



A Recent Survey of Arabic Named Entity Recognition on Social Media

Brahim Ait Ben Ali^{1*}, Soukaina Mihi¹, Ismail El Bazi², Nabil Laachfoubi¹

¹ Hassan First University of Settat, Faculty of Sciences and Techniques, IR2M Laboratory, Settat 26000, Morocco

² Sultan Moulay Slimane University, National School of Business and Management, Beni Mellal 23000, Morocco

Corresponding Author Email: b.aitbenali@uhp.ac.ma

<https://doi.org/10.18280/ria.340202>

ABSTRACT

Received: 26 December 2019

Accepted: 7 February 2020

Keywords:

named entity recognition, Arabic dialect, NLP, social media, formal and informal text

In the last few years, significant amounts of text data have emerged on the different social media platforms. A tendency to extract valuable information from these data for useful purposes has been created and developed. The Named Entity Recognition (NER), as a sub-task of the Natural Language Processing (NLP), remains primordial in order to perform these extractions and the classification of entity names from the text regardless of its structure "formal or informal". Nevertheless, the most recent solutions for NER are confronted with the difficulty of adapting to the informal texts used on social media platforms. This work aims at providing a literature review of the various papers published in the field of NER on social media starting from 2014 until now, by identifying the particular characteristics surrounding the Arabic dialect compared to the English language.

1. INTRODUCTION

The Named Entities Recognition is a sub-task of automatic natural language processing used to identify then to extract useful information such as the names of persons, places, organizations starting from an unstructured text [1]. This approach was first introduced during the "Sixth message understanding conference -MUC6- [2], whose main theme was essentially the extraction of information from a formal text. Several papers on these systems were first concerned by The English language, followed by a series of publications on other languages, namely German, Spanish, Dutch, Japanese and Indian, etc. Concerning the Arabic language, the application of this concept was not initiated until 2005 [3].

It is claimed that most research projects around this theme in the different languages have reached a very high level of performance similar to those of human subjects, particularly in English [4].

Considering a large amount of textual data, which has emerged with the advent of social media and given the importance of the information hidden behind, the researchers have been particularly interested in the extraction and the classification of the named entities according to different predefined categories. However, these systems, as they are currently used to analyze the content of informal texts from social media, for example, show a decrease in the performance score [5]. This is due to spelling errors, grammatical errors, the use of abbreviations, non-compliance with punctuation, etc.

The identification of NEs in the raw information is not an easy task in Arabic languages because Arabic languages are not capitalized. Resources such as gazetteers, dictionaries, POS markers, morphological analyzers are not freely accessible. There are many variations in spelling style. Work on NERs in Arabic languages, particularly colloquial Arabic, is a complex and challenging task as well as limited because of the lack of resources, but it has begun to emerge.

This observation shows the necessity to create and develop

new algorithms able to satisfy efficiently, the difficulties of analyzing informal texts from social media.

Motivations for conducting this survey: In the world, today, where social media are omnipresent, dealing with informal forms is becoming increasingly crucial. In recent years, NER in social media has received considerable attention because of its many new challenges. Hence, a significant number of studies have applied to NER in English and have successively advanced state-of-the-art performance. This trend motivates us to conduct a survey to report on current NER systems in social media for the English language. By comparing these systems, we attempt to try in the future to consider the possibility of integrating these system architectures that have a very high degree of performance to handle the dialectal Arabic language.

On the other hand, to review the literature regarding the several approaches carried out in the field of NER in dialect Arabic, whose number of published articles is restricted and very limited compared to those in the English language considered as the most authoritative language. For this purpose, we will analyze the different approaches related to the publications on dialectal Arabic language in the NER field and their assessment/evaluation with those of the English language to identify the lacunae and challenges, which affect its progress and NER performance. These challenges remain both numerous and varied, especially in social media.

Contributions of this survey: We intensively examine articles of NER in the English language to enlighten and provide guidance to researchers and practitioners in this field. Besides, we study the best-known methods of NER problems newly applied to the English language. Finally, we introduce readers to the challenges facing NER systems in Arabic and describe future orientations in this area. To the best of our knowledge, this is the first survey focusing on Arabic NER on social media and comparing to previous systems applied to English.

The rest of our paper will be as described below: In section

2, the interest will be focused on the methodology used to select works that meet our objectives as mentioned above, when In section 3, we will discuss the NER and its different applications, then in section 4, we will present the most representative datasets for the NER task, afterward, we describe the most accurate systems used to extract entities in order to compare their obtained results in section 6. We end with a concluding part.

2. METHODOLOGY

In order to find the articles in this study, we searched, Science Direct, Web of sciences, Google Scholar, and Semantic Scholar. Our query terms included named entity recognition in social media for both two languages Arabic and English, machine/deep learning models for named entity recognition on social media. We filtered the retrieved items for each request by several quotations and read at least the top three. We maintained for our study the articles that are based on machine learning and deep learning for entity recognition and the models that represent high performance for NER in a given dataset.

3. BACKGROUND

Before reviewing the systems used in the field of NER, we first provide a formal formulation of the problem of NER. We then present the broader role of NERs. Next, we present the future directions

3.1 What is NER?

A named entity is a word or expression that uniquely describes an element among a set of other elements with similar attributes. Names of named entities include names of organizations, persons, and places in the global domain, names of genes, proteins, drugs, and diseases in the biomedical field. Named entities recognition (NERs) is defined as the process of identifying and classifying the entities named in the text into predefined entity categories [6].

Formally, considering a series of tokens $S = \{W_1, \dots, W_2, W_N\}$, NER aims to output a list of tuples, each tuple specify the entity referred to a named entity in S here, $I_s \in [1, N]$ and $I_e \in [1, N]$ are the start and end indexes of a named entity mentioning; t is the entity type of a set of predetermined categories.

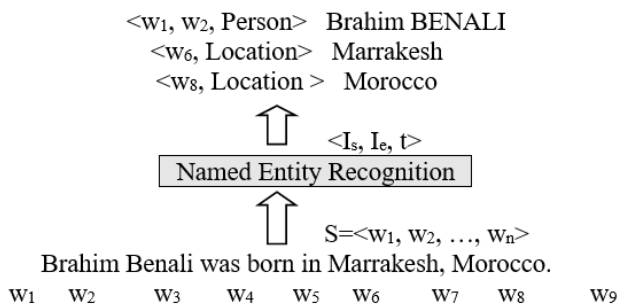


Figure 1. An illustration of the named entity recognition task

Figure 1 illustrates an example where a NER system identifies three named entities based on the given phrase. Once

NER was initially defined in MUC-6 [2], then the task is to identify the names of persons, organizations, places, and time, currency, and percentages in the text. Notice it focuses on a small set of coarse entities and one type for each named entity. This type of NERs task is called coarse-grained NERs [2, 7]. Lately, a number of the fine-grained NER tasks [8, 9] concentrate on a wider set of entity types for which mention can be assigned to several of the fine-grained types.

3.2 The broader role of NER

The general implications of NER related to NLP research are too prominent to list. Some examples of applications for which NER is useful are presented in this section.

•**Information Retrieval:** This consists of identifying and extracting relevant documents from a data set based on an incoming request. A research study that was performed by the author [10], noted that approximately 71% of all search engine queries contain NE. Information extraction can be facilitated by the NER in two phases [11] first, find the NE in the request; second, identify the NE in the searched documents, and then retrieve the relevant documents with consideration for their classified NEs and how they are linked with the request [12, 13]. For instance, the word (Aljazeera) may be identified as an organization name or a name that corresponds to the word island; correct classification will allow easy retrieval of pertinent documents.

•**Question Answering:** This is quite similar to information extraction, but with over advanced findings. A question-answering approach utilizes questions as an entry and provides short and precise answers in return [14]. Besides, in order to facilitate the recognition of NEs in a query, the NER task can be used during its analysis phase. This will, therefore, enable us to identify and locate the relevant documents and even to provide an appropriate answer [14]. Question answering systems may benefit considerably from NERs, as the answers to many factual questions require NEs [15] (e.g., answers to questions who (ماهي/من هو) typically involve persons or organizations, where (اين) questions include places, and when (متى) questions include temporal expressions) [16].

•**Machine Translation:** This refers to the automatic translation of a text from one natural language into another. NEs need particular attention when deciding which parts of NEs should be translated and which parts should be transliterated [17]. For instance, people's names tend to be transliterated in the case of a location name, the name portion and the category portion (e.g., mountains) are usually transliterated and translated, each one respectively. The quality of the NER system plays an important role when it comes to determining the overall quality of the machine translation system. Therefore, NE translation is essential for most multilingual application systems.

•**Text Clustering:** The clustering of the research results can exploit the NER by sorting the resulting clusters according to the ratio of entities that each cluster contains [11, 18]. This improves the process of analyzing the nature of each cluster and also enhances the clustering approach in terms of selected characteristics. For instance, time expressions, as well as location NEs, could be used as factors that will indicate an indication of when and where the events mentioned in a document occurred.

•**Navigation Systems:** Such systems, which make it easier to navigate using digital maps, have now become an essential part of our lives. They give indications, information on

neighboring places, which may be related to other online resources, and circulation conditions. In such systems, points of interest (also known as waypoints) are NEs that are stored in a database with their geo-coordinates [19]. They cover areas of interest that are generally important to, among others, tourists, visitors, and rescuers, making it possible to locate places such as parking lots, shops, hospitals, restaurants, universities, schools, landmarks, etc.

3.3 Future direction

With the progress of modeling languages and the increasing need for real-world applications, we expect NERs to get further attention from the research community. Furthermore, NERs are generally seen as a preprocessing for downstream applications, which means that a specific NER task is determined by the required downstream application, e.g., the kinds of named entities and if there is a necessity to identify nested entities.

4. ARABIC NAMED RECOGNITION CHALLENGES

4.1 Arabic challenges

The United Nations know Arabic as one of the world's main languages as it is mostly used worldwide by 300 million people in 28 countries. It is part of the Semitic language group, which also includes Hebrew and Amharic, Ethiopia's primary language.

Three kinds of the Arabic language are available:

- **Classical Arabic:** The language of the Holy Quran;
- **Modern Standard Arabic (MSA):** Language of official documents, newsletters, education. Traditionally, it is the same language as all Arabs.
- **Colloquial Arabic:** or dialectal Arabic, it is the informal language, which people use to communicate daily; it varies from one country to another. In principle, five colloquial Arabic dialects can be distinguished: Egyptian, Levantine, Maghreb, Gulf, and Iraqi [20].

Arabic can be written from right to left. Contrary to English or French, however, there are no "capital" letters.

The morphology of Arabic is very complex [21]; from a root, we can get words that are different lexically and semantically.

For example, the words "madrassa" and "modarissa" are extracted from the root "d-,r-,s- درس" which is written in the same way in Arabic مدرسة, but with different meanings (madrassa=school, modarissa = educator).

4.2 Arabic twitter challenges

Social media networks contain a tremendous amount of unstructured data. Actually, tweets present multiple challenges for analytics comparing to standard text.

In the Arab world, the majority of people write social media content in an informal way, sometimes in a mixte of bilingual languages, using the Latin words inside an Arabic tweet. Besides, we can find non-Arabic words written in Arabic letters, more often in Maghreb dialects, people use Latin letters to write colloquial Arabic words.

Another challenge consists of repeating letters inside or at the end of a word, for example, when calling a name. For named entity, an example of repeating a letter when calling someone, for example, أميين = أمين.

4.3 Datasets

4.3.1 W-Nut 2015

First, they retrieved tweets through the Twitter Streaming API from May 23-29, 2014, followed by using langid.py (<https://github.com/saffsd/langid.py>) to delete all non-English tweets. The tokenization was performed via the CMU-ARK (<https://github.com/myleott/ark-twokenize-py>) tokenizer. Then, to ensure that the tweets had a high probability of requiring lexical normalization, they filtered out tweets containing less than two non-standard words (i.e., words not in their dictionary).

The two training and dev datasets (Table 1) for NER were taken from previous work on Twitter NER [22]. However, the test dataset was randomly sampled from December 2014 through Feb 2015.

Table 1. An overview of datasets for the English language

Data	Tweets	Tokens
Train	1,795	37,899
Dev	599	11570
Test	1,000	16,261

4.3.2 W-Nut 2016

The new test set includes annotated tweets with named entities as of a later period than the 2015 data. The new test set deployed for the 2016 task iteration includes 3,856 tweets in total, which is roughly two times the quantity of annotated data available for the named entity recognition on Twitter. In their preprocessing phase, they employ Twitter tokenization, POS tagging for tokenized data via TwitIE (<https://gate.ac.uk/wiki/twitie.html>).

This dataset (Table 2) provided a set of 3,856 training tweets and a set of 1,000 development tweets corresponding to all the tests of the 2015 edition of the challenge.

Table 2. W-NUT 2016 datasets for the English language

Dataset	Tweets	Named Entities
Training	3,856	1,768
Development	1,000	661
Test	3,850	3,473

4.3.3 W-Nut dataset 2017

The data sets provided by the organizers of the shared task W-NUT 2017 on the recognition of emerging and rare named entities [23]. The statistics for the data sets are presented (Table 3). They tagged and provided access to 2,295 texts taken from three various sources (Reddit, Twitter, YouTube, and StackExchange comments), covering both. The tweets were encoded using Twokenizer and processed by GATE (<http://gate.ac.uk/sale/tao/index.html>) for crowdsourcing. The corpus was not checked for lewdness and possibly offensive content.

Table 3. Twitter dataset for the English language

Data set	tweets	named entities
Training	3,394	1,975
Development	1,009	833
Test	1,287	1,041

4.3.4 W-Nut dataset 2018

The organizers offered three distinct English datasets: a training set, a development set, and a test set (Table 4). The

data consists of multilingual English tweets and contains NEs in both languages English and Arabic [24].

In order to label the data, they built a CrowdFlower (https://crowdfunder.com/) job from scratch (The JavaScript code and HTML/CSS can be found here: https://github.com/tavo91/ner_annot). The task enables the annotators to choose either one or more words for a single NE. Once the annotators selected a word, the tool proposes to incorporate the words around the actual selection.

Table 4. Data set statistics for ENG and MSA-DA

Language	Train	Dev.	Test
MSA-DA	10,102	1,122	1,110
ENG	8,733	1,857	2,158

All tweets in the MSA-EGY dataset total 12,334 tweets. This is split into three sets: train, dev, and test (10,102, 1,122, and 1,110 tweets, respectively). The MSA-EGY tweets were collected from the chronology of 12 Egyptian political figures. In general, politicians usually tend to use NEs more frequently when writing their tweets. Hence, the percentage of NE tokens in the MSA-EGY dataset is much more significant compared to the rate of NE tokens in the ENG dataset.

4.3.5 Darwish dataset

They use the training and test data sets developed by Darwish [25] to test their model. This dataset has been labeled with three different types of named entities: location, person, and organization. The training dataset consists of randomly chosen tweets from May 3-12, 2012. The test data set includes tweets that were randomly selected from November 23, 2011, to November 27, 2011. Note that these two datasets were tagged according to the Linguistic Data Consortium's tagging guidelines. As shown in Table 5, it gives the statistics of the dataset.

Table 5. Twitter evaluation data statistics

	Tokens	PER	LOC	ORG
Twitter Train	55k	788	713	449
Twitter Test	26k	464	587	316

4.4 Preprocessing

Text preprocessing is an essential task of natural language processing. It consists of applying operations to create another form from the text inputted. The first step is data-cleaning, which is carried out after removing punctuation, diacritics, and stop words (words that have no significant semantic relationship with the context in which they exist) from the text, diacritics suppression, punctuation suppression. Second, Arabic normalization phase in order to have a unified and unique form of letters (اآا will be replaced by ا, ؤ,س,ع will be replaced by, ع and ؤ will be replaced by ؤ).

After that, the word stemming in Arabic that's means the process of deleting all prefixes and suffixes from a word is almost used to generate the stem or root. It is a process of converting plurals into singulars or deriving a verb from the gerund form. Other possibilities include deriving the root from model words. This process of derivation is necessary for

classifiers and index builders/researchers because it reduces dependency on particular word forms and reduces the potential size of vocabularies, which otherwise might have to contain all possible forms.

5. NER SYSTEMS ON SOCIAL MEDIA: LITERATURE REVIEW

5.1 NER systems

Most of the NER systems employ either Rule-based or data-based approaches. Concerning the Rule-based systems, they are a combination of rewriting rules and lexicons (lists of first names, countries, etc.). However, data-oriented systems are based on a probabilistic model learned from previously annotated corpora.

- **Rule-based approaches:** it is based on human expertise to construct analysis models manually. Most often, these models are presented in the form of contextual rules, which can detect the named entities. For this reason, this approach is also called the “**Linguistic approach.**”

Experts manually produce these rules based on linguistic descriptions, specific characteristics, proper name lexicons, and trigger words. They consist of regular expressions that will allow identifying and classifying named entities.

- **Hybrid approach:** in order to take advantage of the two methods, the hybrid approach combines the Rule-based and data-based approaches. The rules for extracting named entities are either automatically learned and then filtered manually via a language module, written manually, and then corrected and improved automatically via an automatic learning module.
- **Learning-based approaches:** This approach uses statistically based rules for extracting named entities and is deduced autonomously. In general, acquisition of these rules is carried out from a large corpus - under supervised or semi-supervised conditions.

In this study, the focus will be on the papers that have addressed the latter approach, which is the subject of our work.

Many of the systems that have been analyzed use CRF and SVM machine learning techniques.

5.2 Support Vector Machine

Support Vector Machines (SVM) [26] are widely known techniques for supervised machine learning. SVMs have been utilized efficiently in a number of problems associated with pattern recognition in fields like bioinformatics and biometrics. Concerning text processing, SVMs achieved the highest results in text categorization and are widely used in NLP-related problems in various languages, including Arabic for such methods as named-entity recognition and sentiment analysis.

The SVM (Figure 2) creates a binary classifier that generates +1 or -1 from a sample vector $x \in R^n$. The decision is based on the given hyperplane separation as follow:

$$c(x) = \begin{cases} +1 & \text{if } w \cdot x + b > 0, w \in R^n, b \in R, \\ -1 & \text{otherwise} \end{cases} \quad (1)$$

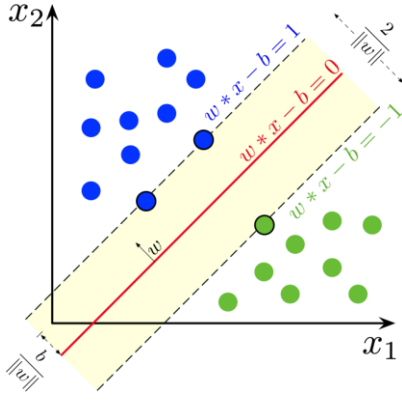


Figure 2. Illustration of support-vector machine

5.3 Conditional random fields

CRF [27] is a probabilistic framework for segmenting and labeling sequence data. This is based on undirected graphical models in which nodes denote the sequence of y labels matching the x sequence. The CRF model attempts to find the y tag that maximizes the conditional likelihood $p(y/x)$ for an x sequence. The CRF model consists of a feature-based model where the features having binary values like:

$$f_k(y_{t-1}, y_t, x) = 1 \text{ for } x = ' \text{Morocco}' \text{ and } y_t = ' B \\ - Loc', \text{ and } 0 \text{ otherwise.}$$

The CRF model considers itself as a generalization of the maximum entropy and Hidden Markov Model (HMM) and may be formulated as follows:

$$p(y/x) = \frac{1}{Z(x)} \exp\left(\sum_t \sum_k \gamma_i \cdot f_k(y_{t-1}, y_t, x)\right)$$

where, γ_i denotes the weights attributed to the various features during the training phase and $Z(x)$ is a normalization factor obtainable as follows:

$$Z(x) = \sum_{y \in Y} \exp\left(\sum_t \sum_k \gamma_i \cdot f_k(y_{t-1}, y_t, x)\right)$$

5.4 Machine learning techniques for NER

Many studies have been conducted in the field of NER on social media by using machine learning algorithms.

In fact, Akhtar et al. [28] provided a technique based on multiobjective differential evolution for the feature selection in the named entities recognition on Twitter, While Yamada et al. [29] through his research provided a new approach to enhance the efficiency of Twitter's NER task by utilizing entity linking, as a method to identify entity mentions in text and solve them into matching entries in the Wikipedia knowledge base. Their approach is a supervised machine-learning environment, which uses high-level knowledge from several extensive. Regarding Tian et al. [30], they used a CRF model that utilized the Wapiti toolkit. The feature model includes state-of-the-art standard features. They first developed a model trained with training data and tested with dev data, and then they assessed it with dev_2015 shared task data. On the other hand, Toh et al. [31] considered the NER as a sequential labeling task and utilized conditional random fields. A number of post-processing steps (e.g., rule-based matching) were

implemented to fine-tune the output of the system. In parallel to the brown clusters, K-means clusters were similarly used; the latter were produced based on word embedding. Godin et al. [32] provided a semi-supervised system, which identifies ten types of named entities. They used a neural network model for the classification process. As features, they used the word embedding, which was trained on more than 400 million micro-post Twitter to produce word representations. They increase performance by using dropout and the rectified linear units (ReLU) to form the network. In addition to this, Yang and Kim [33] developed an approach based on CRF in which both Brown clusters and words embedded are used, the latter trained using canonical correlation analysis as a feature. Moreover, Derczynski et al. [34] used feature extraction, which relied on large Brown clusters, gazetteers suitable for input data, and freebase remote monitoring. The mapping was adapted to the drift using examples of training that were remote in time and weighted down. The classifier was a straight-chain CRF with hyper-parameters set for Twitter. Cherry et al. [35] employed a MIRA-trained semi-Markov labeler with Gazetteer, Brown Clusters, and Word Embedding functions. The word embedding was constructed on phrases through the Word2Vec phrase search utility and was edited via an automatic encoder to be able to predict membership of the Gazetteers. Sikdar and Gambäck [36] introduced a system, which is based on classification utilizing a Conditional Random Fields, a supervised machine learning approach. This system uses an extensive feature set developed specifically for the task, with eight types of features based on actual characters and token internal data, five types of features built through context and chunk information, and five types of features based on lexicon-type information such as stop word matching, word frequencies, and entries in the shared task lexicon and Babelfy. Partalas et al. [37] provided a system, which uses three types of characteristics: lexical and morpho-syntactic characteristics, contextual enrichment characteristics utilizing linked open data, and features based on the distributed representation of words. The system also exploits word clustering to improve performance. The learning algorithm was solved by applying learning to search (L2S), which resembles a reinforcement-learning algorithm. Le et al. [38] provided a system, which is relying on supervised machine learning and trained with a sequential tagging algorithm, which uses conditional random fields to train a classifier for Twitter NE retrieval. Their model employs six distinct classes of features such as (1) spelling, (2) lexical and (3) syntactic as well as (4) parts of speech, (5) polysemia, and (6) n-grams to create a representation of the features. Jain et al. [39] employed conditional random fields with numerous handcraft features. They, also, mostly concentrated on the English-Spanish data. Also, Sikdar et al. [40] used a CRF model. They provided this model with features from both external and internal resources. Besides, they included the linguistic identification labels from the data sets from previous versions of this workshop. Whereas, Janke et al. [41] relied on conditional random fields with many features. A number of these features were also used for neural networks, but they performed best with the CRF approach. Finally, Claeser et al. [42] employed an SVM classifier within a Radial Basis nucleus. They manufactured by hand a wide array of features and also contained gazetteers.

Most researchers, who have conducted their research on the Dialect Arabic language, have used the Dataset of Kareem Darwish [25] by proceeding through the machine learning

techniques.

Indeed, the researchers [43] developed a system using simple effective language-independent approaches based on using extensive gazetteers, domain adaptation, and a two-pass semi-supervised method for enhancing the NER on microblogs. While Zayed and El-Beltagy [44] implemented a hybrid approach to extract Arab person names from tweets and resolving their ambiguity by utilizing contextual bigram models. The adopted approach attempts not to use language-dependent resources. The evaluation of the presented approach shows a 7% improvement in the F-score compared to the best reported results in the literature. From another part, Zirikly and Diab [45] adopted a supervised machine learning approach by using the Conditional Random Fields sequence labeling, word embedding, and word representations to extract named entities on social media.

5.5 Deep learning techniques for NER

The researches carried out and concerning the English language were numerous and varied. They yielded significant results. Indeed, Limsopatham and Collier [46] proposed a novel technique that enables the neural network, especially B-LSTM, to learn and use spelling features specifically. They concentrate on spelling features that are effective and broadly used in many NER systems. Espinosa et al. [47] developed an unsupervised learning approach that uses deep neural networks and a knowledge base (i.e., DBpedia) to initiate sparse entity types with lightly tagged data. Dugas and Nichols [48] developed a system, which uses a bi-directional LSTM-CNN model with word embedding on a large scale-web corpus. Additionally, the system utilizes self-built lexicons based on a combination of a partial matching algorithm and text standardization to deal with the significant vocabulary issue in web texts. Gerguis et al. [49] installed a system using a set of features such as word integration, brown clusters, speech party tags, shape characteristics, gazetteers, and local context to create a functionality representation and a set of experiences for network design. A categorization scheme based on Wikipedia has been introduced to extract lists of fine-grained entities from a few sample entries to serve as gazetteers for toponymic names. As a result, the model employs a long-short memory (LSTM) recurrent neural network model to learn a Named Entity (NE) classifier from the entity representation. Finally, Winata et al. [50] utilized a bi-directional LSTM model for both characters and words. They mostly centered on Out-of-Vocabulary (OOV) words using the FastText library.

In 2017, many studies were focused on the combination of LSTM and CRF, which has yielded meaningful results in terms of performance. Indeed, Jansson and Liu [51] suggested combining latent Dirichlet allocation (LDA) modeling with deep learning about character and word level incorporations. Sikdar and Gambäck [52] built a system that uses three initial classifier templates using CRF, support vector machines (SVM), and an LSTM model. The results of these three classifiers were then used as characteristics to form another CRF classifier that functioned as a group. Lin et al. [53] concatenated character representation level, word-level representation, and word syntax representation (i.e., POS tags, dependency roles, word positions, head positions) to provide a full representation of the words. Rodríguez-Somolinos [54] discussed a deep learning approach for the task of recognition of named entities on Twitter data, which expands a fundamental neural network for sequence labeling by using

sentence level features and learning by transfer. A multi-task method for NER has been suggested by Aguilar et al. [55]. This method consists of using CNN to capture spelling characteristics and word shapes at the character level. For syntactic and contextual information at the word level, for example, POS and word embedding, the model applies an LSTM approach.

Within such a framework, Trivedi et al. [56] suggested a new architecture based on the gate of the representation of a token based on characters and words. The character and word representations were collected with a Convolutional Neural Network (CNN) and a bidirectional LSTM, both respectively. They also applied multi-tasking training on the output layer and transferred the training to a CRF classifier as defined by Akhtar et al. [28]. Furthermore, they gave a mapping of gazetteers to their template. At the moment when Wang et al. [57] suggested a common bi-directional LSTM-CRF network that utilizes attention to the embedding layer. They also pretreated the data before they fed it into the network. Finally, Geetha et al. [58] provided a bidirectional LSTM on the output layer with a conditional random field.

The focus on dialect Arabic language also included deep learning techniques to improve the performance of the systems developed [59].

In addition, Helwe and Elbassuoni [60] employed a bidirectional LSTMs and CRF on the top of the overall approach with the help of character-level embedding, pre-trained word embedding and they obtained the state-of-the-art results in building an Arabic named entity recognition system for social media without the use of any extensive gazetteer and many of hand-engineering features. Gridach [61] developed an approach based on the BiLSTM-CRF network with the help of word embedding, brown clusters, and gazetteers. Moreover, Attia et al. [62] provided a semi-supervised co-training approach in the field of deep learning on social media. This approach, called deep co-learning, uses a small amount of tagged data, which is supplemented by partially labeled data that is automatically produced from Wikipedia.

6. DISCUSSION (2014-2018)

In order to make a reasonable comparison, we have grouped the systems that have been trained and tested on the same dataset in the same table.

This section outlines the approach adopted by each team. Overall, we found different trends between the types of systems reported in 2016 and 2015. The most significant change is the utilization of LSTM-based systems. There were four of the seven submissions based on LSTM, compared to zero in 2015. In the previous year, conditional random fields were the most critical use of ML for named feature extraction.

A total of eight teams contributed to the Shared Task of Named Entity Recognition on the Alan Ritter datasets [38]. Broad diversity of different approaches was adopted to address this task. Table 6, outlines the features utilized by each team and the machine learning approach adopted. Several teams used word embedding and Brown's clusters as features. However, there was one team MULTIMEDIA LAB that utilized absolutely no handcrafted features, entirely based on word embedding with a feed-forward neural network (FFNN) approach [32]. The other new methods to Twitter's NER comprise a semi-Markov MIRA tagger developed collaboratively by the NRC team (Cherry and Guo, 2015) and

the utilization of entity linking as a feature by OUSIA [29]. Every other team utilized CRFs. In addition to a CRF, the IITP

team employed a technique called differential evolution to obtain an optimal set of features.

Table 6. Features and machine learning methods implemented by each group of researchers

Approach	POS	Orthographic	Gazetteers	Brown clustering	Word embedding	Machine learning
BASELINE	-	X	X	-	-	CRFSuite
HALLYM [33]	X	-	-	X	Correlation analysis	CRFSuite
IITP [28]	X	X	X	-	-	CRF++
LATTICE [30]	X	X	-	X	-	CRF wapiti
MULTIMEDIALAB [32]	-	-	-	-	Word2vec	FFNN
NLANGP [31]	-	X	X	X	Word2vec & Glove	CRF++
NRC [35]	-	-	X	X	Word2vec	Semi-markov MIRA
OUSIA [29]	X	X	X	-	X	Entity linking
USFD [34]	X	X	X	X	-	CRF L-BFGS

Except for the MULTIMEDIALAB, which did not use dev2015 as training data, the majority of the systems utilized the training data with the two dev given to training their system. The NRC uses only training.

Table 7 presents the findings for each team for the classification of the ten named entity.

As of 2016, the researchers became interested in deep-learning techniques. The table below shows the mains systems that have used it in the retrieval process (Table 8).

A significant change concerning 2015 relates to the increased usage of neural network approaches by the participants. Various teams, among them the winning team, CAMBRIDGELTL, employed bi-directional LSTM. Some teams reached competitive performance by implementing a wide range of linguistic and knowledge characteristics using conditional random fields (e.g., NTNU) or research learning methods (TALOS).

Table 8 shows the different features and machine learning methods implemented by each group of researchers in the field of NER systems on social media, as discussed in section 6 on the W-NUT 2016 datasets.

Indeed, the authors [46, 47] who have used LSTM obtained

a high-performance score from 44.77% to 52.41%.

Table 9 presents the findings for each team for the classification of the named entity.

Table 10 illustrates that the combination of the LSTM and CRF algorithms allows improving the performance level of Named Entities Recognition Systems.

Table 11 shows the performance levels of the different systems as developed and applied to the English language.

Table 7. The results attained by previous NER systems

Approach	Precision	Recall	F1
OUSIA	57.66	55.22	56.41
NLANGP	63.62	43.12	51.40
NRC	53.24	38.58	44.74
MULTIMEDIALAB	49.52	39.18	43.75
USFD	45.72	39.64	42.46
IITP	60.68	29.65	39.84
HALLYM	39.59	35.10	37.21
LATTICE	55.17	9.68	16.47
BASELINE	35.56	29.05	31.97

Table 8. Features and machine learning methods implemented by each group of researchers

Approach	POS	Orthographic	Gazetteers	B. clustering	W. embedding	ML
BASELINE	-	X	X	-	-	CRFSuite
CAMBRIDGELTL [46]	-	-	-	-	-	LSTM
AKORA [47]	-	-	-	-	-	LSTM
NTNU [36]	X	X	X	-	-	CRF
TALOS [37]	X	X	X	X	GLOVE	L2S
DEEPNNER [48]	-	-	-	-	MULTIPLE	LSTM-CNN
ASU [49]	-	-	X	-	-	LSTM
UQAM-NTL [38]	X	X	X	-	-	CRF

Table 9. The results attained by previous NER systems

Approach	Precision	Recall	F1
CAMBRIDGELTL	60.77	46.07	52.41
TALOS	58.51	38.12	46.16
AKORA	51.70	39.48	44.77
NTNU	53.19	32.13	40.06
ASU	40.58	37.58	39.02
DEEPNNER	54.97	28.16	37.24
DeepER	45.40	31.15	36.95
HJPWHU	48.90	28.76	36.22
UQAM-NTL	40.73	23.52	29.82
LIOX	40.15	12.69	19.26
BASELINE	38.12	11.96	19.11

The table presents the principal components and approaches utilized by the participants. Ext. Res. refers to external resources like pre-formed word embedding, gazetteers, and so on. Hand Feats refers to craft features like capitalization.

The outcomes from the five systems that participated in this task demonstrate that entity recognition is, in fact, more complex than the very frequent entity recognition challenges commonly identified in the Named Entity Recognition Challenges. Further work is, therefore, necessary for this area, and this shared task is only a small start. In the future, data sets of this type may need to be extended, possibly with other classes of entities for particular domains. In addition, they expect that other NLP tasks will address the challenge of creating more diverse reference datasets to increase their coverage of the use of rare and new languages.

Table 10. Features and machine learning methods implemented by each group of researchers

Approach	POS	Gazetteers	Suffix/prefix	W. Embedding	ML
BASELINE	-	X	X	-	CRF
ARCADA [51]	X	-	-	Word & character	LSTM -CRF
FLYTXT [52]	X	-	X	Word	LSTM -CRF- SVM
SJTU-ADAPT [53]	X	X	-	Word	LSTM-CRF
SPINNINGBYTES [54]	X	-	-	Word & sentence	LSTM-CRF
UH-RiTUAL [55]	X	X	-	Word	LSTM-CRF

Table 11. The results attained by previous NER systems

Approach	F1
ARCADA [22]	39.98
FLYTXT [28]	38.35
SJTU-ADAPT [25]	40.42
SPINNINGBYTES [7]	40.78
UH-RiTUAL [1]	41.86
BASELINE	38.46

The gap between the performance levels achieved by the different systems mentioned in the table (Table 11) can be considered as slight. For example, the gap between the first two systems, UH-RiTUAL, and SPINNINGBYTES, equals 1.08% when the difference between ARCADA and FLYTXT is no more than 1.63%. Table 12 shows the different features and machine learning methods implemented by each group of researchers in the field of NER systems on social media, as discussed in section 6 on the datasets and trained on both datasets; English and dialectal Arabic [63].

Table 12. Features and machine learning methods implemented by each group of researchers

Approach	Preprocess	Ext. Res.	HandcraftFeatures	CNN	B-LSTM	CRF	Other
BASELINE					X	X	-
IIT BHU [57]	-	X	-	X	X	X	MTL
CAIRE++ [60]	-	-	-	-	X	-	FastText
FAIR [58]	X	-	-	-	X	X	Attention
LINGUISTS [33]	-	X	X	-	-	X	-
FLYTXT [50]	-	X	-	-	-	X	-
SEMANTIC [25]	-	-	-	-	X	X	-
BATS [34]	-	X	X	-	-	X	-
FRAUNHOFER [10]	-	X	X	-	-	-	SVM
GHHT [4]	-	X	-	-	X	X	-

Concerning both Tables 13 and 14, they show the performance levels of the different systems as developed and applied to the English language and Dialect Arabic.

We remark, based on Table 13, that the values of the performance results achieved by the different systems applied in the English language are high and similar to each other, which allows us to identify, therefore, a small gap between each of the two consecutive systems. For example: “the gap between the first two systems IIT BHU and CAIRE equals 1% when the gap between both lasts BATS and FRAUNHOFER FKIE is no more than 0.5%”.

In addition, the results obtained by the different systems applied to the dialect Arabic language in Table 14 reveal high values "FAIR 71.61%, GHHT 70.0938%..." with low average-gap not exceeding 3% between every two consecutive systems. For example, “the gap between the first two systems FAIR and GHHT equals 1.52% when the gap between the both lasts BATS and SEMANTIC is no more than 0.6%”.

After analyzing the results, it seems that Conditional random fields were the most prevalent aspect among the approaches. Besides, using the CRF in combination with a bi-directional LSTM system (with some variations) produced the highest scores among the participants. As a result, the highest F1- score obtained for the ENG-SPA was 63.7628% while for the MSA-EGY was 71.6154%. In comparison to the formal text (i.e., the news wire), the achieved scores are considerably smaller because of code-switching and noise from the SM environment. This demonstrates the necessity for more robust approaches capable of detecting and processing named entities under severe conditions.

Table 13. Results obtained for English

Approach	F1
IIT BHU	63.7628
CAIRE++	62.7608
FAIR	62.6671
LINGUISTS	62.1307
FLYTXT	59.2501
SEMANTIC	56.7205
BATS	54.1612
FRAUNHOFER FKIE	53.6514
BASELINE	51.8569

Table 14. Results obtained for Arabic dialect

Approach	F1
FAIR	71.6154
GHHT	70.0938
LINGUISTS	67.4419
BATS	65.6207
SEMANTIC	65.0276
BASELINE	62.5478

7. ANALYSIS

We described in detail the task configuration and data sets employed in the respective shared tasks, and also represented the approach adopted by the participating systems. The shared tasks were significantly broader in scope compared to what had been tried earlier in the literature, resulting in two main advantages. Firstly, we are in a position to make more robust

findings of the real potential of the various approaches. Secondly, by analyzing the outcomes from the participating systems, we are capable of suggesting potential research directions for building the systems for dialect Arabic.

8. CONCLUSIONS

In conclusion, and based on this recent literature review, we can summarize that the different languages, especially English, have been studied in the NER research with various degrees of priority. Except for Arabic, whose numbers of studies are still insufficient and very limited. This is due to a number of reasons, such as the lack of resources (e.g., an annotated dataset, rich and numerous gazetteers...), the complexity of the Arabic language, etc.... All this requires multiple efforts to extend, enlarge, and diversify the datasets to improve the performance scores of the different approaches.

We have examined the state of the art of NER systems in Arabic in quite considerable detail. It is important to note that the reference list given here may not be complete. Our objective was to review the main aspects of the Arabic NER and to discuss the main publications that have made use of these ideas. It is expected that this study will provide access to the main branches of the literature related to NER research in Arabic and allow for research in the relevant and fruitful field.

REFERENCES

- [1] Nadeau, D., Sekine, S. (2007). A survey of named entity recognition. *Communications*, 30(1): 14-30. <https://doi.org/10.1075/li.30.1.03nad>
- [2] Sundheim, B., Grishman, R. (1996). Message understanding conference -6: A brief history. *COLING '96: Proceedings of the 16th conference on Computational linguistics*, pp. 466-471. <https://doi.org/10.3115/992628.992709>
- [3] Huang, F. (2005). Multilingual named entity extraction and translation from text and speech. PhD thesis, Lang. Technol. Institute, Sch. Comput. Sci. Carnegie Mellon Univ.
- [4] Finkel, J.R., Grenager, T., Manning, C. (2005). Incorporating non-local information into information extraction systems by Gibbs sampling. *ACL-05 - 43rd Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pp. 363-370. <https://doi.org/10.3115/1219840.1219885>
- [5] Ritter, A., Cherry, C., Dolan, B. (2010). Unsupervised modeling of twitter conversations. *NAACL HLT 2010 - Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 172-180.
- [6] Sharnagat, R. (2014). Named entity recognition literature survey. *Cent. Indian Lang. Technol.*, pp. 1-47.
- [7] Sang, E.F.T.K., De Meulder, F. (2003). Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. *Proceedings of CoNLL-2003, Edmonton, Canada*, pp. 142-147.
- [8] Ling, X., Weld, D.S. (2003). Fine-grained entity recognition. *AAAI'12: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, July, pp. 94-100.
- [9] Del Corro, L., Abujabal, A., Gemulla, R., Weikum, G. (2015). FINET: Context-aware fine-grained named entity typing. *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, pp. 868-878.
- [10] Guo, J., Xu, G., Cheng, X., Li, H. (2009). Named entity recognition in query. *SIGIR '09: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 267-274. <https://doi.org/10.1145/1571941.1571989>
- [11] Benajiba, Y., Rosso, P., Diab, M. (2009). Arabic named entity recognition: A feature-driven study. *IEEE Transactions on Audio, Speech and Language Processing*, 17(5): 926-934. <https://doi.org/10.1109/TASL.2009.2019927>
- [12] Fauzi, M.A., Arifin, A.Z., Yuniarti, A. (2017). Arabic book retrieval using class and book index based term weighting. *International Journal of Electrical and Computer Engineering*, 7(6): 3705-3710. <https://doi.org/10.11591/ijece.v7i6.pp3705-3710>
- [13] Darwish, K., Magdy, W. (2014). Arabic information retrieval foundations and trends R in information retrieval. *Foundations and Trends in Information Retrieval*, 7(4): 239-342. <https://doi.org/10.1561/15000000031>
- [14] Shaheen, M., Ezzeldin, A.M. (2014). Arabic question answering: Systems, resources, tools, and future trends. *Arabian Journal of Science and Engineering*, 39(6): 4541-4564. <https://doi.org/10.1007/s13369-014-1062-2>
- [15] Trigui, O., Belguith, L.H., Rosso, P., Ben Amor, H., Gafsaoui, B. (2012). Arabic question answering for Machine Reading evaluation. *CEUR Workshop Proc.*, pp. 2-5.
- [16] Brini, W., Ellouze, M., Tirgui, O., Mesfar, S., Hadrich Belguith, L., Rosso, P. (2009). Factoid and definitional arabic question answering system. *Arabian Journal for Science and Engineering*, 39(6): 4541-45. <https://doi.org/10.1016/j.infsof.2008.09.005>
- [17] Yasser Al-Onaizan, Hassan, H., Sorensen, J. (2005). An integrated approach for Arabic-English named entity translation. *Semitic '05: Proceedings of the ACL Workshop on Computational Approaches to Semitic Languages*, pp. 87-93. <https://doi.org/10.3115/1621787.1621803>
- [18] Sangaiah, A.K., Fakhry, A.E., Abdel-Basset, M., El-henawy, I. (2019). Arabic text clustering using improved clustering algorithms with dimensionality reduction. *Cluster Computing*, 22: 4535-4549. <https://doi.org/10.1007/s10586-018-2084-4>
- [19] Kim, S.Y., Kim, S.H., Cho, H.G. (2012). Developing a system for searching a shop name on a mobile device using voice recognition and GPS information. *Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication, ICUIMC'12*, pp. 1-8. <https://doi.org/10.1145/2184751.2184785>
- [20] Kwaik, K.A., Saad, M., Chatzikyriakidis, S., Dobnik, S. (2019). Shami: A corpus of levantine Arabic dialects. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan*, pp. 3645-3652.
- [21] Al Huneety, A.I. (2015). The phonology and morphology of wadi mousa Arabic. Phd thesis, Sch. Humanit. Lang. Soc. Sci. Univ. Salford, UK.
- [22] Ritter, A., Clark, S., Etzioni, O. (2014). Named entity

- recognition in tweets: An experimental study. EMNLP '11: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1524-1534.
- [23] Derczynski, L., Nichols, E., van Erp, M., Limsopatham, N. (2018). Results of the WNUT2017 shared task on novel and emerging entity recognition. Proceedings of the 3rd Workshop on Noisy User-generated Text, Copenhagen, Denmark, pp. 140-147. <https://doi.org/10.18653/v1/W17-4418>
- [24] Aguilar, G., AlGhamdi, F., Soto, V., Diab, M., Hirschberg, J., Solorio, T. (2019). Named entity recognition on code-switched data: Overview of the CALCS 2018 shared task. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 138-147. <https://doi.org/10.18653/v1/W18-3219>
- [25] Darwish, K. (2013). Named entity recognition using cross-lingual resources: Arabic as an example. ACL 2013 - 51st Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, pp. 1558-1567.
- [26] Cortes, C., Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(5): 273-297. <https://doi.org/10.1007/BF00994018>
- [27] Lafferty, J., McCallum, A. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Computer Vision*. 2001, June (2001), 282-289. https://doi.org/10.1007/978-0-387-31439-6_100233
- [28] Akhtar, M.S., Sikdar, U.K., Ekbal, A. (2015). IITP: Multiobjective differential evolution based twitter named entity recognition. Proceedings of the Workshop on Noisy User-generated Text, Beijing, China, pp. 61-67. <https://doi.org/10.18653/v1/w15-4308>
- [29] Yamada, I., Takeda, H., Takefuji, Y. (2015). Enhancing named entity recognition in twitter messages using entity linking. Proceedings of the Workshop on Noisy User-generated Text, Beijing, China, pp. 136-140. <https://doi.org/10.18653/v1/w15-4320>
- [30] Tian, T., Dinarelli, M., Tellier, I. (2015). Data adaptation for named entity recognition on tweets with features-rich CRF. Proceedings of the Workshop on Noisy User-generated Text, pp. 68-71. <https://doi.org/10.18653/v1/W15-4309>
- [31] Toh, Z., Chen, B., Su, J. (2015). Improving Twitter named entity recognition using word representations. Proceedings of the Workshop on Noisy User-generated Text, pp. 141-145. <https://doi.org/10.18653/v1/W15-4321>
- [32] Godin, F., Vandersmissen, B., De Neve, W., Van de Walle, R. (2015). Multimedia Lab @ ACL WNUT NER Shared Task: Named entity recognition for twitter microposts using distributed word representations. Proceedings of the Workshop on Noisy User-generated Text, Beijing, China, pp. 146-153. <https://doi.org/10.18653/v1/W15-4322>
- [33] Yang, E., Kim, Y. (2015). Hallym: Named entity recognition on Twitter with word representation. Proceedings of the ACL 2015 Workshop on Noisy User-generated Text, Beijing, China, pp. 72-77.
- [34] Derczynski, L., Augenstein, I., Bontcheva, K. (2015). USFD: Twitter NER with drift compensation and linked data. arXiv e-prints, pp. 48-53.
- [35] Cherry, C., Guo, H., Dai, C. (2015). NRC: Infused phrase vectors for named entity recognition in Twitter. Proceedings of the Workshop on Noisy User-generated Text, Beijing, China, pp. 54-60. <https://doi.org/10.18653/v1/W15-4307>
- [36] Sikdar, U.K., Gambäck, B. (2016). Feature-rich twitter named entity recognition and classification. Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT), Osaka, Japan, pp. 164-170.
- [37] Partalas, I., Lopez, C., Derbas, N., Kalitvianski, R. (2016). Learning to search for recognizing named entities in Twitter. Proceedings of the 2nd Workshop on Noisy User-generated Text, Osaka, Japan, pp. 171-177.
- [38] Le, F.S., Mallek, F., Sadat, F. (2016). UQAM-NL : Named entity recognition in Twitter messages. Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT), Osaka, Japan pp. 197-202.
- [39] Jain, D., Kustikova, M., Darbari, M., Gupta, R., Mayhew, S. (2019). Simple features for strong performance on named entity recognition in code-switched Twitter data. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 103-109. <https://doi.org/10.18653/v1/W18-3213>
- [40] Sikdar, U.K., Barik, B., Gambäck, B. (2018). Named entity recognition on code-switched data using conditional random fields. Proceedings of The Third Workshop on Computational Approaches to Code-Switching, Melbourne, Australia, pp. 115-119.
- [41] Janke, F., Li, T., Rincón, E., Guzmán, G., Bullock, B., Toribio, A.J. (2019). The University of Texas system submission for the code-switching workshop shared task 2018. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 120-125. <https://doi.org/10.18653/v1/W18-3216>
- [42] Claeser, D., Kent, S., Felske, D. (2019). Multilingual named entity recognition on Spanish-English code-switched tweets using support vector machines. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 132-137. <https://doi.org/10.18653/v1/W18-3218>
- [43] Darwish, K., Gao, W. (2014). Simple effective microblog named entity recognition: Arabic as an example. Proceedings of the 9th International Conference on Language Resources and Evaluation, LREC 2014, pp. 2513-2517.
- [44] Zayed, O.H., El-Beltagy, S.R. (2015). Named entity recognition of persons' names in Arabic tweets. International Conference Recent Advances in Natural Language Processing, RANLP, (2015), pp. 731-738.
- [45] Zirikly, A., Diab, M. (2015). Named entity recognition for Arabic social media. Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing, Denver, Colorado, pp. 176-185. <https://doi.org/10.3115/v1/w15-1524>
- [46] Limsopatham, N., Collier, N. (2016). Bidirectional LSTM for named entity recognition in Twitter messages. 2nd Workshop on Noisy User-generated Text, pp. 145-152. <https://doi.org/10.17863/CAM.7201>
- [47] Espinosa, K.J., Batista-Navarro, R., Ananiadou, S. (2016). Learning to recognise named entities in tweets by exploiting weakly labelled data. Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT),

- Osaka, Japan, pp. 153-163.
- [48] Dugas, F., Nichols, E. (2016). DeepNNER: Applying BLSTM-CNNs and extended lexicons to named entity recognition in Tweets. Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT), Osaka, Japan, pp. 178-187.
- [49] Gerguis, M.N., Salama, C., El-Kharashi, M.W. (2016). ASU: An experimental study on applying deep learning in Twitter named entity recognition. Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT), Osaka, Japan, pp. 188-196.
- [50] Winata, G.I., Wu, C., Madotto, A., Fung, P. (2019). Bilingual character representation for efficiently addressing out-of-vocabulary words in code-switching named entity recognition. 3rd Workshop in Computational Approaches in Linguistic Code-switching, pp. 110-114.
- [51] Jansson, P., Liu, S. (2017). Distributed representation, LDA topic modelling and deep learning for emerging named entity recognition from social media. Proceedings of the 3rd Workshop on Noisy User-generated Text, Copenhagen, Denmark, pp. 154-159. <https://doi.org/10.18653/v1/w17-4420>
- [52] Sikdar, U.K., Gambäck, B. (2018). A feature-based ensemble approach to recognition of emerging and rare named entities. Proceedings of the 3rd Workshop on Noisy User-generated Text, Copenhagen, Denmark, pp. 177-181. <https://doi.org/10.18653/v1/W17-4424>
- [53] Lin, B.Y., Xu, F., Luo, Z., Zhu, K. (2018). Multi-channel BiLSTM-CRF model for emerging named entity recognition in social media. Proceedings of the 3rd Workshop on Noisy User-generated Text, Copenhagen, Denmark, pp. 160-165. <https://doi.org/10.18653/v1/w17-4421>
- [54] Rodríguez-Somolinos, A. (2017). From visual perception to inference in the French evidential markers *il m'est avis que*, *apparemment*, and *il paraît que*. *Journal of Historical Linguistics*, 7(1): 111-133. <https://doi.org/10.1075/jhl.7.1-2.05rod>
- [55] Aguilar, G., Carrel, T., Pasic, M., Niederhauser, U., Turina, M. (2017). A multi-task approach for named entity recognition in social media data. Proceedings of the 3rd Workshop on Noisy User-generated Text, Copenhagen, Denmark, pp. 148-153. <https://doi.org/10.18653/v1/W17-4419>
- [56] Trivedi, S., Rangwani, H., Kumar Singh, A. (2019). IIT (BHU) submission for the ACL shared task on named entity recognition on code-switched data. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, pp. 148-153. <https://doi.org/10.18653/v1/W18-3220>
- [57] Wang, C., Cho, K., Kiela, D. (2018). Code-switched named entity recognition with embedding attention. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 154-158. <https://doi.org/10.18653/v1/W18-3221>
- [58] Geetha, P., Chandu, K., Black, A.W. (2019). Tackling code-switched NER: Participation of CMU. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 126-131. <https://doi.org/10.18653/v1/W18-3217>
- [59] Galal, M., Madbouly, M.M., El-Zoghby, A. (2019). Classifying Arabic text using deep learning. *Journal of Theoretical and Applied Information Technology*, 97(23): 3412-3422.
- [60] Helwe, C., Elbassuoni, S. (2019). Arabic named entity recognition via deep co-learning. *Artificial Intelligence Review*, 52(1): 197-215. <https://doi.org/10.1007/s10462-019-09688-6>
- [61] Gridach, M. (2016). Character-aware neural networks for Arabic named entity recognition for social media. 6th Workshop on South and Southeast Asian Natural Language Processing, pp. 23-32.
- [62] Attia, M., Samih, Y., Maier, W. (2019). GHHT at CALCS 2018: Named entity recognition for dialectal arabic using neural networks. Proceedings of the Third Workshop on Computational Approaches to Linguistic Code-Switching, Melbourne, Australia, pp. 98-102. <https://doi.org/10.18653/v1/W18-3212>
- [63] Solorio, T., Blair, E., Maharjan, S., Bethard, S., Diab, M., Ghoneim, M., Hawwari, A., AlGhamdi, F., Hirschberg, J., Chang, A., Fung, P. (2015). Overview for the first shared task on language identification in code-switched data. Proceedings of the First Workshop on Computational Approaches to Code Switching, Doha, Qatar, pp. 62-72. <https://doi.org/10.3115/v1/w14-3907>