

Detecting drugs and adverse events from Spanish health social media streams

Isabel Segura-Bedmar, Ricardo Revert, Paloma Martínez

Computer Science Department,
Carlos III University of Madrid, Spain

{isegura, rrevert, pmf}@inf.uc3m.es

Abstract

To the best of our knowledge, this is the first work that does drug and adverse event detection from Spanish posts collected from a health social media. First, we created a gold-standard corpus annotated with drugs and adverse events from social media. Then, Textalytics, a multilingual text analysis engine, was applied to identify drugs and possible adverse events. Overall recall and precision were 0.80 and 0.87 for drugs, and 0.56 and 0.85 for adverse events.

1 Introduction

It is well-known that adverse drug reactions (ADRs) are an important health problem. Indeed, ADRs are the 4th cause of death in hospitalized patients (Wester et al., 2008). Thus, the field of pharmacovigilance has received a great deal of attention due to the high and growing incidence of drug safety incidents (Bond and Raehl, 2006) as well as to their high associated costs (van Der Hooft et al., 2006).

Since many ADRs are not captured during clinical trials, the major medicine regulatory agencies such as the US Food and Drug Administration (FDA) or the European Medicines Agency (EMA) require healthcare professionals to report all suspected adverse drug reactions. However, some studies have shown

that ADRs are under-estimated due to the fact that they are reported by voluntary reporting systems (Bates et al., 2003; van Der Hooft et al., 2006; McClellan, 2007). In fact, it is estimated that only between 2 and 10 per cent of ADRs are reported (Rawlins, 1995). Healthcare professionals must perform many tasks during their workdays and thus finding the time to use these surveillance reporting systems is very difficult. Also, healthcare professionals tend to report only those ADRs on which they have absolute certainty of their existence. Several medicines agencies have implemented spontaneous patient reporting systems in order for patients to report ADRs themselves. Some of these systems are the MedWatch from the FDA, the Yellow Cards from the UK Medicines agency (MHRA) or the website¹ developed by the Spanish Agency of Medicines and Medical devices (AEMPS). Unlike reports from healthcare professionals, patient reports often provide more detailed and explicit information about ADRs (Herxheimer et al., 2010). Another important contribution of spontaneous patient reporting systems is to achieve patients having a more central role in their treatments. However, despite the fact that these systems are well-established, the rate of spontaneous patient reporting is very low probably because many

¹ <https://www.notificaram.es/>

patients are still unaware of their existence and even may feel embarrassed when describing their symptoms.

In this study, our hypothesis is that health-related social media can be used as a complementary data source to spontaneous reporting systems in order to detect unknown ADRs and thereby to increase drug safety. In recent days, social media on health information, just like has happened in other areas, have seen a tremendous growth (Hill et al., 2013). Examples of social media sites include blogs, online forums, social networking, and wikis, among many others. In this work, we focus on health forums where patients often exchange information about their personal medical experiences with other patients who suffer the same illness or receive similar treatment. Some patients may feel more comfortable sharing their medical experiences with each other rather than with their healthcare professionals. These forums contain a large number of comments describing patient experiences that would be a fertile source of data to detect unknown ADRs.

Although there have been several research efforts devoted to developing systems for extracting ADRs from social media, all studies have focused on social media in English, and none of them have addressed the extraction from Spanish social media. Moreover, the problem is that these studies have not been compared with each other, and hence it is very difficult to determine the current “state-of-art” of the techniques for ADRs extraction from social media. This comparison has not been performed due to the lack of a gold-standard corpus for ADRs. Thus, the goal of our work is twofold: i) to create a gold-standard corpus annotated with drugs and adverse events and ii) to develop a system to automatically extract mentions of drugs and adverse events from Spanish health-related social media sites. The corpus is composed by patients’ comments from Forumclinic², a health online networking website

in Spanish. This is the first corpus of patient comments annotated with drugs and adverse events in Spanish. Also, we believe that this corpus will facilitate comparison for future ADRs detection from Spanish social media.

This is a preliminary work, in which we have only focused on the automatic detection of mentions of drugs and adverse events. Our final goal will be to develop a system to automatically extract drugs and their side effects. We hope our system will be beneficial to *AEMPS* as well as to the pharmaceutical industry in the improvement of their pharmacovigilance systems.

2 Related Work

In recent years, the application of Natural Language Processing (NLP) techniques to mine adverse reactions from texts has been explored with promising results, mainly in the context of drug labels (Gurulingappa et al., 2013; Li et al., 2013; Kuhn et al., 2010), biomedical literature (Xu and Wang, 2013), medical case reports (Gurulingappa et al., 2012) and health records (Friedman, 2009; Sohn et al., 2011). However, as it will be described below, the extraction of adverse reactions from social media has received much less attention.

In general, medical literature, such as scientific publications and drug labels, contains few grammatical and spelling mistakes. Another important advantage is that this type of texts can be easily linked to biomedical ontologies. Similarly, clinical records present specific medical terminology and can also be mapped to biomedical ontologies and resources. Meanwhile social media texts are markedly different from clinical records and scientific articles, and thereby the processing of social media texts poses additional challenges such as the management of meta-information included in the text (for example as tags in tweets) (Bouillot et al., 2013), the detection of typos and unconventional spelling, word shortenings (Neunedert et al, 2013; Moreira et al., 2013) and slang and emoticons (Balahur, 2013), among others. Moreover, these texts are often very short

² <http://www.forumclinic.org>

and with an informal nature, making the processing task extremely challenging.

Regarding the identification of drug names in text, during the last four years there has been significant research efforts directed to encourage the development of systems for detecting these entities. Concretely, shared tasks such as DDIEExtraction 2013 (Segura-Bedmar et al., 2013), CHEMDNER 2013 (Krallinger et al., 2013) or the i2b2 Medication Extraction challenge (Uzuner et al., 2010) have been held for the advancement of the state of the art in this problem. However, most of the work on recognizing drugs concerns either biomedical literature (for example, MedLine articles) or clinical records, thus leaving unexplored this task in social media streams.

Leaman et al., (2010) developed a system to automatically recognize adverse effects in user comments. A corpus of 3,600 comments from the DailyStrength health-related social network was collected and manually annotated with a total of 1,866 drug conditions, including beneficial effects, adverse effects, indications and others. To identify the adverse effects in the user comments, a lexicon was compiled from the following resources: (1) the COSTART vocabulary (National Library of Medicine, 2008), (2) the SIDER database (Kuhn et al., 2010), (3) MedEffect³ and (4) a list of colloquial phrases which were manually collected from the DailyStrength comments. The final lexicon consisted of 4,201 concepts (terms with the same CUI were grouped in the same concept). Finally, the terms in the lexicon were mapped against user comments to identify the adverse effects. In order to distinguish adverse effects from the other drug conditions (beneficial effects, indications and others), the systems used a list of verbs denoting indications (for example, help, work, prescribe). Drug name recognition was not necessary because the evaluation focused only on a set of four drugs: carbamazepine, olanzapine,

trazodone and ziprasidone. The system achieved a good performance, with a precision of 78.3% and a recall of 69.9%.

An extension of this system was accomplished by Nikfarjam and Gonzalez (2011). The authors applied association rule mining to extract frequent patterns describing opinions about drugs. The rules were generated using the Apriori tool⁴, an implementation of the Apriori algorithm (Agrawal and Srikant, 1994) for association rule mining. The system was evaluated using the same corpus created for their previous work (Leaman et al., 2010), and which has been described above. The system achieved a precision of 70.01% and a recall of 66.32%. The main advantage of this system is that it can be easily adapted for other domains and languages. Another important advantage of this approach over a dictionary based approach is that the system is able to detect terms not included in the dictionary.

Benton et al., (2011) created a corpus of posts from several online forums about breast cancer, which later was used to extract potential adverse reactions from the most commonly used drugs to treat this disease: tamoxifen, anastrozole, letrozole and exemestane. The authors collected a lexicon of lay medical terms from websites and databases about drugs and adverse events. The lexicon was extended with the Consumer Health Vocabulary (CHV)⁵, a vocabulary closer to the lay terms, which patients usually use to describe their medical experiences. Then, pairs of terms co-occurring within a window of 20 tokens were considered. The Fisher's exact test (Fisher, 1922) was used to calculate the probability that the two terms co-occurred independently by chance. To evaluate the system, the authors focused on the four drugs mentioned above, and then collected their adverse effects from their drug labels. Then, precision and recall were calculated by comparing the adverse effects from drug labels and the adverse effects obtained by the system.

³ <http://www.hc-sc.gc.ca/dhp-mps/medeff/index-eng.php>

⁴ <http://www.borgelt.net/apriori.html>

⁵ <http://consumerhealthvocab.org>

The system obtained an average precision of 77% and an average recall of 35.1% for all four drugs.

UDWarning (Wu et al., 2012) is an ongoing prototype whose main goal is to extract adverse drug reactions from Google discussions. A knowledge base of drugs and their adverse effects was created by integrating information from different resources such as SIDER, DailyMed⁶, Drugs.com⁷ and MedLinePlus. The authors hypothesized that unknown adverse drug effects would have a high volume of discussions over the time. Thus, the systems should monitor the number of relevant discussions for each adverse drug effect. However, to the best of our knowledge, the UDWarning's component devoted to the detection of unrecognized adverse drug effects has not been developed yet.

Bian et al., (2012) developed a system to detect tweets describing adverse drug reactions. The systems used a SVM classifier trained on a corpus of tweets, which were manually labeled by two experts. MetaMap (Aronson and Lang, 2010) was used to analyze the tweets and to find the UMLS concepts present in the tweets. The system produced poor results, mainly because tweets are riddled with spelling and grammar mistakes. Moreover, MetaMap is not a suitable tool to analyze this type of texts since patients do not usually use medical terminology to describe their medical experiences.

As it was already mentioned, the recognition of drugs in social media texts has hardly been tackled and little research has been conducted to extract relationships between drugs and their side effects, since most systems were focused on a given and fixed set of drugs. Most systems for extracting ADRs follow a dictionary-based approach. The main drawback of these systems is that they fail to recognize terms which are not included in the dictionary. In addition, the dictionary-based approach is not able to handle the large number of spelling and grammar errors in social media texts. Moreover, the detection of

ADRs has not been attempted for languages other than English. Indeed, automatic information extraction from Spanish-language social media in the field of health remains largely unexplored. Additionally, to the best of our knowledge, there is no corpus annotated with ADRs in social media texts available today.

3 Method

3.1 Corpus creation

In order to create the first corpus in Spanish annotated with drugs and adverse events, we reviewed the main health-related social networks in Spanish language to select the most appropriate source of user comments. This corpus will be used to evaluate our system.

Twitter was initially our preferred option due to the tremendous amount of tweets published each day (nearly 400 millions). However, we decided to discard it because Twitter does not seem to be the preferred source for users to describe their ADRs. Gonzalez et al. (2013) gathered a total of 42,327 in a one-month period, from which only 216 described ADRs. Although Facebook is the most popular social media and many Facebook groups dedicated to specific diseases have emerged in the last years, we discarded it because most of these groups usually have restricted access to their members. Online health-related forums are an attractive source of data for our corpus due to their high dynamism, their great number of users as well as their easy access. After reviewing the main health forums in Spanish, we chose ForumClinic, an interactive program for patients, whose main goal is to provide rigorous information about specific diseases (such as breast cancer, HIV, bipolar disorder, depression, schizophrenia, ischemic heart disease, among others) and their treatments. Also, this platform aims to increase the participation of patients maintaining a discussion forum where patients can exchange information about their experiences. Figure 1 shows the distribution of user comments across the main twelve categories defined in the forum. We

⁶ <http://dailymed.nlm.nih.gov/dailymed/>

⁷ <http://www.drugs.com/>

implemented a web crawler to gather all user comments published in ForumClinic to date.

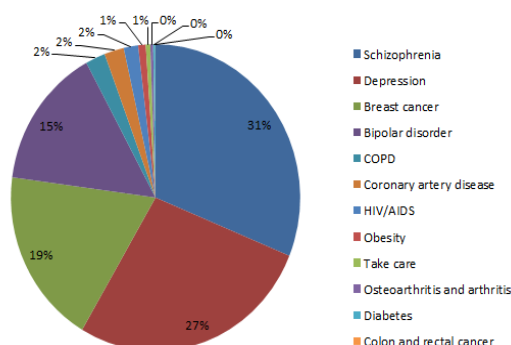


Figure 1 Distribution of user comments.

Then, we randomly selected a sample of 400 comments that were manually labeled with drugs and adverse events by two annotators with expertise in Pharmacovigilance. It should be noted that adverse events and ADRs do not refer to the same: while an adverse event may or may not be caused by a drug, an ADR is an adverse event that is suspected to be caused by a drug. A drug is a substance used in the treatment, cure, prevention or diagnosis of diseases. The corpus includes generic and brand drugs as well as drug families. Disagreements between the annotators were discussed and reconciled during the harmonization process, where a third annotator helped to make the final decision (some examples are shown in Table 1). All the mentions of drugs and adverse events were annotated, even those containing spelling or grammatical errors (for example, *hemorragia*). Nominal anaphoric expressions, which refer to previous adverse events or drugs in the comment, were also included in the annotation. The annotators found 187 drugs (from which 40 were nominal anaphors and 14 spelling errors) and 636 adverse events (from which 48 were nominal anaphors and 17 spelling errors). The corpus is available for academic purposes⁸.

To measure the inter-annotator agreement we used the F-measure metric. This metric approximates the kappa coefficient (Cohen, 1960)

when the number of true negatives (TN) is very large (Hripcsak and Rothschild, 2005). In our case, we can state that the number of TN is very high since TN are all the terms that are not true positives, false positives nor false negatives. The F-measure was calculated by comparing the two corpora created by the two first annotators. The corpus labelled by the first annotator was considered the gold-standard. As it was expected, drugs exhibit a high IAA (0.89), while adverse events point to moderate agreement (0.59). As drugs have specific names and there are a limited number of them, it is possible to create a limited and controlled vocabulary to gather many of the existing drugs. On the other hand, patients can express their adverse events in many different ways due to the variability and richness of natural language.

Sentence	Final Decision
<i>De entre los distintos antiretrovirales, transcriptasa inversa, proteasa, integrasa y fusión, qué grupo sería el más potente y cual el menos.</i>	Names in bold type refer to four families of inhibitors (that is, drug families), and thereby, they should be annotated.
<i>Como complemento proteico recomendamos el de los laboratorio <i>Vegenat</i>. Si compras los <u>complementos del Decathlon</u>, asegúrate que contenga proteínas.</i>	The mention “complementos del Decathlon” should not be annotated as a drug since it is not a brand-marked drug.

Table 1: Some examples of disagreements between annotators

3.2 Constructing a dictionary for drugs and adverse events

Since our goal is to identify drugs and adverse events from user comments, the first challenge is to create a dictionary that contains all of the drugs and known adverse events.

CIMA⁹ is an online information center about medicines that provides all the daily updated official information about drugs. CIMA is

⁸ <http://labda.inf.uc3m.es/SpanishADRCorpus>

⁹ <http://www.aemps.gob.es/cima/>

maintained by the Spanish Agency for Medicines and Health Products (AEMPS). It includes information on all drugs authorized in Spain and their current authorization status. CIMA contains a total of 16,418 brand drugs and 2,228 generic drugs. Many brand drug names include additional information such as dosages, mode and route of administration, laboratory, among others (for example, “ESPIDIFEN 400 mg GRANULADO PARA SOLUCION ORAL SABOR ALBARICOQUE” or “ESPIDIFEN 600 mg GRANULADO PARA SOLUCION ORAL SABOR LIMON EFG, 20 sobres”). Since it is unlikely that these long names are used by patients, we implemented a method to shorten them by removing their additional information (for example, “ESPIDIFEN”). After applying this method, the resulting list of brand drug names consisted of 3,662 terms. The main limitation of CIMA is that it only provides information about drugs authorized in Spain. That is, CIMA does not contain information about drugs approved only in Latin America. CIMA is free and offers a downloadable version in XML format. Thus, it provides the information in a well-structured format that makes it possible to directly extract generic and brand drug names as well as other related information such as their ATC codes, their pharmaceutical company, among others. Unfortunately, CIMA does not provide information about drug groups. For this reason, we decided to consider the WHO ATC system¹⁰, a classification system of drugs, as an additional resource to obtain a list of drug groups.

MedDRA¹¹ is a medical terminology dictionary about events associated with drugs. It is a multilingual terminology, which includes the following languages: Chinese, Czech, Dutch, French, German, Hungarian, Italian, Japanese, Portuguese and Spanish. Its main goal is to provide a classification system for efficient communication of ADRs data between countries. The main advantage of MedDRA is that its

¹⁰ http://www.whooc.no/atc_ddd_index/

¹¹ <http://www.meddra.org/>

structured format allows easily obtaining a list of possible adverse events. MedDRA is composed of a five levels hierarchy. We collected the terms from the most specific level, "Lowest Level Terms" (LLTs)". This level contains a total of 72,072 terms, which express how information is communicated in practice.

By analyzing the information from these resources, we found that none of them contained all of the drugs and adverse events. Patients usually use lay terms to describe their symptoms and their treatments. Unfortunately, many of these lay terms are not included in the above mentioned resources. Therefore, we decided to integrate additional information from other resources devoted to patients to build a more complete and comprehensive dictionary. There are several online websites that provide information to patients on drugs and their side effects in Spanish language. For example, MedLinePlus and Vademecum contain information about drugs and their side effects. These websites allow users to browse by generic or drug name, providing an information leaflet for each drug in a HTML page. Since these leaflets are unstructured, the extraction of drugs and their adverse effects is a challenging task. While drug names are often located in specific fields (such as title), their adverse events are usually descriptions of harmful reactions in natural language. We only developed a web crawler to browse and download pages related to drugs from Vademecum since this website provided an easier access to its drug pages than MedLinePlus. We plan to augment the list of drugs and adverse events by crawling MedLinePlus in future work.

After extracting drugs and adverse events from these different resources, we created a dictionary of drugs and adverse events. Table 2 shows the statistics of our final dictionary.

Resource	Total
Generic drugs from CIMA	2,228
Brand drugs from CIMA	3,662

<i>Drug group names from the ATC system</i>	466
<i>Drug names (which are not in CIMA) from Vademecum</i>	1,237
<i>Total Drugs:</i>	7,593

Table 2: Number of drugs in the dictionary.

Resource	Total
Adverse events from MedDRA	72,072
<i>Adverse events from Vademecum (which are not in MedDRA)</i>	2,793
<i>Total adverse events:</i>	74,865

Table 3: Number of adverse events in the dictionary.

3.3 Using Textalytics and gazetteers to identify drugs and adverse events

Textalytics¹² is a multilingual text analysis engine to extract information from any type of texts such as tweets, posts, comments, news, contracts, etc. This tool offers a wide variety of functionalities such as text classification, entity recognition, concept extraction, relation extraction and sentiment analysis, among others. We used a plugin that integrates Textalytics with GATE. In this paper, we applied entity recognition provided by Textalytics, which follows a dictionary-based approach to identify entities in texts. We created a dictionary for drugs and adverse events from CIMA and MedDRA. This dictionary was integrated into Textalytics. Additionally, the lists of drugs and adverse events collected from the others resources (ATC system and Vademecum) were used to create GATE gazetteers.

4 Results and error analysis

We evaluated the system on the corpus annotated with drugs and adverse events. The results of this study show a precision of 87% for drugs and 85% for adverse events, and a recall of 80% for drugs and 56% for adverse events.

We performed an analysis to determine the main sources of error in the system. A sample of 50 user comments were randomly selected and analyzed. Regarding the detection of adverse events, the major cause of false negatives was the use of colloquial expressions to describe an adverse event. Phrases like “*me deja ko* (it makes me KO)” or “*me cuesta más levantarme* (it’s harder for me to wake up)” were used by patients for expressing their adverse events. These phrases are not included in our dictionary. A possible solution may be to create a lexicon containing this kind of idiomatic expressions. The second highest cause of false negatives for adverse events was due to the different lexical variations of the same adverse event. For example, ‘*depresión* (depression)’ is included in our dictionary, but their lexical variations such as “*depremido* (depress)”, “*me deprimio* (I get depressed)”, “*depresivo* (depressive)” or “*deprimente* (depressing)” were not detected by our system since they are not in our dictionary. Nominalization may be used to identify all the possible lexical variations of a same adverse event. Another important error source of false negatives was spelling mistakes (eg. *hemorragia* instead of *hemorragia*). Many users have great difficulty in spelling unusual and complex technical terms. This error source may be handled by a more advanced matching method capable of dealing with the spelling error problem. The use of abbreviations (“*depre*” is an abbreviation for “*depression*”) also produces false negatives. Techniques such as lemmatization and stemming may help to resolve this kind of abbreviations.

False positives for adverse events were mainly due to the inclusion of MedDRA terms referring to procedures (such as therapeutic, preventive or laboratory procedures) and tests in our dictionary. MedDRA includes terms for diseases, signs, abnormalities, procedures and tests. We should have not included those terms referring to procedures and tests since they do not represent adverse events.

¹² <https://textalytics.com/>

The main source of false negatives for drugs seems to be that users often misspelled drug names. Some generic and brand drugs have complex names for patients. Some examples of misspelled drugs are *avilify* (*Abilify*) or *rivotril* (*ribotril*). Another important cause of false negatives was due to the fact that our dictionary does not include drugs approved in other countries than Spain (for example, *Clorimipramina*, *Ureadin* or *Paxil*). However, ForumClinic has a large number of users in Latin America. It is possible that these users have posted comments about some drugs that have only been approved in their countries. The third largest source of errors was the abbreviations for drug families. For instance, *benzodiacepinas* (*benzodiazepine*) is commonly used as *benzos*, which is not included in our dictionary. An interesting source of errors to point out is the use of acronyms referring to a combination of two or more drugs. For instance, *FEC* is a combination of *Fluorouracil*, *Epirubicin* and *Cyclophosphamide*, three chemotherapy drugs used to treat breast cancer. This combination of drugs is not registered in the resources (CIMA and Vademecum) used to create our dictionary.

Most false positives for drugs were due to a lack of ambiguity resolution. Some drug names are common Spanish words such as “*Alli*” (a slimming drug) or “*Puntual*” (a laxative). These terms are ambiguous and resolve to multiple senses, depending on the context in which they are used. Similarly, some drug names such as “alcohol” or “oxygen” can take a meaning different than the one of pharmaceutical substance. Another important cause of false positives is due to the use of drug family names as adjectives that specify an effect. This is the case of *sedante* (*sedative*) or *antidepressivo* (*antidepressant*), which can refer to a family of drugs, but also to the definition of an effect or disorder caused by a drug (*sedative effects*).

5 Conclusion

In this research, we created the first Spanish corpus of health user comments annotated with drugs and adverse events. The corpus is available

for research. In this work, we only focused on the detection of the mentions of drugs and adverse events, but not the relationships among them. In future work, we plan to extend the system to detect the relationships between drugs and their side effects. Also, we would like to identify their indications and beneficial effects.

Acknowledgments

This work was supported by the EU project TrendMiner [FP7-ICT287863], by the project MULTIMEDICA [TIN2010-20644-C03-01], and by the Research Network MA2VICMR [S2009/TIC-1542].

References

- Rakesh Agrawal and Ramakrishnan Srikant. 1994. Fast algorithms for mining association rules. In *Proc. 20th Int. Conf. Very Large Data Bases*, 1215:487-499.
- Alan R Aronson and Francois-Michel Lang. 2010. An overview of MetaMap: historical perspective and recent advances. *Journal of the American Medical Informatics Association*, 17(3):229-236.
- Alexandra Balahur. 2013. Sentiment Analysis in Social Media Texts. *WASSA 2013*, 120.
- David W. Bates, R Scott Evans, Harvey Murff, Peter D. Stetson, Lisa Pizziferri and George Hripesak. 2003. Detecting adverse events using information technology. *Journal of the American Medical Informatics Association*, 10(2):115-128.
- Adrian Benton, Lye Ungar, Shawndra Hill, Sean Hennessy, Jun Mao, Annie Chung, Charles E. Leonarda and John H. Holmes. 2011. Identifying potential adverse effects using the web: A new approach to medical hypothesis generation. *Journal of biomedical informatics*, 44(6): 989-996.
- Jiang Bian, Umit Topaloglu and Fan Yu. 2012. Towards large-scale twitter mining for drug-related adverse events. In *Proceedings of the 2012 international workshop on Smart health and wellbeing*, 25-32.
- CA. Bond and Cynthia L. Raehl. 2006. Adverse drug reactions in United States hospitals. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 26(5):601-608.

- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychol Meas* ;20:37e46.
- Ronald A. Fisher. 1922. On the interpretation of χ^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society*, 85(1):87-94.
- Flavien Bouillot, Phan N. Hai, Nicolas Béchet, Sandra Bringay, Dino Ienco, Stan Matwin, Pascal Poncelet, Mathieu Roche and Maguelonne Teisseire. 2013. How to Extract Relevant Knowledge from Tweets?. *Communications in Computer and Information Science*.
- Carol Friedman. 2009. Discovering novel adverse drug events using natural language processing and mining of the electronic health record. In *Artificial Intelligence in Medicine*. LNAI 5651:1 -5.
- Graciela H. Gonzalez, Matthew L Scotch and Garrick L Wallstrom. Mining Social Network Postings for Mentions of Potential Adverse Drug Reactions. HHS-NIH-NLM (9/10/2012 - 8/31/2016).
- Harsha Gurulingappa, Abdul Mateen-Rajput and Luca Toldo. 2012. Extraction of potential adverse drug events from medical case reports. *Journal of biomedical semantics*. 3(1):15.
- Harsha Gurulingappa, Luca Toldo, Abdul Mateen-Rajput, Jan A. Kors, Adel Taweel and Yorki Tayrouz. 2013. Automatic detection of adverse events to predict drug label changes using text and data mining techniques. *Pharmacoepidemiology and drug safety*, 22(11):1189-1194.
- A Herxheimer, MR Crombag and TL Alves. 2010. Direct patient reporting of adverse drug reactions. A twelve-country survey & literature review. *Health Action International (HAI). Europe*. Paper Series Reference 01-2010/01.
- Shawndra Hill, Raina Merchant and Lile Ungar. (2013). Lessons Learned About Public Health from Online Crowd Surveillance. *Big Data*, 1(3):160-167.
- George Hripcsak and Adam S. Rothschild. 2005. Agreement, the F-measure, and reliability in information retrieval. *J Am Med Inform Assoc*.12:296e8.
- Martin Krallinger, Florian Leitner, Obdulia Rabal, Miguel Vazquez, Julen Oyarzabal and Alfonso Valencia. 2013. Overview of the chemical compound and drug name recognition (CHEMDNER) task. In *BioCreative Challenge Evaluation Worksho*. 2:2-33.
- Michael Kuhn, Monica Campillos, Ivica Letunic, Lars J. Jensen and Peer Bork. 2010. A side effect resource to capture phenotypic effects of drugs. *Molecular systems biology*, 6(343):1-6.
- Robert Leaman, Laura Wojtulewicz, Ryan Sullivan, Annie Skariah, Jian Yang and Graciela Gonzalez. 2010. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In *Proceedings of the 2010 workshop on biomedical natural language processing*. 117-125. Association for Computational Linguistics.
- Anne J. Leendertse, Antoine C. Egberts, Lennar J. Stoker, & Patricia M.L.A. van den Bemt. 2008. Frequency of and risk factors for preventable medication-related hospital admissions in the Netherlands. *Archives of internal medicine*, 168(17), 1890.
- Qi Li, Louise Deleger, Todd Lingren, Haijun Zhai, Megan Kaiser, Laura Stoutenborough Anil G Jegga, Kevin B Cohen and Imre Solti. 2013. Mining FDA drug labels for medical conditions. *BMC medical informatics and decision making*, 13(1):53.
- Mark McClellan. 2007. Drug Safety Reform at the FDA-Pendulum Swing or Systematic Improvement?. *New England Journal of Medicine*, 356(17):1700-1702.
- Silvio Moreira, Joao Filgueiras, Bruno Martins, Francisco Couto and Mario J. Silva. 2013. REACTION: A naive machine learning approach for sentiment classification. In *2nd Joint Conference on. Lexical and Computational Semantics*. 2:490-494.
- Melanie Neunerdt, Michael Reyer and Rudolf Mathar. 2013. A POS Tagger for Social Media Texts trained on Web Comments. *Polibits*, 48:59-66.
- Azadeh Nikfarjam and Graciela H. Gonzalez. 2011. Pattern mining for extraction of mentions of adverse drug reactions from user comments. In *AMIA Annual Symposium Proceedings*, 2011:1019-1026. American Medical Informatics Association.
- Isabel Segura-Bedmar, Paloma Martinez and María Herrero-Zazo. 2013. SemEval-2013 Task 9: Extraction of Drug-Drug Interactions from

- Biomedical Texts (DDIExtraction 2013). 3206(65): 341-351.
- Cornelis S. van Der Hooft, Miriam CJM Sturkenboom, Kees van Grootheest, Herre J. Kingma and Bruno HCh Stricker. 2006. Adverse drug reaction-related hospitalisations. *Drug Safety*, 29(2):161-168.
- Hamish Cunningham. 2002. GATE, a general architecture for text engineering. *Computers and the Humanities*, 36(2):223-254.
- M Rawlins. 1995. Pharmacovigilance: paradise lost, regained or postponed? The William Withering Lecture 1994. *Journal of the Royal College of Physicians of London*, 29(1): 41-49.
- Sunghwan Sohn, Jean-Pierre A. Kocher, Christopher G. Chute and Guergana K. Savova. 2011. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. *Journal of the American Medical Informatics Association*, 18(Suppl 1):i144-i149.
- Özlem Uzuner, Imre Solti and Eithon Cadag. 2010. Extracting medication information from clinical text. *Journal of the American Medical Informatics Association*. 17(5):514-518.
- Rong Xu and QuanQiu Wang. 2013. Large-scale extraction of accurate drug-disease treatment pairs from biomedical literature for drug repurposing. *BMC Bioinformatics*, 14(1):181.
- Karin Wester, Anna K. Jönsson, Olav Spigset, Henrik Druid and Staffan Hägg. 2008. Incidence of fatal adverse drug reactions: a population based study. *British journal of clinical pharmacology*, 65(4):573-579.