命名实体识别研究综述

徐嘉艺

（中国地质大学（武汉），经济管理学院，武汉 430074）

摘　要：命名实体识别是信息抽取、自然语言处理、文本挖掘领域中的重要研究部分。随着互联网的发展和大数据时代的到来，如何从海量的文本数据中抽取有价值的信息变得愈发重要，命名实体识别从而也迎来了新的发展。本文梳理了命名实体识别领域从萌发之初发展至今的研究过程，总结了该领域主要技术方法的相关研究，分析了该领域目前的研究热点及研究难点，最后提出了关于该领域的未来发展的思考。

关键词：命名实体识别；文本挖掘；自然语言处理；深度学习

Overview of Named Entity Recognition Technology

Xu Jiayi

(School of Economics and Management,China University of Geosciences,Wuhan 430074)

Abstract: Named entity recognition is an important research part in information extraction, natural language processing and text mining. With the development of the Internet and the arrival of the era of big data, how to extract valuable information from massive text data has become increasingly important, and named entity recognition has ushered in a new development. This paper summarizes the development process of named entity recognition from the beginning to the present, summarizes the related research on the main technologies and methods in this field, analyzes the current research hot spots and difficulties in this field, and finally puts forward some thoughts on the future development of this field.

Key words: named entity recognition; text mining; natural language processing; deep learning

目录

1 引言

命名实体识别（Name Entity Recognition，NER）的概念于 MUC-6（the Sixth Message Understanding Conference）第一次被提出[1]，是人工智能领域中自然语言处理（Natural Language Processing，NLP）领域中的一个子任务，其主要作用为从文本中识别出实体（例如组织、人员、地点等专有名词）以及含有特殊意义的时间、货币等数量短语并将其归类[2]。举例说明：“美国商会（US Chamber of Commerce）会长苏珊·克拉克在当地时间 1 月 11 日举行的“美国商业状况”年度演讲中称，竞争对美国的未来至关重要，包括“应对中国挑战”。”这句话中包含的实体有：日期实体“1 月 11 日”、组织机构实体“美国商会”、人名实体“苏珊·克拉克”。

命名实体识别技术最早主要被用于识别某些特殊名词（实体、时间、数量词）[2]，随着研究的推进，学术领域开始关注对于开放域（open domain）的信息抽取研究，命名实体识别的任务不再局限于一般的实体分类，越来越多的实体类型被提出，研究学者们对其进行了更详细的任务划分：除了一般的实体识别之外，还可以针对某些特定领域（例如金融、生物、电影领域等等）进行专门的信息抽取。

随着互联网技术的发展以及大数据时代的来临，海量非结构化的互联网文本数据中蕴含着大量有价值信息，命名实体技术是进行文本挖掘与文本分析的核心技术，可以对数据进行有效的信息查找以及信息抽取。命名实体识别技术由最开始的基于规则和字典的方法逐渐发展为统计学方法以及目前基于深度学习的方法。其在自然语言处理领域中有着广泛的应用，例如构建知识图谱（Knowledge Graph）[3]、机器翻译（Machine Translation）[4]、网络搜索（Web Search）[5]等。

2 命名实体的定义

国外对于命名实体识别的研究开展较早，1991 年，Rau 在第七届 IEEE 人工智能应用会议上发表了关于“从文本中抽取公司名称”的文章[8]。1995 年，MUC-6（the Sixth Message Understanding Coference）首次指出将命名实体（Named Entity，NE）作为研究对象[1]，MUC-7 指出命名实体识别的任务是从文本中识别出人名、组织名称、地理位置名称以及时间、货币和百分比表达式等信息[2]。ACE（Automatics Content Etraction）在之后则对机构名与地名进行的细化，增加了地理-政治实体、设施[6]和交通工具、武器[7]的实体类型。CoNLL-2002 和 CoNLL2003 会议基于原有的 MUC 定义，指出实体类型的划分除了人名、组织名与机构名之外还应包括其他命名实体[9-10]。

除上文提到的会议之外，也有研究学者发表了对命名实体定义的观点。Constantine 等[11]提出命名实体就是专有名词（Proper Noun，PN）。Alfonseca 等[12]认为命名实体代

表我们用于解决特定问题的研究对象。Sekine 等[13-14]希望 NER 能够广泛涵盖更多的应用，将实体类别扩展到了 200 个。Borrega 等[15]规定了命名实体必须为名词和名词短语且不应当产生歧义，提出了强命名实体（Strong Named Entities，SNE）和弱命名实体（Weak Named Entities，WNE）的概念。Nadeau 等[16]指出命名实体代表严格指示词（rigid designators）。严格指示词是指，若对于一个对象 *x*，在所有存在的世界中，指示词 *d* 均表示 *x*，不存在指示词 d 所表示的其他对象，即 *x* 的指示词 *d* 是严格的。最后，Marrero 等[17]在前人的基础上，将命名实体总结为语法类别、严格指示、唯一标识和应用目的四种类别。作者首先假设每种类别都能被用作为命名实体的定义标准，然后使用分析、举例等方式否定其作为标准的可行性。最后得出：应用方面的需求目的，是定义命名实体唯一可行的标准。

3 命名实体识别技术的研究历史

关于命名实体识别技术的研究，最早开始于 1991 年 Rau 在第七届 IEEE 人工智能应用会议上发表的关于“从文本中抽取公司名称”的文章[8]。Rau 主要使用了启发式和手工编写规则的算法进行信息抽取。基于英语语法的特点，识别英文文本中的命名实体往往不需要考虑复杂的分词操作。而对于中文文本的命名实体识别，在处理文本之前必须经过词法、句法分析，所以中文处理的难度要高于英文。关于国外命名实体识别技术的研究，1999 年 Bikel 等[18]提出了基于隐马尔可夫模型的识别技术。基于该方法的识别技术在 MUC-6 的测试文本集中的测试结果达到了：97%（地名识别）、94%（机构名识别）、95%（人名识别）。2009 年 Liao 等[19]提出了基于条件随机场的识别模型，使用半监督学习的方法；Ratinov 等[20]提出的未标注文本训练词类模型的算法，在 CoNLL-2003 的文本集中 F1 值达到了 90.8%的准确度。

针对中文文本的命名实体识别技术开展较晚，从 20 世纪 90 年代初期开始，国内一些学者对中文命名实体（例如：人名、地名、机构名等）识别进行了一些研究。1995 年孙茂松等[21]提出了中文姓名的识别方法，其主要采用统计学的方法计算姓氏和人名用字概率。张小衡等[22]提出了对中文机构名称进行识别与分析的方法，其采用了编写人工规则的方法对高校名进行了实验研究。Zhang 等[23]在 ACI2000 上演示了其针对中文命名实体及实体间关系开发的信息抽取系统，该系统使用基于记忆的学习（Memory Based Learning，MBL ）的算法获取规则，用以抽取命名实体及实体间的关系。2004 年 Tsai 等[24]提出 基于最大嫡的混合的方法；2007 年冯元勇等[25]提出基于单字提示特征的中文命名实体识别快速算法；2008 年郑逢强等[26]将《知网》中的义原作为特征加入到最大嫡模型中，以此来训练产生性能更好的模型。

4 命名实体识别研究现状

近年来国内外关于命名实体识别技术的研究主要包括基于统计机器学习的方法和深度学习的方法。

基于机器学习的方法一般是通过标注好的文本进行训练，利用训练好的模型进行识别[27]。其中，常用的模型有隐马尔可夫模型[28] 、最大熵模型、决策树、支持向量机等。基于深度学习的识别方法在近年来运用广泛，Yao 等[29]提出了一种基于 CNN 的适合医学文本内容的训练的命名实体识别方法，这种方法不需要构建词典，并且同时可以保证较高的准确率。Strubell 等[30] 提出了迭代扩张卷积神经网络（iterated dilated convolutional neural networks，IDCNN）命名实体识别的方法，与下文提到的目前最具有表现力的 LSTM 模型相比，该模型只需要 O（N）的时间复杂度，在保持与 LSTM 相当的精度的条件下，可提升八倍的速度。Yang 等人[31]分别采用字符级 CNN 和词级别 CNN 的的方式进行命名实体识别，在字符级 CNN 中使用单层 CNN，词级采用多层 CNN，最后利用 Softmax 或者条件随机场的方式实现实体的标注。Kong 等[32]提出了一种完全基于 CNN 的模型，充分利用 GPU 并行性来提高模型效率，模型中构造多级 CNN 来捕获短期和长期上下文信息，在保证较高识别准确率的情况下大幅提高了效率。

循环神经网络循环神经网络（Recurrent neural network，RNN）也可以用于命名实体识别，RNN 的变体 LSTM 在命名实体识别方面取得了显著的成就。Huang 等人[33]将双向长短期记忆网络（bi-directional long -short-term memory，BiLSTM）-CRF 应用于自然语言处理基准序列标记数据集。Zhang 等[34]提出了针对中文 NER 的 Lattice LSTM 模型。与基于字符的方法相比，该模型显式地利用了词序列信息，达到了最佳结果。Han 等[35]针对专业领域内命名实体识别通常面临领域内标注数据缺乏的问题，将生成对抗网络与长短期记忆网络模型相结合，在各项指标上显著优于其他模型。近年来，基于深度学习的命名实体识别研究除了基于卷积神经网络和循环神经网络的方法外，还出现了一些更新的技术。首先是 Transformer 模型[36]，它不再使用传统的神经网络思想，使用到的只有注意力机制[35]。2018 年，BERT 模型被提出[37]，在命名实体识别领域，Dai 等[38]在中文电子病历表识别的应用上使用了 BERT+ BiLISM+CRF 的网络结构，取得了很好的效果，Li[39]等人使用了多层变种网络结构进行中文临床 命名实体识别，同样取得了很好的识别效果。文献[40]中利用预训练的 BERT 模型结合 BiLSTM，提高了在微博中文数据集上命名实体识别的准确率。Li 等[41] 针对现有的 Lattice LSTM 结构复杂的问题，提出了 FLAT，在性能和效率上均有提升。Yoon 等[42]提出一个新型模型，由多个双向 LSTM 网络构成，每个网络可作为一个单独的任务识别某一种制定的实体类型，多个任务将各自学习到的知识进行转移，获得更准确的预测。

5 命名实体识别的评测会议

5.1 MUC（Message Understanding Conference）评测

1995 年举办的第六届 MUC 会议使得命名实体识别成为一项明确且重要的研究任务，对 NER 的发展具有重大意义。MUC 会议是一系列面向信息抽取研究的会议。其特点是以评测的形式定义会议主题，参会者必须是评测比赛的参赛者。评测的比分使用信息检索领域常用的正确率（Precision）、召回率（Recall）和 F1 值（F1 score）。该会议的这种评测的形式后来被广泛使用在各类信息检索和自然语言处理会议当中。

MUC-6 中定义的实体较为广泛，包括命名实体（ENAMEX）、时间表达式（TIMEX）和数量表达式（NUMEX）[1]。其中命名实体又分为人名、地名和机构名。由于语料规模较小（30 篇文档），类型单一（均是新闻）， 因此参会的方法大多都取得了较好的识别效果，F1 值最高达到 96.42%，而且人名的识别效果要明显好于地名和机构名。由于采用的是华尔街日报语料，任务只是对英语句中命名实体的识别。因此 1996 年春进一步举办了 MET（the multilingual entity task），将命名实体识别的任务扩展到汉语、西班牙语和日语。然而其中汉语的识别效果相对较低，F1 值最高只有 84.51%。1998 年的 MUC-7[2]和 MET-2，继承并修订了 MUC-6 的标注规范，增加了训练语料的规模，然而此次评测的英文识别效果并不如 MUC-6，F1 值最高达到 93.39%，中文的识别效果却有了明显的提高，F1 值达到 86%。

MUC-6 的研究均采用了基于规则的方法，如词法规则，包括词法规则、词性规则、短语规则等。大多数方法都是根据命名实体前后的提示词和上下文来制定字符序列匹配规则。MUC-7 的大部分研究都是基于规则的，就像 MET-2 的中文命名实体识别一样。这些方法组成了一组有限的规则和模式，然后自动从文本中找到匹配这些规则或模式的字符串，并将它们标记为各种命名实体。 基于规则的方法特别适合于识别时间和数量的表达式。但是，由于命名实体的构造规则是多变的，所以用有限的规则来识别几乎无限的命名实体是不合适的。基于规则的方法非常特定于领域，尽管它们在会议度量中表现良好，但它们不适用于一般领域中的文本。在 MUC-7 中，出现了一些基于统计机器学习方法的初步尝试，如最大熵 ME 和隐马尔可夫 HMM。

5.2 ACE（Auyomatic Content Extraction）项目

ACE 项目从 1999 年开始举办，目的是发展从人类语言中抽取信息的自动内容抽取技术，处理的信息形式不限于文本，还包括了音频和图像。该项目的语言数据和标注规范由语言 LDC（the Linguistic Data Consortium）[43]提供。ACE 研究的是目标对象（例如：实体、关系和事件）而不仅仅是 MUC 文本中的单词。不同的是，MUC 侧重于识别实体的名称，而 ACE 侧重于具有这些名称的实体。

ACE 分为多个阶段，第一阶段（1999-2001）主要关注实体检测与跟踪（entity detection and tracking，EDT），该阶段实际上包含四个子任务：实体识别、实体属性识别、实体参考识别和参考内容识别，后两个子任务是实体跟踪。在实体识别方面，ACE 将实体分为五类：人名、机构名、地缘政治实体、地名、设施[6]。地缘政治实体和设施实际上是地名和机构名称的延伸，这部分实体识别也不包括在 MUC 中。

ACE 第二阶段（2002-2003）在 EDT[44]中加入转喻，并将研究内容扩展到真实体相关性[45]的关系检测与表征（RDC）。ACE-2003 进一步增加了汉语和阿拉伯语的研究。

ACE 第三阶段（2004 - 2008）进一步拓展了研究领域。 Ace-2004 开始包括主要任务：实体识别（EDT）、关系识别（RDR）和事件识别（VDR）[46]。 两种实体类别被添加到 EDT：车辆和武器。此外，实体链接跟踪（LNK）已经被添加到三种语言。ACE-2005 增加了对价值和时间表达式的识别，并进一步定义了事件（events）[47]。到目前为止，ACE 有五个主要任务[48]。

ACE-2007 增加了西班牙语[49]，并增加了一个新的任务：实体翻译[50-51]，而其他任务与 ACE-2005[52]相比变化不大。ACE-2008 将研究语言限制在英语和阿拉伯语，并提出除了局部（文档内）实体识别之外，还要进行全局（跨文档）实体识别[53-54]。

5.3 CoNLL（Conference on Computational Natural Language Learning）会议

CoNLL （Computational Natural Language - Language Learning）是由 ACL 的自然语言理解特别兴趣小组（Natural Language understanding group，SIGNLL）组织的年度学术会议，自 1997 年以来一直在举办，并于 1999 年成立了类似于 MUC 的共享任务。 评估任务的主题每年都不同，其中 CoNLL-2002 和 CoNLL-2003 的主题是独立于语言的命名实体识别。也就是找到一种方法，在不同的语言中有效地识别命名实体，这比 MUC 要困难得多。CoNLL-2002 和 CoNLL-2003 定义的命名实体包括人名、地名、机构名、时间和数量，其中 CoNLL-2002 的语料库为西班牙语和荷兰语，而 CoNLL-2003 的语料库为英语和德语。

两届会议对命名实体识别的研究与中央大学的研究有明显的不同。大多数参与的方法使用统计机器学习方法。如隐马尔可夫 HMM[55]、最大熵 ME[56-57]、支持向量机 SVM[58]、条件随机场 CRF[59]、AdaBoost[60]等。一些连接主义模型包括 Winnow 方法[61]，投票感知机[62]，和短期和长期记忆网络 LSTM[63] 也试了一下。在基于规则的方法中，仍然只使用转换学习 TBL 方法，但效果并不理想[64]。多种方法的整合也是当时研究的一个重要尝试[65-67]。Carreras 等人[58]使用 Adaboost.MH 方法在 CoNLL-2002 的西班牙语和荷兰语的 F1 值分别为 81.39 和 77.05，识别效果最好。而在在 CoNLL -2003 中，最好的方法都采用或集成了最大熵模型，在英语和德语中排名第一的[66]的 F1 值分别为 88.76 和 72.41。

关于 CoNLL 对命名实体识别的研究，统计机器学习已经成为主流，各种重要的机器学习方法也被尝试过。虽然当时的语料库规模并不大，但机器学习仍然表现出较强的性能。其中，统计机器学习是唯一可行且有效的独立于语言的命名实体识别方法。此时，选择和改进合适的机器学习方法，选择更合适的文本表示特征已成为命名实体识别研究的主要思路和趋势。另外，CoNLL 是基于双语的，因此在面对双语的情况下，机器学习方法的研究难度明显增加，特别是机器学习方法非常依赖语料库中的特征表示。没有双语语料库的大规模支持，不可避免地很难获得一个效果良好的通用双语 NER 系统。因此，长期以来，NER 的研究多集中在单一语言上，很少涉及双语甚至多语。近年来大数据环境的推动，双语、多语语料库规模不断扩大，跨语言 NER 研究才迸发萌芽。

5.4 SIGHAN（ the special interest group for chinese language processing）Bakeoff

中文命名实体识别比英文更为复杂和困难，这主要体现在中文文本中缺少表示词边界的分隔符号。 命名实体的识别效果很大程度上受到自动分词结果的影响[68]，而中文自动分词效果也经常受到命名实体识别的影响。在中文信息处理中，类似命名实体识别的研究首次出现，以提高中文自动分词的效果[69-70]。早期中国人命名实体识别主要集中在某一类命名实体上，如人名[71]、组织名[72]，多采用基于规则的方法[73-75]。

中国命名实体识别在 2003 年的“863 评价”中首次提出，并一直持续到 2005 年。汉语自动分词是 2003 年和 2004 年才出现的一个子任务。评价结果中 F1 值最高的仅为 82.38%[76]。SIGHAN Bakeoff-2006 开始把 NER 研究作为一个重要的研究领域，并组织大规模的评估会议。

SIGHAN 也是 ACL 的一个特殊兴趣组，主要研究中文自动分词。在汉语分词中，未注册词（OOV）是影响分词效果的一个非常重要的因素，而命名实体是影响分词效果最显著的一个因素。因此，命名实体识别是汉语自动分词中不可避免的问题。2006 年，SIGHAN 正式将 NER 问题作为其评估竞赛（bakeoff）的一项任务。Bakeoff-2006 提供了三套汉语语料库（MSRA、LDC 和 CITYU），并参照 conl -2002 的体系，定义了四种类型的命名实体：个人名称、地名、机构名称和地缘政治实体（GPE）。通过 Bakeoff-2007，减去 LDC 语料库，将命名实体简化为最常见的三个类别：人名、地名和组织名。在 Bakeoff 的评价中，统计机器学习方法仍占主导地位，总体上取得了较好的效果。在 Bakeoff-2006 和 Bakeoff-2007 的评价中，大多数领先的 NER 系统采用 CRF 模型[84-89]和 ME 模型。Bakeoff 的评价有几个重要发现：语料库中未知词的比例直接影响识别效果；在三种命名实体中，组织名称识别是最困难的。在训练语料库中加入外部数据对识别效果有很大影响。

SIGHAN Bakeoff-2010 没有继续关注命名实体识别，而是关注与命名实体相关的另一个重要问题：命名实体消歧（NED）[84]。在 SIGHAN Bakeoff 2012 中，命名实体识

别和消歧（NERD）作为一个全新的问题，成为本次 Bakeoff 的评估任务[85]。大多数参与的 NERD 系统分为 NER 和 NED，其中 CRF 模型仍是 NER 的首选[86-87]。从那时起，SIGHAN Bakeoff 就不再使用 NER 作为评估任务。

6 数据集与标注方法

6.1 公开数据集

常用的命名实体识别数据集有 CoNLL 2003，CoNLL 2002，ACE 2004，ACE 2005 等。数据集的具体介绍如下：

（1）CoNLL 2003 数据集[90]包括 1393 篇英语新闻文章和 909 篇德语新闻文章，英语语料库是免费的，德国语料库需要收费。英语语料取自路透社收集的共享任务数据集。数据集中标注了 4 种实体类型 ：PER，LOC，ORG，MISC。

（2）CoNLL 2002 数据集[91]是从西班牙 EFE 新闻机构收集的西班牙共享任务数据集。数据集标注了 4 种实体类型 ：PER，LOC，ORG，MISC。

（3）ACE 2004 多语种训练语料库[92]版权属于语言数据联盟（Linguistic Data Consortium，LDC），ACE 2004 多语言培训语料库包含用于 2004 年自动内容提取（ACE）技术评估的全套英语、阿拉伯语和中文培训数据。语言集由为实体和关系标注的各种类型的数据组成。

（4）ACE 2005 多语种训练语料库[92]版权属于 LDC，包含完整的英语、阿拉伯语和汉语训练数据，数据来源包括：微博、广播新闻、新闻组、广播对话等，可以用来做实体、关系、事件抽取等任务。

（5） OntoNotes 5.0 数据集[93]版权属于 LDC，由 1745 K 英语、900 K 中文和 300 K 阿拉伯语文本数 据组成，OntoNotes 5.0 的数据来源也多种多样，来自 电话对话、新闻通讯社、广播新闻、广播对话和博客 等。实体被标注为 PERSON，ORGANIZATION，LOCATION 等 18 个类型。

（6）MUC 7 数据集[94]是发布的可以用于命名实体识别任务，版权属于 LDC，下载需要支付一定费用。数据取自北美新闻文本语料库的新闻标题，其中包含 190 K 训练集、64 K 测试集。

（7）Twitter 数据集是由 Zhang 等[95]提供，数据收集于 Twitter，训练集包含了 4000 推特文章，3257 条推特用户测试。该数据集不仅包含文本信息还包含了图片信息。

对于中文领域来说，三类基本实体的数据来源多为评测会议数据集，多由新闻文本组成，如表 1 所示。

表 1 中文领域的公开数据集及实体类型总结

Table 1 Summary of public datasets and entity types in the Chinese domain

公开数据集	实体类型
BOSON 数据集	时间、地点、人名、组织、公司、产品
1998 人民日报数据集	人名、地名、组织名
MSRA 数据集	地点、机构、人物
SIGHAN 2005-MSR	日期、百分数、数字、人名、地名等12类
SIGHAN 2005-PKU	人名、地名、组织名
OntoNotes	Vehicle、Person、Organization、Weapon 等7大类及 45 个二级子类
Resume	中文简历等字段
Weibo	人名、地名、组织名
CLUENER	地址、书名、公司、电影、游戏等10类

6.2 标注方法

针对不同的数据集可能采用不同的标注方法，在此介绍几种常见的标注方法：

（1）BIO 标注法，是 CoNLL 2003 采用的标注法，I 表示内部，O 表示外部，B 表示开始。如若语料中某个词标注 B/I-XXX，B/I 表示这个词属于命名实体的开始或内部，XXX 表示命名实体的类型。O 表示属于命名实体的外部，即它不是一个命名实体。

中 B-school 国 I-school 地 I-school 质 I-school
大 I-school 学 I-school

图 1 BIO 标注规范
Fig.1 BIO annotation specification

（2）BIOES 标注法是 BIO 方法的扩展，其中 B 表示这个词处于一个命名实体的开始，I 表示内部，O 表示外部，E 表示这个词处于一个实体的结束，S 表示这个词是单独形成一个命名实体。BIOES 是目前最通用的命名实体标注方法。

王 S-person 好 O 战 S-event ， O
请 O 以 O 战 S-event 喻 O

图 2 BIOES 标注规范
Fig.2 BIOES annotation specification

（3）Markup 标注法,是 OntoNotes 数据集使用的标注方法，它直接用标签把命名实体标注出来，然后通过 TYPE 字段设置相应的类型。

ENAMEX TYPE=”school”>中国地质大学 ENAMEX>
位于
ENAMEX TYPE=”location”>湖北省武汉市 ENAME>

图 3 Markup 标注规范
Fig.3 Markup annotation specification

7 命名实体识别的评价指标

目前，命名实体识别任务常采用的评价指标有 精确率（Precision）、召回率（Recall）、F1 值（F1-Measure）等。

7.1 精确率

精确率是指对给定数据集，分类正确样本个数和总样本数的比值。

$$Percision = \frac{TP + TN}{TP + FN + FP + TN}$$

式中，TP 指将正预测为真，FN 指将正预测为假，FP 指将反预测为真，TN 指将反预测为假。

7.2 召回率

召回率是指分类器中判定为真的正例占总正例的比率。

$$Recall = \frac{TP}{TP + FN}$$

7.3 F1 值

F1 值是指精确率和召回率的调和平均指标，是平衡准确率和召回率影响的综合指标。

$$\frac{1}{F} = \frac{1}{Recall} + \frac{1}{Precision}$$

8 命名实体识别主要技术方法

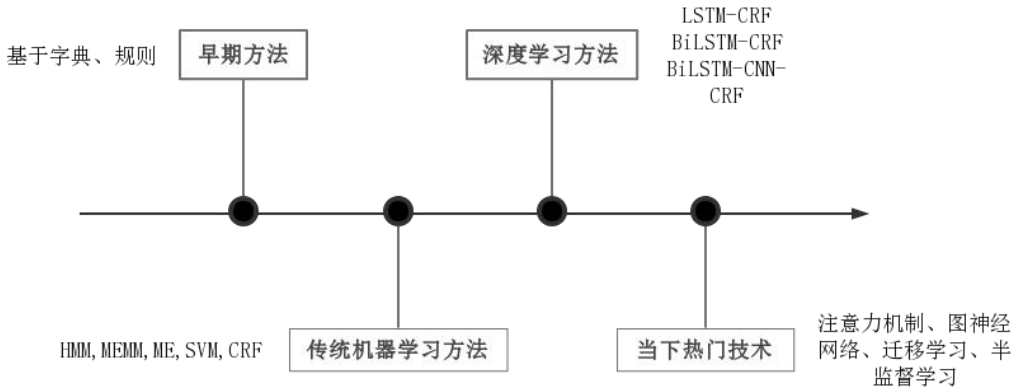


图 4 命名实体识别技术研究发展趋势

Fig.4 NER technology research development trend

8.1 基于规则和字典的模式匹配方法

模式匹配方法应用最早，也被称作 NER 专家系统方法（Expert System，ES）。ES 要求包含专业最高水平知识，提取专家知识并将其转换为规则形式。基于规则和字典的方法是最初代的命名实体识别使用的方法，这些方法多采用由语言学家通过人工方式，依据数据集特征构建的特定规则模板或者特殊词典。规则包括关键词、位置词、方位词、中心词、指示词、统计信息、标点符号等。词典是由特征词构成的词典和外部词典共同组成，外部词典指已有的常识词典。制定好规则和词典后，通常使用匹配的方式对文本进行处理以实现命名实体识别。Rau 等[8]首次提出将人工编写的规则与启发式想法相结合的方法，实现了从文本中自动抽取公司名称类型的命名实体。这种基于规则的方法的局限性是显而易见的。它不仅消耗大量的人力，而且不容易在其他实体类型或数据集中扩展，也不能适应数据的变化。

模式匹配方法包括：

（1）构造大而完整的词典。例如，针对民族名的特点，部分学者构建了维吾尔语 NER[96]的维吾尔语名数据字典。如果文本中的任何一个实体没有被包含在字典中，它将被手动记录在字典中，以便下次识别。

（2）在字典的基础上，根据实体提取添加实体构建规则。典型规则包括关键词、位置词、中心词等元素。例如，NER[97]的中文翻译采用了普通人姓名的形成规律——全名如[姓+名]，姓如[姓+职位]，[老（小）+姓]来识别；化学物质 NER 采用化学物质的组成模式——化学介词+化学词头+化学符号[98]，使用正则表达式提取化学物质的名称。

基于模块化匹配方法的 NLP 系统，如 Sheffield University NLP[99]开发的 NLP 框架 GATE，具有明确的 NER 规范。GATE 下的 JAPE 组件是一种特定于 GATE 的模式匹配语言。它的语法类似于正则表达式，其构造规则由文本中实体的特征决定。结构规则的不同也会产生冲突，如[武汉长江大桥]可分为[武汉市长|长江大桥]或[武汉长江大桥|中]真实主体，目前这两种语法规范主要采用基于前向匹配或后向匹配或[100]结合这两种算法来解决这些冲突。也可以参考英语词干算法[101]的原始原理。将实体出现的频率作为实体分割优先级的依据，但并不灵活。

模式匹配方法具有较高的准确性，但许多实体识别规则的制定依赖于领域专家，且领域之间基本没有重用。此外，领域词典需要定期维护，新实体的不断出现和它们的不规则性使构建一个完整的词典变得困难。虽然也有不足之处，但仍采用模式匹配的方法，因为某些领域的实体规则可被耗尽 95%，规则仍然是提取判断文档部分实体的首选。同时，在机器学习和深度学习 NER 模型中加入规则和词汇可以提高准确率。

8.2 基于传统机器学习的方法

在基于机器学习的方法中，命名实体识别被当作是序列标注问题。与分类问题相比，

序列标注问题中当前的预测标签不仅与当前的输入特征相关，还与之前的预测标签相关，即预测标签序列之间是 有强相互依赖关系的。采用的传统机器学习方法主要包括：隐马尔可夫模型（Hidden Markov Model，HMM）、最大熵（Maximum Entropy，ME）、最大熵马尔可夫模型（Maximum Entropy Markov Model，MEMM）、支持向量机 Support Vector Machine，SVM）、条件随机场（Conditional Random Fields，CRF）等。

其中，ME 结构紧凑，普适性较好，其缺点主要是训练时间复杂性高，甚至导致训练代价难以承受，另外由于需要明确的归一化计算，导致开销比较大。HMM 对转移概率和表现概率直接建模，统计共现概率。ME 和 SVM 在正确率上要比 HMM 高一些，但是 HMM 在训练和识别时的速度要快一些。MEMM 对转移概率和表现概率建立联合概率，统计条件概率，但由于只在局部做归一化容易陷入局部最优。CRF 模型统计全局概率，在归一化时考虑数据在全局的分布，而不是仅仅在局部进行归一化，因此解决了 MEMM 中标记偏置的问题。在传统机器学习中，CRF 被看作是命名实体识别的主流模型，优点在于在对一个位置进行标注的过程中 CRF 可以利用内部及上下文特征信息。还有学者通过调整方法的精确率和召回率对传统机器学习进行改进。Culotta 和 McCallum[102]计算从 CRF 模型提取的短语的置信度得分，将这些得分用于对实体识别进行排序和过滤。Carpenter[103]从 HMM 计算短语级别的条件概率，并尝试通过降低这些概率的阈值来增加对命名实体识别的召回率。对给定训练好的 CRF 模型，Minkov 等[104]通过微调特征的权重来判断是否是命名实体，更改权重可能会奖励或惩罚 CRF 解码过程中的实体识别。

8.3 基于深度学习的方法

随着深度学习的不断发展，命名实体识别的研究重点已转向深层神经网络 （Deep Neural Network，DNN），该技术几乎不需要特征工程和领域知识[105-107]。Collobert 等[108]首次提出基于神经网络的命名实体识别方法，该方法中每个单词具有固定大小的窗口，但未能考虑长距离单词之间的有效信息。为了克服这一限制，Chiu 和 Nichols 提出了一种双向 LSTM-CNNs 架构，该架构可自动检测单词和字符级别的特征。Ma 和 Hovy[110]进一步将 其扩展到 BiLSTM-CNNs-CRF 体系结构，其中添加了 CRF 模块以优化输出标签序列。Liu 等[111]提出了一种称为 LM-LSTM-CRF 的任务感知型神经语言模型，将字符感知型神经语言模型合并到一个多任务框架下，以提取字符级向量化表示。这些端到端模型具备从数据中自动学习的功能，可以很好地识别新实体。

部分学者将辅助信息和深度学习方法混合使用进行命名实体识别。Liu 等[112]在混合半马尔可夫条件随机场（Hybrid Semi-Markov Conditional Random Fields，HSCRFs）的体系结构的基础上加入了 Gazetteers 地名词典，利用实体在地名词典的匹配 结果作为命名实体识别的特征之一。一些研究尝试在标签级别跨数据集共享信息，Greenberg [113]

提出了一个单一的 CRF 模型，使用异构标签集进行命名实体识别，此方法对平衡标签分布的领域数据集有实用性。Augenstein 等[114]使用标签向量化表示在任务之间进一步播信息。Beryozkin 等[115]建议使用给定的标签层次结构共同学习一个在所有标签集中共享其标签层的神经网络，取得了非常优异的性能。

表 2 总结了自 1991 年至 2021 年来命名实体识别的方法对比。

表 2 命名实体识别方法对比[116]

Table 2 Comparison of named entity recognition methods

命名实体识别技术方法	年份	方法	数据集	模型	方法特点	评测指标	评测值/%
传统规则	1991		财经新闻	启发式算法+规则	较为准确地自动提取实体，但构造规则的方法会耗费大量的人力，可移植性很差。	Acc	97.5
传统机器学习	2004	HMM	GENIA V3.0	HMM+实体识别器	集成了构词模式，形态模式，词性，中心名词，特殊动词，别称 6 个特征，特征丰富，需要人工构造特征。	F-measure	66.6
	2011	CRF	Tweets	KNN 分类器+CRF	在半监督框架下进行实体识别，有效缓解训练数据匮乏的问题。人工构造特征复杂。	F-measure	80.2
深度学习	2011	CNN	CoNLL 2003	Conv-CRF	首次引入 CNN 进行实体识别，但丢失了长距离单词的有效信息。	F-measure	88.67
	2019		CCKS-2017	RD-CNN-CRF	将实体识别视为序列标注任务，利用残差膨胀卷积捕获上下文，有效提高训练效率。	F-measure	88.51
	2021		CCKS-2017	ALL CNN	构建多级 CNN+注意力机制捕获不同尺度的上下文信息，提高模型效率。	F-measure	90.49
	2015	RNN	CoNLL 2003	BI-LSTM-CRF	首次应用 BI-LSTM，捕获过去和未来的特征，但需要大量的特征工程。	F-measure	88.83
	2016		CoNLL 2003	LSTM-CNNs-CRF	不需要人工构造特征，将 BI-LSTM 与 CNN 结合到一起，完全端到端的模型。	F-measure	91.21
	2018		CoNLL 2003	并行 RNN 模型	采用多个独立的 BI-LSTM，大大减少参数量，提高训练效率。	F-measure	91.48
	2020		CoNLL 2003	CNN-BI-LSTM-CRF	研究单词和字符特征对实体识别的有效性，采用两层 BI-LSTM 减少输入序列以克服长输入序列难以预测的问题。	F-measure	91.10
	2019		CoNLL 2003	TENER	引入相对位置编码，可以分别在词级与字符级表示。	F-measure	91.52
	2019	Transformer	CoNLL 2003	BERT	采用 transformer-encoder 结构，可以深度挖掘上下文相关信息，但模型参数量大，训	F-measure	92.80

			练速率较慢。		
2021	CoNLL 2003	SelfAtt-BI-LSTM-CRF	引入自注意力机制更好地处理实体之间的长距离依赖关系。	Acc	90.47
2021	CoNLL 2017	MHA-BiLSTM-CRF	将中文字符特征与临床知识特征相结合，对医学临床文本更具有针对性。	F-measure	91.97

9 研究热点

陈曙东等[117]通过调研近三年来 ACL、AAAI、EMNLP、COLING、NAACL 等自然语言处理顶级会议中命名实体识别相关的论文，总结并选择了若干具有代表性的研究热点进行展开介绍,分别是匮乏资源命名实体识别、细粒度命名实体识别、嵌套命名实体识别、命名实体链接。

9.1 匮乏资源命名实体识别

命名实体识别通常需要大规模的标注数据集， 例如标记句子中的每个单词，这样才能很好地训练 模型。然而这种方法很难应用到标注数据少的领域，如生物、医学等领域。这是因为资源不足的情况下，模型无法充分学习隐藏的特征表示，传统的监督学习方法的性能会大大降低。

近来，越来越多的方法被提出用于解决低资源命名实体识别。一些学者采用迁移学习的方法，桥接富足资源和匮乏资源，命名实体识别的迁移学习方法可以分为两种：基于并行语料库的迁移学习和基于共享表示的迁移学习。利用并行语料库在高资源和低资源语言之间映射信息，Chen 和 Feng 等[118-119]提出同时识别和链接双语命名实体。Ni 和 Mayhew 等[120]创建了一个跨语言的命名实体识别系统，该系统通过将带注释的富足资源数据转换到匮乏资源上，很好地解决了匮乏资源问题。Zhou 等[121]采用双对抗网络探索高资源和低资源之间有效的特征融合，将对抗判别器和对抗训练集成在一个统一的框架中进行，实现了端到端的训练。

还有学者采用正样本—未标注样本学习方法（Positive-Unlabeled，PU），仅使用未标注数据和部分不完善的命名实体字典来实现命名实体识别任务。Yang 等学者[122]采用 AdaSampling 方法，它最初将所有未标记的实例视为负实例，不断地迭代训练模型，最终将所有未标注的实例划分到相应的正负实例集中。Peng 等学者[123]实现了 PU 学习方法在命名实体识别中的应用，仅使用未标记的数据集和不完备的命名实体字典来执行命名实体识别任务，该方法无偏且一致地估算任务损失，并大大减少对字典大小的要求。

因此，针对资源匮乏领域标注数据的缺乏问题，基于迁移学习、对抗学习、远监督学习等方法被充分利用，解决资源匮乏领域的命名实体识别难题，降低人工标注工作量，

也是最近研究的重点。

9.2 细粒度命名实体识别

为了智能地理解文本并提取大量信息，更精确地确定非结构化文本中提到的实体类型很有意义。通常这些实体类型在知识库的类型层次结构中可以形成类型路径[124]，知识库中的类型通常为层次结构的组织形式，即类型层次。

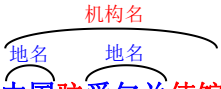
大多数命名实体识别研究都集中在有限的实体类型上，MUC-7[2]只考虑了 3 类：人名、地名和组织机构名，CoNLL-2003[10]增加了其他类，ACE[6]引入了地缘政治、武器、车辆和设施 4 类实体，Ontonotes 类型增加到 18 类，BBN 有 29 种实体类型。Ling 等[125]定义了一个细粒度的 112 个标签集，将标签问题表述为多类型多标签分类。

学者们在该领域已经进行了许多研究，通常学习每个实体的分布式表示，并应用多标签分类模型进行类型推断。Neelakantan 和 Chang [126]利用各种信息构造实体的特征表示。Yaghoobzadeh 等[127]重点关注实体的名 称和文本中的实体指代项，并为实体和类型对设计了两个评分模型。这些工作淡化了实体之间的内部关系，并单独为每个实体分配类型。Jin 等[128]以实体之间的内部关系为结构信息，构造实体图，进一步提出了一种网络嵌入框架学习实体之间的相关性。最近的研究表明以卷积方式同时包含节点特征和图结构信息，将实体特征丰富到图结构将获益颇多[129-130]。此外，还有学者考虑到由于大多数知识库都不完整，缺乏实体类型信息，例如在 DBpedia 数据库中 36.53%的实体没有类型信息。因此对于每个未标记的实体，Jin 等[131]充分利用其文本描述、类型 和属性来预测缺失的类型，将推断实体的细粒度类型问题转化成基于图的半监督分类问题，提出了使用分层多图卷积网络构造 3 种连通性矩阵，以捕获实体之间不同类型的语义相关性。

此外，实现知识库中命名实体的细粒度划分也是完善知识库的重要任务之一。细粒度命名实体识别现有方法大多是通过利用实体的固有特征（文本描述、属性和类型）或在文本中实体指代项来进行类型推断，最近有学者研究将知识库中的实体转换为实体图，并应用到基于图神经网络的算法模型中。

9.3 嵌套命名实体识别

通常要处理的命名实体是非嵌套实体，但是在实际应用中，嵌套实体非常多。大多数命名实体识别会忽略嵌套实体，无法在深层次文本理解中捕获更细粒度的语义信息。如图 5 所示，在“3 月 3 日，中国驻爱尔兰使馆提醒旅爱中国公民重视防控，稳妥合理加强防范。”句子中提到的中国驻爱尔兰使馆是一个嵌套实体，中国和爱尔兰均为地名，而中国驻爱尔兰使馆为组织机构名。普通的命名实体识别任务只会识别出其中的地名“中国”和“爱尔兰”，而忽略了整体的组织机构名。



3 月 3 日，中国驻爱尔兰使馆提醒旅爱中国公民重视防控，稳妥合理加强防范

图 5 嵌套实体示例

Fig.5 Example of nested entity

学者们提出了多种用于嵌套命名实体识别的方法。Finkel 和 Manning[132]基于 CRF 构建解析器，将每个命名实体作为解析树中的组成部分。Ju 等[133]动态堆叠多个扁平命名实体识别层，并基于内部命名实体识别提取外部实体。如果较短的实体被错误地识别，这类方法可能会遭受错误传播问题的困扰。嵌套命名实体识别的另一系列方法是基于超图的方法。Lu 和 Roth[134]首次引入了超图，允许将边缘连接到不同类型的节点以表示嵌套实体。Muis 和 Lu[135]使用多图表示法，并引入分隔符的概念用于嵌套实体检测。但是这样需要依靠手工提取的特征来识别嵌套实体，同时遭受结构歧义问题的困扰。Wang 和 Lu[136]提出了一种使用神经网络获取分布式特征表示的神经分段超图模型。Katiyar 和 Cardie[137]提出了一种基于超图的计算公式，并以贪婪学习的方式使用 LSTM 神经网络学习嵌套结构。这些方法都存在超图的虚假结构问题，因为它们枚举了代表实体的节点、类型和边界的组合。Xia 等提出了 MGNER 架构，不仅可以识别句子中非重叠的命名实体，也可以识别嵌套实体，此外不同于传统的序列标注任务，它将命名实体识别任务分成两部分开展，首先识别实体，然后进行实体分类。

嵌套实体识别充分利用内部和外部实体的嵌套信息，从底层文本中捕获更细粒度的语义，实现更深层次的文本理解，研究意义重大。

9.4 命名实体链接

命名实体链接主要目标是进行实体消歧，从实体指代项对应的多个候选实体中选择意思最相近的一个实体。这些候选实体可能选自通用知识库，例如维基百科、百度百科，也可能来自领域知识库，例如军事知识库、装备知识库。图 6 给出了一个实体链接的示例。短文本“美海军陆战队 F/A-18C 战斗机安装了生产型 AN/APG-83 雷达”，其中实体指代项是“生产型 AN/APG-83 雷达”，该实体指代项在知识库中可能存在多种表示和含义，而在此处短文本，其正确的含义为“AN/APG-83 可扩展敏捷波束雷达”。

图 6 实体链接示例

Fig.6 Example of named entity linking

实体链接的关键在于获取语句中更多的语义，通常使用两种方法。一种是通过外部

语料库获取更多的辅助信息，另一种是对本地信息的深入了解以获取更多与实体指代项相关的信息。Tan 等[138]提出了一种候选实体选择方法，使用整个包含实体指代项的句子而不是单独的实体指代项来搜索知识库，以获得候选实体集，通过句子检索可以获取更多的语义信息，并获得更准确的结果。Lin 等[137]寻找更多线索来选择候选实体，这些线索被视为种子实体指代项，用作实体指代项与候选实体的桥梁。Dai 等[138]使用社交平台 Yelp 的特征信息，包括用户名、用户评论和网站评论，丰富了实体指代项相关的辅助信息，实现了实体指代项的歧义消除。因此，与实体指代项相关的辅助信息将通过实体指代项和候选实体的链接实现更精确的歧义消除。

另一些学者使用深度学习研究文本语义。Francis-Landau 等[139]使用卷积神经网络学习文本的表示形式，然后获得候选实体向量和文本向量的余弦相似度得分。Ganea 和 Hofmann[140]专注于文档级别的歧义消除，使用神经网络和注意力机制来深度表示实体指代项和候选实体之间的关系。Mueller 和 Durrett[141]将句子左右分开，然后分别使用门控循环单元和注意力机制，获得关于实体指代项和候选 实体的分数。Ouyang 等提出一种基于深度序列匹配网络的实体链接算法，综合考虑实体之间的内容相似度和结构相似性，从而帮助机器理解底层数据。目前，在实体链接中使用深度学习方法是一个热门的研究课题。

10 研究难点

关于命名实体识别领域当前的研究难点，陈曙东等[117]主要将其总结为三点：领域命名实体识别的局限性、命名实体表述多样性和歧义性、命名实体的复杂性和开放性。

10.1 领域命名实体识别的局限性

目前命令实体识别只是在有限的领域和有限的实体类型中取得了较好的成绩，如针对新闻语料中的人名、地名、组织机构名的识别。但这些技术无法很好地迁移到其他特定领域中，如军事、医疗、生物、小语种语言等。一方面，由于不同领域的数据往往具有领域独特特征，如医疗领域中实体包括疾病、症 状、药品等，而新闻领域的模型并不适合；另一方面，由于领域资源匮乏造成标注数据集缺失，导致模型训练很难直接开展。因此，采用半监督学习、远监督学习、无监督学习方法实现资源的自动构建和补足，以及迁移学习等技术的应用都可作为解决该问题的核心研究方向。

10.2 命名实体表述多样性和歧义性

自然语言的多样性和歧义性给自然语言理解带来了很大挑战，在不同的文化、领域、背景下，命名实体的外延有差异，是命名实体识别技术需要解决的根本问题。获取大量

文本数据后，由于知识表示粒度不同、置信度相异、缺乏规范性约束等问题，出现命名实体表述多样、指代不明确等现象。因此，需要充分理解上下文语义来深度挖掘实体语义进行识别。可以通过实体链接、融合对齐等方法，挖掘更多有效信息和证据，实现实体不同表示的对齐、消除歧义，从而克服命名实体表述多样性和歧义性。

10.3 命名实体的复杂性和开放性

传统的实体类型只关注一小部分类型，例如“人名”、“地名”、“组织机构名”，而命名实体的复杂性体现在实际数据中实体的类型复杂多样，需要识别细粒度的实体类型，将命名实体分配到更具体的实体类型中。目前业界还没有形成可遵循的严格的命名规范。命名实体的开放性是指命名实体内容和类型并非永久不变，会随着时间变化发生各种演变，甚至最终失效。命名实体的开放性和复杂性给实体分析带来了巨大的挑战，也是亟待解决的核心关键问题。

11 未来发展方向

命名实体识别从提出以来，一直是信息检索、数据挖掘、自然语言处理等领域中一个重要的研究领域。从 MUC 到 ACE 再到 CoNLL，一系列重要的评测会议划定了 NER 的基本研究范围，也提出了大量经典的重要的研究方法。与大多数 NLP 问题类似，NER 的发展基本经历了一种从规则向统计的转向。早期的规则方法已经不再流行，但其研究思路仍然给人以宝贵的启示，且规则和统计相结合的方法仍不时得到有效的尝试。如今的 NLP 领域，统计机器学习方法日臻完善，NER 也在这辆高速列车上走向成熟，而深度学习带来的机器学习新热潮，将会使 NER 在统计机器学习的道路上继续高速地推进。

然而从目前已有的研究成果来看，NER 研究还远不是一个得到完善解决甚至将要完善解决的问题，各领域下对命名实体定义的模糊，实验结果在 80%~90%徘徊的 F1 值，使得 NER 仍然是一个有挑 战性的研究领域。一方面，大数据环境下，机器学习乃至深度学习仍将是最有效的 NER 方法。而另一方面，虽然机器学习带来了 NER 的火热发展，但大量研究固化于调整经典模型、挑选更多特征、扩大语料规模这种三角模式，是值得研究者们反思的。NER 的研究不应局限于 F1 值的提高上，从更多的角度来思考 NER 这一问题，才能使这个研究领域获得更全面的发展。比如当语料规模不足时，不是考虑扩大语料的规模，而是使用迁移学习方法来解决[143]。NER 在其他学科上的应用也是未来一个重要的研究方向。将已有的 NER 方法有效地应用在各种领域的文本上，帮助各种学科获取其所关注的命名实体，这本就是 NER 研究的意义和价值所在。

参考文献

[1] Grishman R . Message Understanding Conference-6: A Brief History[C]// Proceedings of the 16th conference on Computational linguistics. 1996.

[2] Chinchor N . MUC-7 Named entity task definition version 3.5[J]. 1997.

[3] Xie R ,Liu Z , Jia J , et al. Representation learning of knowledge graphs with entity descriptions. 2016.

[4] Hartley B . Improving machine translation quality with automatic named entity recognition. Association for Computational Linguistics, 2003.

[5] Zhu J ,Uren V , Motta E . ESpotter: Adaptive Named Entity Recognition for Web Browsing[C]// Third Biennial Conference on Professional Knowledge Management. 2005.

[6] LDC. Entity detection and tracking-phase 1 ace pilot study task definition[EB/OL]. [2017-03-10]. https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/edt-phase1-v2.2.pdf.

[7] LDC. Annotation guidelines for entity link tracking (LNK) Version 3.0 20040401[EB/OL]. [2017-03-10]. https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/english-lnk-v3.0.PDF.

[8] Rau LF . Extracting company names from text[C]// Artificial Intelligence Applications, 1991. IEEE, 1991.

[9] Sang K, Tjong EF.Introduction to the CoNLL-2002 Shared Task[J]. COLING-02 proceedings of the 6th conference on Natural language learning - Volume 20, 2002.

[10] Sang EFTK,De MeulderF.Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition[J]. arXiv, 2003.

[11] Spyropoulos C D . Automatic adaptation of proper noun dictionaries through cooperation of machine learning and probabilistic methods[C]// International Acm Sigir Conference on Research & Development in Information Retrieval. ACM, 2000.

[12] Alfonseca E , Manandhar S. An Unsupervised Method for General Named Entity Recognition And Automated Concept Discovery. 2002.

[13] Sekine S , Sudo K , Nobata C. Extended Named Entity Hierarchy. 2002.

[14] Sekine S, Nobata C. Definition, dictionaries and tagger for extended named entity hierarchy[J]. Proc of Lrec, 2004.

[15] Borrega O ,Mariona Taulé, M Antònia Martí. What do we mean when we speak about Named Entities[J]. 2007.

[16] Nadeau D, Sekine S. A survey of named entity recognition and classification[J]. Lingvisticae Investigationes, 2007, 30(1): 3-26.

[17] Marrero M, Urbano J, S Sánchez-Cuadrado, et al. Named Entity Recognition: Fallacies, challenges

and opportunities[J]. Computer Standards & Interfaces, 2013, 35(5):482-489.

[18] Bikel D M , Schwartz R , Weischedel R M . An Algorithm that Learns What's in a Name[J]. Machine Learning, 1999, 34.

[19] Liao W ,Veeramachaneni S. ASimple Semi-supervised Algorithm For Named Entity Recognition. Association for Computational Linguistics, 2009.

[20] Stevenson S, Carreras X . Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL-2009). 2009.

[21] 孙茂松, 黄昌宁, 高海燕,等. 中文姓名的自动辨识[J]. 中文信息学报, 1995, 9(2):12.

[22] 张小衡, 王玲玲. 中文机构名称的识别与分析[J]. 中文信息学报, 1997, 11(4):22-33.

[23] Zhang Y , Zhou J F . A trainable method for extracting Chinese entity names and their relations. 2000.

[24] Tsai T H , Wu S H , Lee C W , et al. Mencius: A Chinese Named Entity Recognizer Using Maximum Entropy-based Hybrid Model. 2004.

[25]冯元勇, 孙乐, 李文波,等. 基于单字提示特征的中文命名实体识别快速算法[C]// 第三届全国信息检索与内容安全学术会议论文集. 2007:104-110.

[26] 郑逢强, 林磊, 刘秉权,等. 《知网》在命名实体识别中的应用研究[J]. 中文信息学报, 2008, 22(5):5.

[27] 李嘉欣, 王平. 中文命名实体识别研究方法综述[J]. 计算机时代, 2021(4):4.

[28] Patil N V , Patil A S , Pawar B V . HMM based Named Entity Recognition for inflectional language[C]// 2017 International Conference on Computer, Communications and Electronics (Comptelix). 2017.

[29] Yao L, Liu H, Liu Y, et al. Biomedical named entity recognition based on deep neutral network[J]. Int. J. Hybrid Inf. technol, 2015, 8(8): 279–288.

[30] Strubell E , Verga P , Belanger D , et al. Fast and Accurate Entity Recognition with Iterated Dilated Convolutions[J]. 2017.

[31] Yang J , Liang S , Zhang Y . Design Challenges and Misconceptions in Neural Sequence Labeling[C]// 2018.

[32] Kong J , Zhang L , Jiang M , et al. Incorporating multi-level CNN and attention mechanism for Chinese clinical named entity recognition[J]. Journal of Biomedical Informatics, 2021, 116:103737.

[33] Huang Z , Wei X , Kai Y . Bidirectional LSTM-CRF Models for Sequence Tagging[J]. Computer Science, 2015.

[34] Zhang Y , Yang J . Chinese NER Using Lattice LSTM[J]. 2018.

[35] Zhang H , YGuo, Li T . Domain Named Entity Recognition Combining GAN and BiLSTM-Attention-CRF[J]. Journal of Computer Research and Development, 2019.

[36] Wolf T , Debut L , Sanh V , et al. Transformers: State-of-the-Art Natural Language Processing[C]// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. 2020.

[37] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:, 1810, 04805: 2018.

[38] Dai Z, Wang X, Ni P, et al. Named entity recognition using bert bilstm crf for chinese electronic health records[C]//2019 12th international congress on image and signal processing, biomedical engineering and in

formatics (cisp-bmei). IEEE, 2019: 1−5.

[39] Li X, Zhang H, Zhou X H. Chinese clinical named entity recognition with variant neural structures based on

BERT methods[J]. Journal of biomedical informatics, 2020, 107: 103422.

[40] 毛明毅, 吴晨, 钟义信, 陈志成. 加入自注意力机制的 BERT 命名实体识别模型 [J]. 智能系统学报,2020, 15(04): 772–779.

Mao Mingyi, Wu Chen, Zhong Yixin, Chen Zhicheng. Self attention mechanism based Bert named entity recognition model[J]. CAAI Transactions on Intelligent Systems, 2020, 15(04): 772–779.

[41] Li X, Yan H, Qiu X, et al. FLAT:Chinese NER Using Flat-Lattice Transformer[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 6836-6842.

[42] Yoon W, So C H, Lee J, et al. Collabonet: collaboration of deep neural networks for biomedical named entity recognition[J]. BMC bioinformatics, 2019, 20(10): 55–65.

[43] Doddington G . The automatic content extraction (ACE) program-tasks, data, and evaluation [J]. Proc Lrec, 2004.

[44]LDC. Entity detection and tracking – Phase 1 EDT and metonymy annotation guidelines [EB/OL].[2017-03-10].https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/edt-guidelines-v2-5.pdf.

[45]LDC. Annotation guidelines for relation detection and characterization (RDC) Version 3.6 - 4.22.2002[EB/OL]. [2017-03-10]. https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/rdc-g

uidelines-v3.6.pdf.

[46] NIST. The ACE 2004 evaluation plan evaluation of the recognition of ACE entities, ACE relations and ACE events[EB/OL]. [2017-03-10].http://itl.nist.gov/iad/mig/tests/ace/2004/doc/ace

04-evalplan-v7.pdf.

[47] LDC. ACE (Automatic Content Extraction) English annotation guidelines for events [EB/OL]. [2017-03-10].https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/english-events-

guidelines-v5.4.3.pdf.

[48] NIST. The ACE 2005 evaluation plan evaluation of the detection and recognition of ACE enti -ties, values, temporal expres- sions, relations, and events[EB/OL]. [2017-03-10]. http://itl.nist. gov/iad/mig/tests/ace/2005/doc/ace05-evalplan.v3.pdf.

[49] LDC. ACE (Automatic Content Extraction) Spanish annotation guidelines for entities [EB/OL]. [2017-03-10]. https://www.ldc. upenn.edu/sites/www.ldc.upenn.edu/files/spanish-entities-guidel ines-v1.6.pdf.

[50] LDC. GALE Arabic translation guidelines V2.3[EB/OL]. [2017-03-10]. http://projects.ldc. upenn.edu/gale/Translation/specs/GALE_Arabic_translation_guidelines_v2.3.pdf.

[51] LDC. GALE Chinese translation guidelines V2.3[EB/OL].[2017-03-10]. http://projects.ldc.u -penn.edu/gale/Translation/specs/GALE_Chinese_translation_guidelines_v2.3.pdf.

[52] NIST. The ACE 2007 evaluation plan evaluation of the detecion and recognition of ACE entities, values, temporal expres- sions, relations, and events[EB/OL]. [2017-03-10]. http://itl.nist. gov/iad/mig/tests/ace/2007/doc/ace07-evalplan.v1.3a.pdf.

[53] LDC. ACE 2008: Cross-document annotation guidelines (XDOC)[EB/OL]. [2017-03-10]. https://www.ldc.upenn.edu/sites/ www.ldc.upenn.edu/files/ace08-xdoc-1.6.pdf.

[54] NIST. Automatic content extraction 2008 evaluation plan assessment of detection and recognition of entities and relations within and across documents[EB/OL]. [2017-03-10]. http:// itl.nist. gov/iad/mig/tests/ace/2008/doc/ace08-evalplan.v1.2d.pdf.

[55] Burger J D, Henderson J C, Morgan W T. Statistical named entity recognizer adaptation[C]// Proceedings of the 6th Con- ference on Natural Language Learning, Stroudsburg: Associa- tion for Computational Linguistics, 2002, 20: 1-4.

[56] Chieu H L, Ng H T. Named entity recognition with a maximum entropy approach[C]// Conference on Natural Language Learn- ing at HLT-NAACL, 2003: 160-163.

[57] Curran J R, Clark S. Language independent NER using a maximum entropy tagger[C]// Conference on Natural Language Learning at HLT-NAACL, 2003: 164-167.

[58] Mayfield J, McNamee P, Piatko C. Named entity recognition using hundreds of thousands of features[C]// Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Computational Linguis- tics, 2003: 184-187.

[59] McCallum A, Li W. Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons[C]// Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Computational Linguistics, 2003,4: 188-191.

[60] Carreras X, Marquez L, Padró L. Named entity extraction using adaboost[C]// Proceedings of the 6th Conference on Natural Language Learning, Stroudsburg: Association for Computa- tional Linguistics, 2002, 20: 1-4.

[61] Zhang T, Johnson D. A robust risk minimization based named entity recognition system[C]// Proceedings of the Seventh Con- ference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Computational Linguistics, 2003, 4:
204-207.
[62] Carreras X, Màrquez L, Padró L. Learning a perceptron-based named entity chunker via online recognition feedback[C]// Pro- ceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Com- putational Linguistics, 2003, 4: 156-159.
[63] Hammerton J. Named entity recognition with long short-term memory[C]// Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Computational Linguistics, 2003, 4: 172-175.
[64] Black W J, Vasilakopoulos A. Language independent named entity classification by modified transformation-based learning and by decision tree induction[C]// proceedings of the 6th Con- ference on Natural Language Learning, Stroudsburg: Associa- tion for Computational Linguistics, 2002, 20: 1-4.
[65] Florian R. Named entity recognition as a house of cards: Classifier stacking[C] // Proceedings of the 6th Conference on Natural Language Learning, Stroudsburg: Association for Computa- tional Linguistics, 2002, 20: 1-4.
[66] Florian R, Ittycheriah A, Jing H, et al. Named entity recognition through classifier combination[C]// Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Computational Linguistics, 2003, 4:168-171.
[67] Klein D, Smarr J, Nguyen H, et al. Named entity recognition with character-level models[C]// Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, Stroudsburg: Association for Computational Linguistics, 2003, 4:180-183.
[68] 赵军. 命名实体识别、排歧和跨语言关联[J]. 中文信息学报,2009, 23(2): 3-17.
[69] Chang J, Chen S, Chen Y, et al. A multiple-corpus approach to identification of Chinese surname-names[C]// Proceedings of Natural Language Processing Pacific Rim Symposium, 1991:87-91.
[70] Wang L, Li W, Chang C. Recognizing unregistered names for mandarin word identification[C]//Proceedings of the 14th Con- ference on Computational Linguistics, Stroudsburg: Association for Computational Linguistics, 1992, 4: 1239-1243.
[71] 张俊盛, 陈舜德, 郑萦等.多语料库作法之中文姓名辨识[J]. 中文信息学报, 1992, 6(3): 9-17.
[72] 张小衡, 王玲玲.中文机构名称的识别与分析[J]. 中文信息学报, 1997, 11(4): 21-32.
[73] 宋柔, 朱宏.基于语料库和规则库的人名识别法[C]//全国第二届计算机语言学联合学术会议.北京: 北京语言学院出版 社, 1993: 150-154.
[74] 郑家恒,刘开瑛.自动分词系统中姓氏人名处理策略探讨[C]//全国第二届计算机语言学联合学术会议. 北京: 北京语 言学院出版社, 1993: 139-143.

[75] 孙茂松, 黄昌宁, 高海燕, 等. 中文姓名的自动辨识[J].中文信息学报, 1995, 9(2): 16-27.

[76] 孙镇, 王惠临. 命名实体识别研究进展综述[J]. 现代图书情报技术, 2010, 26(6): 42-47.

[77] Zhou J S, He L, Dai X Y, et al. Chinese Named Entity Recognition with a Multi-Phase Model[C]// Proceedings of the Fifth SIGHAN Workshop on Chinese Language Processing. Stroudsburg: Association for Computational Linguistics, 2006: 213-216.

[78] Chen A, Peng F, Shan R, et al. Chinese named entity recognition with conditional probabilistic models[C]// Proceedings of the Fifth SIGHAN Workshop on Chinese Language Processing. Stroudsburg: Association for Computational Linguistics, 2006:173-176.

[79] Chen W L, Zhang Y J, Isahara H. Chinese named entity recognition with conditional random fields[C]// Proceedings of the Fifth SIGHAN Workshop on Chinese Language Processing. Stroudsburg: Association for Computational Linguistics, 2006:118-121.

[80] Mao X N, Dong Y, He S K, et al. Chinese word segmentation and named entity recognition based on conditional random fields[C]// Proceedings of the Sixth SIGHAN Workshop on Chinese Language Processing, 2008: 90-93.

[81] Zhao H, Kit C. Unsupervised segmentation helps supervised learning of character tagging for word segmentation and named entity recognition[C]// Proceedings of the Fourth International Chinese Language Processing Bakeoff & the First CIPS Chinese Language Processing Evaluation, 2008: 106-111.

[82] Yu X F, Lam W, Chan S K, et al. Chinese NER using crfs and logic for the fourth SIGHAN bakeoff[C]// Proceedings of the Sixth SIGHAN Workshop on Chinese Language Processing, 2008: 102-105.

[83] Zhang S X, Qin Y, Wen J, et al. Word segmentation and named entity recognition for SIGHAN Bakeoff3[C]// Proceedings of the Fifth SIGHAN Workshop on Chinese Language Processing. Stroudsburg: Association for Computational Linguistics, 2006:158-161.

[84] Chen Y, Jin P, Li W, et al. The Chinese persons name disambiguation evaluation: Exploration of personal name disambigua- tion in Chinese news[C]//Processings of the CIPS-SIGHAN Joint Conference on Chinese Language, 2010.

[85] He Z, Wang H, Li S. The Task 2 ofCIPS-SIGHAN 2012 named entity recognition and disambiguation in Chinese Bakeoff[C]// Proceedings of the Second CIPS-SIGHAN Joint Conference on Chinese Language Processing. 2012: 108-114.

[86] Zong H, Wong D F, Chao L S. A template based hybrid model for Chinese personal name disambiguation[C]// Proceedings of the Second CIPS-SIGHAN Joint Conference on Chinese Lan- guage Processing, Tianjin, China, 2012: 121-126.

[87] Tian W, Pan X, Yu Z T, et al. Chinese name disambiguation based on adaptive clustering with the attribute features[C]// Pro- ceedings of the Second CIPS-SIGHAN Joint Conference on Chinese Language Processing, Tianjin, China, 2012: 132-137.

[88] Collins M, Singer Y. Unsupervised models for named entity classification[C]// Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora, 1999: 100-110.

[89] Cucerzan S, Yarowsky D. Language independent named entity recognition combining morphological and contextual evi- dence[C]// Proceedings of the 1999 Joint SIGDAT Conference on EMNLP and VLC, 1999: 90-99.

[90] Y ANG P， LIU W， YANG J.Positive Unlabeled Learning Via Wrapper-based Adaptive Sampling[C]// IJCAI.2017 : 3273-3279.

[91] SANG E F，DE MEULDER F.Introduction to the CoNLL-2003 Shared Task : Language-independent Named Entity Recognition[J].arXiv preprint cs /0306050，2003.

[92]ZHANG Y，YANG J.Chinese Ner Using Lattice Lstm[J].arXiv preprint arXiv : 1805.02023，2018.

[93] REN X， HE W， QU M， et al.Afet : Automatic Fine-grained Entity Typing by Hierarchical Partial-label Embedding [C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing，2016 : 1369-1378.

[94] ZHOU J T，ZHANG H，JIN D，et al.Dual Adversarial Neural Transfer for Low-resource Named Entity Recogni- tion[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics，2019 :3461-3471.

[95] ZHANG Q，FU J，LIU X，et al.AdaptiveCo-attention Network for Named Entity Recognition in Tweets [C] // Thirty-Second AAAI Conference on Artificial Intelligence,2018.

[96] YU G L，XU Y，MA M D，et al.Rule-based named entity recognition in Uyghur[J].China Science & Technology Panorama Magazine，2015(15):33.

[97] ZHANG X Y，WANG T，CHEN H W.A mixed statistical model-based method for Chinese named entity recog- nition[J].Computer Engineering & Science，2006，28(6):135- 139.

[98] LI N，ZHENG R T，JI J M，et al.Research on Chinese chemical name recognition based on heuristic rules[J]. New Technology of Library and Information Service，2010(5):13- 17.

[99] CHENG C， CHENG X Y， HUA J.Research of Chinese named entity recognition using GATE[J].Advanced Mate- rials Research，2012，393/394/395:262-264.

[100] HE H.Introduction to natural language processing[M]. Beijing:Posts and Telecom Press，2019:111-129.

[101] PORTER M.An algorithm for suffix stripping[J].Program Electronic Library and Information，1980，14（3）:130- 137.

[102] CULOTTA A， MCCALLUM A． Confidence Estimation for Information Extraction[C] // Proceedings of HLT-NAACL 2004 : Short Papers，2004 : 109-112．
[103] CARPENTER B．Ling Pipe for 99．99% Recall of Gene Mentions[C]// Proceedings of the Second BioCreative Challenge Evaluation Workshop．BioCreative，2007，23 : 307-309．
[104] MINKOV E,WANG RC， TOMASIC A， et al.NER systems that Suit User ’s Preferences : Adjusting the Recall-precision Trade-off for Entity Extraction[C]// Proceedings of the Human Language Technology Conference of the NAACL,Companion Volume : Short Pa- pers．2006 : 93-96.
[105] LAMPLE G， BALLESTEROS M， SUBRAMANIAN S， et al．Neural Architectures for Named Entity Recognition[J].arXiv preprint arXiv : 1603．01360，2016．
[106] UKOV-GREGORI A，BACHRACH Y， COOPE S．
Named Entity Recognition with Parallel Recurrent Neural Networks[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)，2018 : 69-74．
[107] ZHOU J T，ZHANG H， JIN D， et al． Roseq : Robust Sequence Labeling [J]．IEEE Transactions on Neural Networks and Learning Systems，2019，PP (99) : 1-11．
[108] COLLOBERT R，WESTON J，BOTTOU L，et al．Natural Language Processing (almost) from Scratch[J]．Journal of Machine Learning Research，2011，12 (Aug) : 2493-2537．
[109] CHIU J P C,NICHOLS E.Named Entity Recognition with Bidirectional LSTM-CNNs[J].Transactions of the Association for Computational Linguistics，2016，4 : 357-370．
[110] MA X，HOVY E．End-to-end Sequence Labeling Via Bidirectional Lstm-cnns-crf[J].arXiv preprint arXiv :1603．01354，2016．
[111] LIU L，SHANG J，REN X，et al．Empower Sequence Labeling with Task-aware Neural Language Model [C] // Thirty-Second AAAI Conference on Artificial Intelligence．2018．
[112] LIU T，YAO J G，LIN C Y．Towards Improving Neural Named Entity Recognition with Gazetteers[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics，2019 : 5301-5307．
[113] GREENBERG N，BANSAL T，VERGA P，et al．Marginal Likelihood Training of Bilstm-crf for Biomedical Named Entity Recognition from Disjoint Label Sets [C] // Pro- ceedings of the 2018 Conference on Empirical Methods in Natural Language Processing，2018 : 2824-2829．
[114] AUGENSTEIN I， RUDER S， SGAARD A． Multi-task Learning of Pairwise Sequence Classification Tasks over Disparate Label Spaces[J]．arXiv preprint arXiv : 1802.09913，2018．
[115] BERYOZKIN G，DRORI Y，GILON O，et al．A Joint Named-Entity Recognizer for Heterogeneous Tag-sets Using a Tag Hierarchy[J]．arXiv preprint arXiv : 1905.09135，2019.
[116] 张吉祥,张祥森， 武长旭， 赵增顺.知识图谱构建技术综述[J/OL].计算机工程．https://doi.org/10.19678/j.issn.1000-3428.0061803

[117] CHEN Shudong, OUYANG Xiaoye. Overview of Named Entity Recognition Technology [J]. Radio Communications Technology，2020，46 (3) : 251-260.

[118] CHEN Y， ZONG C， SU K Y. On Jointly Recognizing and Aligning Bilingual Named Entities[C]// Proceedings of the 48th Annual Meeting of the Association for Computa- tional Linguistics. Association for Computational Linguistics，2010 : 631-639.

[119] FENG X，FENG X，QIN B，et al. Improving Low Resource Named Entity Recognition using Cross-lingual Knowledge Transfer[C]// IJCAI. 2018 : 4071-4077.

[120] MAYHEW S， TSAI C T， ROTH D. Cheap Translation for Cross - lingual Named Entity Recognition [C]// Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing，2017: 2536-2545.

[121] ZHOU J T，ZHANG H，JIN D，et al. Dual Adversarial Neural Transfer for Low-resource Named Entity Recogni- tion[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics，2019 :3461-3471.

[122] YANG P， LIU W， YANG J. Positive Unlabeled Learning Via Wrapper-based Adaptive Sampling[C]// IJCAI.2017 : 3273-3279.

[123] PENG M，XING X，ZHANG Q，et al. Distantly Supervised Named Entity Recognition using Positive -Unlabeled Learning[J]. arXiv preprint arXiv : 1906. 01378，2019.

[124]REN X， HE W， QU M， et al. Afet : Automatic Fine-grained Entity Typing by Hierarchical Partial-label Embedding [C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016 : 1369-1378.

[125] LING X，WELD D S. Fine-grained Entity Recognition[C]// Twenty-Sixth AAAI Conference on Artificial Intelligence，2012.

[126] NEELAKANTAN A，CHANG M W.Inferring Missing Entity Type Instances for Knowledge Base Completion : New dataset and methods[J]. arXiv preprint arXiv : 1504.06658，2015.

[127] YAGHOOBZADEH Y， SCHTZE H. Multi-level Representations for Fine-grained Typing of Knowledge Base Entities[J]. arXiv preprint arXiv : 1701. 02025，2017.

[128] JIN H，HOU L，LI J，et al. Attributed and Predictive Entity Embedding for Fine-grained Entity Typing in Knowledge Bases[C]// Proceedings of the 27th International Conference on Computational Linguistics，2018 : 282-292.

[129] DEFFERRARD M， BRESSON X， VANDERGHEYNST P.Convolutional Neural Networks on Graphs with Fast Local- ized Spectral Filtering[C]// Advances in Neural Informa- tion Processing Systems，2016 : 3844-3852.

[130] ATWOOD J， TOWSLEY D. Diffusion - convolutional Neural Networks[C] // Advances in Neural Information Processing Systems，2016 : 1993-2001.

[131] JIN H，HOU L，LI J，et al．Fine-Grained Entity Typing Via Hierarchical Multi Graph Convolutional Networks [C]// Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th In- ternational Joint Conference on Natural Language Processing (EMNLP-IJCNLP)，2019 : 4970-4979．

[132] FINKEL J R，MANNING C D．Nested Named Entity Recognition[C] // Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing : Volume 1-Volume 1．Association for Computational Lin- guistics，2009 : 141-150．

[133] JU M，MIWA M，ANANIADOU S．A Neural Layered Model for Nested Named Entity Recognition[C] // Pro- ceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies，Volume 1 (Long Papers)，2018 : 1446-1459．

[134] LU W，ROTH D．Joint Mention Extraction and Classification with Mention Hypergraphs [C]// Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing，2015 : 857-867．

[135] MUIS A O，LU W．Labeling Gaps Between Words : Recognizing Overlapping Mentions with Mention Separators[J]．arXiv preprint arXiv : 1810．09073，2018．

[136] WANG B，LU W．Neural Segmental Hypergraphs for Overlapping Mention Recognition [J]．arXiv preprint arXiv : 1810．01817，2018．

[137] KATIYAR A，CARDIE C． Nested Named Entity Recognition Revisited[C]// Proceedings of the 2018 Con- ference of the North American Chapter of the Association for Computational Linguistics : Human Language Technolo- gies，Volume 1 (Long Papers)，2018 : 861-871．

[138] TAN C， WEI F， REN P， et al． Entity Linking for Queries by Searching Wikipedia Sentences[J]．arXiv preprint arXiv : 1704．02788，2017．

[137] LIN Y，LIN C Y，JI H．List only Entity Linking[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)，2017: 536-541．

[138] DAI H，SONG Y，QIU L，et al．Entity Linking within a Social Media Platform : A Case Study on Yelp[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing，2018 : 2023-2032．

[139] FRANCIS-LANDAU M， DURRETT G， KLEIN D．Capturing Semantic Similarity for Entity Linking with Convolutional Neural Networks[J].arXiv preprint arXiv :1604．00734，2016．

[140] GANEA O E， HOFMANN T.Deep Joint Entity Disambiguation with Local Neural Attention [J]．arXiv preprint arXiv : 1704．04920，2017．

[141] MUELLER D，DURRETT G．Effective Use of Context in Noisy Entity Linking[C]// Proceedings of the 2018 Con- ference on Empirical Methods in Natural Language Processing，2018 : 1024-1029．