德禄. (2009). 多模态话语分析综合理论框架探索 [Exploring an integrated theoretical framework for multimodal discourse analysis]. 中国外语 [Foreign Languages in China], 6(1), 24–30.
- 张德禄. (2018). 多模态话语分析理论与应用研究 [Theory and applications of multimodal discourse analysis]. 高等教育出版社 [Higher Education Press].
- Brown, T. B., Mann, B., Ryder, N., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8440–8451). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.747>
- Costa-jussà, M. R., et al. (2022). No language left behind: Scaling human-centered machine translation. *arXiv preprint arXiv:2207.04672*. <https://doi.org/10.48550/arXiv.2207.04672>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019* (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Doddington, G., Mitchell, A., Przybocki, M., Ramshaw, L., Strassel, S., & Weischedel, R. (2004). The Automatic Content Extraction (ACE) program—Tasks, data, and evaluation. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)* (pp. 837–840).



- Edge, D., Trinh, H., Cheng, N., et al. (2024). From local to global: A graph RAG approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*.
<https://doi.org/10.48550/arXiv.2404.16130>
- Fillmore, C. J. (1982). Frame semantics. In Linguistic Society of Korea (Ed.), *Linguistics in the morning calm* (pp. 111–137). Hanshin Publishing.
- Gottschalk, S., & Demidova, E. (2018). EventKG: A multilingual event-centric temporal knowledge graph. In *The Semantic Web: ESWC 2018* (pp. 272–287).
https://doi.org/10.1007/978-3-319-93417-4_18
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.
- Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.
- Li, J., Li, D., Savarese, S., & Hoi, S. (2022). BLIP: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *Proceedings of the 39th International Conference on Machine Learning (ICML 2022)* (pp. 12888–12900).
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. Continuum.
- Radford, A., Kim, J. W., Hallacy, C., et al. (2021). Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning (ICML 2021)* (pp. 8748–8763).
- Wang, L., Ma, Y., Zhang, Q., et al. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), Article 186345.
<https://doi.org/10.1007/s11704-024-40231-1>
- Wang, X., Wang, Z., Han, X., Jiang, W., Han, J., Liu, Z., Sun, M., & Li, P. (2020). MAVEN: A massive general domain event detection dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1652–1671). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2020.emnlp-main.130>