# Annotating Spin in Biomedical Scientific Publications: the case of Randomized Controlled Trials (RCTs)

**Anna Koroleva, Patrick Paroubek**

LIMSI-CNRS,

Bât. 508, rue John von Neumann, campus Université d'Orsay, F-91405 Orsay

{koroleva, pap}@limsi.fr

### Abstract

In this paper we report on the collection in the context of the MIROR project of a corpus of biomedical articles for the task of automatic detection of inadequate claims (spin), which to our knowledge has never been addressed before. We present the manual annotation model and its annotation guidelines and describe the planned machine learning experiments and evaluations.

**Keywords:** spin, annotation scheme, biomedical articles

## 1. Introduction

Merriam Webster dictionary defines spin doctor as *"a person (such as a political aide) responsible for ensuring that others interpret an event from a particular point of view"*[1]. In the context of the MIROR[2] project, we address spin in biomedical scientific publications, where it refers to misleading presentation of scientific results, in particular in articles reporting randomized controlled trials (RCTs), an important type of clinical trial. In our case, spin consists in presenting the examined treatment as having greater beneficial effects than the experiments show (Boutron et al. 2010; Boutron et al. 2014; Haneef et al. 2015; Yavchitz et al. 2016 ). Spin in RCTs affects clinical decision-making (Boutron et al. 2014) and results in distorted presentation of research findings in media (Yamamoto & Takagi 2005; Hall et al. 2009). We present here the first steps aiming at proposing an algorithm for automatic spin identification in biomedical abstracts, something which to the best of our knowledge has not been attempted before. We present here the construction and annotation of a corpus of medical publication extracted from PubMed Central[3] (PMC) about RCT and describe the annotation model and guidelines.

### 1.1 On spin types

From previous research on spin classification (Boutron et al. 2010; Lazarus et al. 2015, Yavchitz et al. 2012], we can outline three main types and their subtypes of spin in RCT reports:

1. misleading reporting of study results:

- selective reporting of outcomes (omission of the primary outcome; focus on statistically significant results different from the main outcome);

- occulting adverse events;

- misleading reporting of study design;

- linguistic spin (beautifying formulations);

- discarding limitations;

- selective citation of other studies

2. inadequate interpretation of the results:

- claiming a beneficial effect of the intervention despite statistically non-significant results;

- claiming an equivalent effect of the interventions for statistically non-significant results;

- claiming that the treatment is safe for statistically non-significant safety outcomes;

- concluding a beneficial effect despite no comparison test performed;

- interpretation of the results according to statistical significance instead of clinical relevance;

3. inadequate extrapolation:

- inadequate extrapolation from the population, interventions or outcome actually assessed in the study to a larger population, different interventions or outcomes;

- inadequate implications for clinical practice.

Example of spin putting focus on secondary result ("*improved PFS and response for treatment*") instead of the main result, object of the experiment ("*survival rate*"):

> *"This study demonstrates <u>improved PFS and response for the treatment</u> A compared with comparator B, although this did **not result in improved survival** "*.

Fig. 1. Example of spin (focus on secondary result)

In the rest of this paper, we present our linguistic model of spin (section 2), the annotation scheme (section 3), the the annotation guidelines (section 4), conclusions and plans for future work (section 5).

## 2. Model of spin

To the best of our knowledge, this is the first attempt at addressing the analysis of spin in the biomedical literature from a Natural Language Processing point of view. Spin detection overlaps partially with previous works in NLP, in particular objectivity/subjectivity identification (Wiebe

---

[1]https://www.merriam-webster.com/dictionary

[2]http://miror-ejd.eu/

[3]https://www.ncbi.nlm.nih.gov/pmc/

et al. 2005), sentiment analysis (Pak 2012), fact checking (Nakashole & Mitchell 2014) or deception detection (Hancock et al. 2010; Litvinova et al. 2017); a point to note is that these works address texts of general domain while we deal with spin in biomedical texts. We regard spin detection as a task most closely related to deception detection. Deception is defined as a deliberate act of communicating information that the speaker/author believes to be false, with the intention to induce listeners/readers to believe a distorted presentation of the topic. Strictly speaking, spin is not necessarily a form of deception, as the intention is difficult to establish most of the time, e.g., spin in abstracts may be conditioned by limited space; by author's wish to report the results that he/she perceives to be most important; by unclear/absent reporting guidelines; by lack of training etc. However, spin is similar to deception for what concerns its impact and the method required to detect it from textual content only (Mihalcea et al. 2009).

Spin can be considered as the most serious form of incomplete or incoherent reporting of trial data and results (omission or inconsistent presentation of information). We aim at creating a general model that would be able to represent the information about a trial data and all possible realizations of spin in reporting.

For trial data, we choose to follow the information structure accepted in trial registries (official catalogues for registering clinical trials, containing in a structured form trial data provided by the investigators who carry out a trial).



| Start Date ICMJE | September 2006 |
| Primary Completion Date | *Not Provided* |
| Current Primary Outcome Measures ICMJE (submitted: June 8, 2007) | Sleep quality assessed by the Leed Sleep Evaluation Questionnaire, measured after one month of treatment |
| Original Primary Outcome Measures ICMJE | *Same as current* |
| Change History | No Changes Posted |
| Current Secondary Outcome Measures ICMJE (submitted: June 8, 2007) | • Sleep quality assessed by the Leed Sleep Evaluation - Questionnaire, measured 15 days after treatment withdrawal<br>• Sleep quality assessed by a Sleep diary completed each day during all the study period by the participants<br>• Sleep efficiency measured by ambulatory actigraphy (2 period)<br>• Evolution of melatonin/6sulfatoxymelatonin ratio (before and after treatment)<br>• Clinical General Impression of the clinician, before and after treatment.<br>• Safety of the treatment (adverse event reporting) |

Fig. 2. Excerpt from an RCT description queried on ClinicalTrials.gov with the keywords: '*RCT insomnia France*'

Trial registries may slightly vary regarding the level of detalisation used for information presentation, so we reviewed several registries (ClinicalTrials.gov[4], ISRCTN[5], plus some national registries) and generalized the categories used. We compiled the following list of data describing a trial:

- Information about interventions: intervention name, dosage, administration schedule, treatment duration;

- Information about participants: age, gender, health condition, health type, nationality/ethnicity, recruitment country/region; information regarding intervention assigned;

other information. Can be represented in a form of a list of inclusion and exclusion criteria, that can contain all of the above information;

- Trial methods / trial design: allocation concealment, allocation type, blinding, sample sizes for groups examined, study type, study subtype, trial phase, statistical tests used;

- Trial objectives / outcomes (with their methods of measurement and associated time points);

- Data about registration: registration number, registration time;

- Financing: sponsors;

- Hypothesis, hypothesis type;

- General information: medical domain;

- Summary.

We also introduced some other categories that are not typically present in registries but that are relevant to trial description: limitations and reported statistical measures

In order to be able to capture instances of spin, we further need to reflect the following phenomena:

1. Incomplete reporting, which can take many forms, but we are most interested in omission of information that is normally supposed to be present in a well-reported abstract, such as:

- clear definition of the primary outcome;

- results for primary outcome;

- results for non-significant secondary outcomes;

- information about adverse events (their absence should be stated explicitly).

Omission of some other types of information (design, methods, statistical tests used, etc.) should not be considered as spin but rather as incomplete reporting acting as 'spin facilitator' hindering fact checking.

2. Incoherent reporting:

- primary outcome described in the trial registry differs from the primary outcome described in the text;

- patient population reported in the abstract does not correspond to the population studied in its qualitative characteristics (age, gender, etc.);

- reported results do not correspond to trial design;

- the compared treatments are reported to be similar when the design does not allow to conclude on similarity (i.e. the trial is not a 'non-inferiority' of 'equivalence' trial);

- within-group comparison reported when the trial objective was not to examine changes within groups (i.e. the trial is not a 'before-and-after trial');

- focus on significant secondary outcomes instead of primary outcome;
- positive conclusions are made (efficacy stated, treatment recommended for use) when the primary outcome is not significant.

Incoherence or incompleteness of reporting can be established by checking the completeness of the abstract, discrepancies between abstract and article body or between trial registry entry (if available) and article content. We thus work with two types of documents: articles and registry entries. For articles, the model comprises information about its structure: its division into title, abstract and body text, for registries we rely on their internal structure, in general a tabular form holding short pieces of text or data.

## 3. Annotation scheme

We proposed a description of an algorithm of spin detection elsewhere (Koroleva & Paroubek 2017a). The main steps are the following:

- dividing a given article into title, abstract and body text; finding results and conclusions within the abstract;
- identifying positive evaluation of the studied treatment in results/conclusions of the abstract;
- extracting elements of trial data relevant to spin assessment, such as outcomes, patient population, statistical significance of results;
- extracting relation between elements of trial data, such as an outcome and its statistical significance;
- extracting specific constructions possibly related to spin (see below);
- final assessment of spin: checking if the information in the results and conclusions of the abstract corresponds to the extracted trial data, for example, if the pre-defined outcomes are reported correctly or if the positive evaluation of the treatment is supported by statistically significant results.

We propose here an annotation scheme comprising the information elements relevant for the future algorithm. Our annotation scheme is implemented in XML and includes several levels of information:

1. Document type (article/registry entry).

2. Structural information (for articles). For this annotation level we adopt the existing annotation scheme used in PubMed[6], simplified for our needs. Our scheme includes journal name, article title, authors list, abstract, body text, bibliography. Within abstracts, Results and Conclusions sections are marked.

3. Elements describing the trial (what was studied and how: compared interventions, outcomes, population

6 https://www.ncbi.nlm.nih.gov/corehtml/query/DTD/index.shtml

studied, statistical measures used, etc.): we introduce a separate tag for each type of trial data. This decision is motivated by the fact that we need specific sets of attributes for different types of trial data, and we need to introduce particular relations for specific types of trial information. As outcome is the most important type of trial data for spin detection, for outcomes (or trial objectives) we use several tags that are needed to distinguish between different specific constructions:

The <Prol>primary outcome measure will be</Prol> <Out type="Prim" Status="Declared">**QoL**</Out>, assessed with the ALS Assessment Questionnaire...

Fig. 3. Example of annotation for a primary (attribute type is Prim) outcome (Out) explicitly declared, with the annotation of its linguistic marker (Prol).

The type of an outcome can have three different type attribute values: Prim (primary) / Sec (secondary) / None (undefined). Outcome has also an attribute 'status' which can have two values: Declared when it is explicitly stated in the text to be an outcome (e.g. Fig 3), which is its value by default and Reported, when the outcome is only reported in results or conclusion section without referring explicitly to its nature.

Our <Prol>secondary aim is</Prol> <Aim type="Sec">**to describe the costs**</Aim> associated with RESERVE-DSD.

Fig. 4. The AIM is the objective of the trial.

4. Relations between elements of trial data: relations that link a pair of elements that describe different features of a single concept, e.g. an outcome with its method of measurement or with its time points, or an intervention to its dosage, administration schedule, etc.

5. Particular constructions of interest:

- Positive evaluations of treatment (positive results regarding the treatment);
- Statements of similarity between treatments regarding their efficacy or safety;
- Within-group comparisons (statements of changes that occur within a group receiving the studied treatment, without comparing it to the group receiving the control treatment);
- Recommendations to use treatment.

These include: i) an analysis which shows that the ethnic difference in performance in this 2006/7 <Subj>**cohort of Year 3 students**</Subj> was similar in size to that in <Subj>**previous cohorts on the course**</Subj> [see Additional file 2 ]

Fig. 5. Example of a similarity statement. Subj – trial subjects.

A problem that arises with this type of information consists in deciding which fragment of text should be

annotated. Normally these constructions comprise a whole proposition, but we can as well highlight some words that are the most 'representative' of the meaning of each construction. We choose to annotate the smallest possible fragments that are indicators of relevant constructions. The motivation behind this decision is the need to make the annotation as clear and simple as possible for annotators, and the fact that, having annotation on word level, we can easily expand it to the sentence level.

6. Annotation for spin: annotation level that is meant to capture all the cases of incoherence and incompleteness regarding the types of information enumerated above. This type of annotation resembles most to a well-known task of relation annotation, but here the most important is not to capture relation that holds between two text fragments, but to mark the cases when there is no relation when we expect it to exist. For example, a relation between a declared primary outcome in article text or protocol and a corresponding reported outcome in abstract means no spin, but a declared primary outcome with no related reported outcome is a case of spin. A similarity statement is not spin if the trial was of equivalence type, but it is a spin if there is no text fragment indicating that the trial belongs to equivalence trials.

To annotate this information, we follow the system accepted in TimeML (Pustejovsky et al. 2003) annotation for relations: we introduce empty tags that contain reference to IDs of fragments that are linked in case of good reporting; in case of incoherence/incompleteness, the tag contains ID of the present text fragment. These tags have an attribute 'spin' that is set to 'yes' or 'no'.

Another form of actual spin or of 'spin facilitator' is omitting some information about methods, design or results in the abstract, e.g. not stating clearly the primary outcome. For this type of omission, we do not need to refer to an ID, we only need an empty tag to mark which type of information is omitted in abstract.

Thus, the annotation for spin is done on the lowest level: as a relation between text fragments. We can then calculate the value of 'spin' attribute for the whole text.

Figure 6 shows an example of text with spin (the example comes from the appendix of Boutron et al. 2014) and the process of assessment of outcome-related spin.

<Abstract>Abstract

<...>

<Res>Results

<...> <Out ID="1" Type="None" Status="Reported">**The International Union Against Cancer R0 resection rate**</Out> was 81.9% after treatment A as compared with 66.7% with surgery alone (P = .036). The surgery-only group had more <Out ID="2" Type="None" Status="Reported">**lymph node metastases**</Out> than the treatment A group (76.5% v 61.4%; P = .018). <...> A <Out ID="3" Type="None" Status="Reported">**survival**</Out> benefit could not be shown (hazard ratio, 0.84; 95% CI, 0.52 to 1.35; P = .

466). </Res>

<Concl>Conclusion

This trial showed a significantly increased <Out ID="4" Type="None" Status="Reported">**R0 resection rate**</Out> but failed to demonstrate a <Out ID="5" Type="None" Status="Reported">**survival**</Out> benefit. </Concl>

<BodyText> <...>

The primary end point of this trial was <Out ID="6" Type="Prim" Status="Declared">**overall survival**</Out>. <...></BodyText>

Fig. 6. Example of annotation of spin for an abstract

The first step in annotating this text would be to annotate all outcomes reported in the abstract (IDs 1 − 5) and the declared primary outcome (ID 6). The following steps to fully annotate all types of spin related to primary outcome would be the following:

1) Check and mark if there is a definition of the primary outcome in the abstract. Here it is absent (full text if abstract omitted for the sake of space) − we conclude incomplete reporting.

2) Check and mark if the declared primary outcome is present among the reported outcomes. Here it can be considered to correspond to the outcomes 3 and 5 − we conclude correct reporting.

3) Check and mark if the primary outcome is presented correctly according to its importance: it should be presented in the first place without regard to significance of results; there should be no focus on other outcomes. In this abstract, the insignificant primary outcome is presented after significant secondary ones − we conclude incoherent reporting (focus on secondary outcomes).

## 4. Annotation guidelines

We plan to combine automatic annotation as first stage, and manual annotation to correct and complete the annotation. We do not aim at manually annotating all the types of information. Most of the trial data not directly relevant to spin detection will be marked automatically only in trial registry entries, where information is highly structured. We do not thus cover them in the annotation guidelines.

We described our algorithms of automatic pre-annotation in our previous works (Koroleva & Paroubek 2017a, Koroleva & Paroubek 2017b). These algorithms aim at extracting/annotating the following:

- text structure: separating results and conclusiong sections in abstracts;

- various constructions defining trial outcomes, with special attention to the primary one, for example:

1. The primary outcome is <Out Type="Prim" Status="Declared"> emotional distress ( symptoms of depression , anxiety , diabetes-specific stress , and general perceived stress ) </Out> .

2. This project has one primary aim : to measure <Out Type="Prim" Status="Declared"> the impact of continuity of midwifery care </Out> compared to routine care on restricting excessive gestational weight gain in obese women.

3. Sample size A power calculation was carried out for the primary outcome (<Out Type="Prim" Status="Declared"> health related quality of life measured on the York version of the SF-12</Out>).

- comparative constructions that are often used to report the trial results. These constructions usually include some of the following elements: compared patient groups, compared treatments, outcomes that serve as basis for comparison. We mainly focus on extracting outcomes:

1. <Subj> Patients with TC asthma </Subj> has significantly higher<Out Type="None" Status="Reported"> AQLQ scores </Out> compared to those with NTC asthma.

2. Muscarinic agonists appear to reduce <Out Type="None" Status="Reported"> the potency of beta-agonist bronchodilation </Out> , possibly through an effect on adenylyl cyclase 17.

3. <Out Type="None" Status="Reported"> Levels of hs-CRP </Out> increased modestly in the ABC / 3TC arm compared with the TDF / FTC the arm.

- Description of studied population:

1. We studied <Subj> <Aim Type="None"> 19 consecutive unselected patients who met the ARDS criteria of the American European Consensus Conference 21. </Aim> </Subj>

2. A total of <Subj> 32 patients aged 12 to 17 years with severe , active and refractory JoAS </Subj> were enrolled in a multicenter , randomized , double-blind , placebo-controlled parallel study of 12 weeks.

These annotations, although not perfectly correct and complete, are hoped to reduce workload for annotators: in case pre-annotation is completely correct or completely erroneous, the annotators will simply need to validate/reject it, reducing the number of cases requiring manual annotation.

The current pre-annotated corpus includes 3938 articles on randomized controlled trials in various medical domains, extracted from PubMed Central. This corpus will serve as basis for manual annotation.

We will split manual annotation into several stages that would differ regarding their complexity and thus the skills required from the annotators.

Some of the tasks are relatively easy and can be done by annotators who do not have special knowledge in medical domain. We consider that the tasks that fall into this group are: explicit descriptions of outcomes, mentions of patient population, statistical measures (p-value), confidence intervals.

Some other types of information require some special knowledge of medical domain as understanding of medical terms is needed to correctly interpret the meaning of sentences and categorize text fragments as representing a certain type of trial data/construction. Following tasks fall into this category: reported outcomes, similarity statements, within-group comparisons, evaluations related to treatment.

The final task of spin annotation (i.e. marking parts of the text that represent coherent and complete reporting for chosen concepts, and marking cases when there is incoherence/incompleteness) is an even more difficult task. The concept of spin in biomedical domain is not completely formally defined yet, experts in the domain often disagree on classifying a certain phenomenon as spin or not. For example, some experts regard absence of explicit definition of the primary outcome in the abstract of an article as definite spin, while others consider it to represent incomplete reporting but less important than spin. Besides, mismatch between information in the abstract and in the article (e.g. change of outcomes studied and reported) is not spin if it has valid scientific justification, which should be provided in the article. Extraction of such justifications and assessment of their validity would be necessary to conclude on absence or presence of spin, but it falls outside scope of our work.

Thus, there are several difficulties that we should take into account when developing annotation guidelines:

1) Some of the tasks require at least some level of special medical knowledge, so it is likely that the annotators will not be linguists and will not have experience in corpus creation/annotation. This fact should be taken into account when choosing terminology (no specific linguistic terms) and when defining the task (e.g., be clear about annotating coordinated elements as separate elements and not one element).

2) Choice of the annotation tool to be used should take into account the complexity of the task but also the involvement of non-linguists in annotation process. From the point of view of functionality, the tool should at the very least be able to capture relations, potentially embedded. This requirement makes tools not allowing relation annotation, such as WebAnnotator (Tannier 2012), not appropriate. After testing and comparing several tools, we chose the Glozz platform (Widlöcher & Mathet 2012) as the one that best corresponds to the needs of the task of full linguistic annotation of spin. Glozz is a flexible and powerful tool that allows to annotate units (text fragments), their relations and schemes (which can be seen as higher-level relations that can include one or more units, relations or other schemes) which covers all possible instances of incompleteness or incoherence in reporting.

However, demonstration of text annotation with Glozz to a medical expert showed that it does not meet the requirements of non-linguist annotators: ease of installation of the tool, amount of time needed for training for a person without previous experience in corpus annotation, complexity of guidelines describing the task. Consequently, we decided to replace the task of linguistic annotation of texts by a set of simpler tasks (in the form of questions) such as the following:

- validation/correction of primary outcomes found at the pre-annotation stage;

- validation/correction of reported outcomes found at the pre-annotation stage;

- establishing if two given (extracted at previous stages) outcomes refer to the same concept;

- identification of similarity statements in the Results and Conclusions of abstracts;

- identification of within-group comparisons in the Results and Conclusions of abstracts;

- identification of other positive evaluation of the studied treatment in the Results and Conclusions of abstracts.

We plan to use a web-based survey tool (such as LimeSurvey[7]) to create questionnaires containing these questions, generated on the basis of pre-annotation. Using survey tools for corpus annotation is not typical. Our decision is motivated by several reasons: survey tools are usually available online and thus do not require any complex installation procedures (survey participants can access the survey simply by following a link received by email); survey tools are widely used in medical community and are familiar to the community. This fact reduces time needed for annotators to learn how to use the tool. Besides, breaking the task into simple questions, independent one from another, allows to include into each question a brief guideline on how to answer, thus in most cases annotators will not need to refer to an extensive external annotation guide. Answering simple questions is also likely to cause fewer discrepancies between annotators than full annotation of spin.

3) In case of full linguistic annotation of spin, we should clearly define which pieces of text to annotate. We anticipate some difficulties in cases when elements of trial data get embedded one into another. The guidelines should explain whether to annotate these elements as embedded or as two separate instances linked by a certain type of relation (e.g. outcome and its method of measurement).

4) Given the complexity of the task, we need to clarify the definition of what should be considered to be spin. For this, we need to strictly define the types of spin that we focus on, describe in detail which pieces of information are relevant to these types of spin. Taking into account lack of agreement between experts in detailed definition of spin, for our current annotation project we decided to avoid using the notion "spin" and focus on tasks that are relatively simpler and clearer, such as: annotating

_____
[7] https://www.limesurvey.org/

outcomes; marking if pairs of extracted outcomes refer to the same concept; annotating specific constructions of interest, such as similarity statements or within-group comparisons. This information would allow to estimate with a certain probability that an article does or does not contain spin, but the final decision is left to the human readers of the article.

5) The task of developing guidelines must be fulfilled in close collaboration with experts in medical domain and in the domain of spin in medical texts, in order to verify that all the definitions regarding medical concepts and spin are correct.

## 5. Conclusions and future work

In this paper we described our approach to creation of a corpus of biomedical articles annotated for spin (distorted reporting) and its supporting information. We briefly outlined the proposed algorithm of spin detection and summarized our work on automatic pre-annotation of the corpus. Consequently, we described the annotation scheme that we developed for spin annotation. We discussed the process of creating the annotation guidelines, provided some thoughts as for choice of annotation tool and outlined expected challenges.

Our future tasks include running a pilot survey to validate usability of survey format for our task and evaluate the adequacy and clarity of the questions for annotators. Consequently we will proceed to a full-scale survey project.

## Acknowledgements.

## 6. Bibliographical References

1. Boutron I., Altman D.G., Hopewell S., Vera-Badillo F., Tannock I., Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of Cancer: the SPIIN randomized controlled trial. J Clin Oncol, 32, 4120–4126 (2014)
2. Boutron I., Dutton S., Ravaud P., Altman D.G. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA, 303, 2058–2064 (2010)
3. Conroy N. J., Rubin V. L., & Chen Y., Automatic Deception Detection: Methods for Finding Fake News, Proceedings of ASIST 2015, St. Louis, MO, USA, November (2015)
4. Hall M., Frank E., Holmes G., Pfahringer B., Peter R., Witten I. H. The weka data mining software: An update. SIGKDD Explorations, 11(1) (2009)
5. Hancock J. T., Beaver D. I., Chung C. K., Frazee J., Pennebaker J. W., Graesser A., & Cai Z., Social language processing: A framework for analyzing the communication of terrorists and authoritarian

regimes, Behavioral Sciences of Terrorism and Political Aggression, Vol. 2, No. 2, pp 108–132, May (2010)

6. Haneef R., Lazarus C., Ravaud P., Yavchitz A., Boutron I. Interpretation of results of studies evaluating an intervention highlighted in Google Health News: a cross-sectional study of news. PLoS ONE, 10(10) (2015)

7. Koroleva A., Paroubek P. Automatic detection of inadequate claims in biomedical articles: first steps. Proceedings of Workshop on Curative Power of MEdical Data (2017)

8. Koroleva A., Paroubek P. On the contribution of specific entity detection and comparative construction to automatic spin detection in biomedical scientific publications. Proceedings of The Second Workshop on Processing Emotions, Decisions and Opinions at The 8th Language and Technology Conference (2017)

9. Lazarus C., Haneef R., Ravaud P., Boutron I. Classification and prevalence of spin in abstracts of non-randomized studies evaluating an intervention. BMC Med Res Methodol., 15:85 (2015)

10. Litvinova O., Seredin P., Litvinova T. & Lyell J., Deception detection in Russian texts, Proc. of the Student Research Workshop at EACL2017, Valencia, Spain, April (2017)

11. Mihalcea R., and Strapparava C., The Lie Detector: Explorations in the Automatic Recognition of Deceptive Language. Proceedings of ACL, pp. 309-312, Singapore, (2009).

12. Nakashole N., & Mitchell T.M., Language-Aware Truth Assessment of Fact Candidates, Proceedings of the 52nd Annual Meeting of the ACL, pp 1009–1019, Baltimore, Maryland, USA, June (2014)

13. Pak A., Automatic, adaptive, and applicative sentiment analysis, PhD. dissertation, U. Paris-Sud, https://tel.archives-ouvertes.fr/tel-00717329

14. Pustejovsky J., Castaño J., Ingria R., Saurí R., Gaizauskas R., Setzer A. and Katz G. TimeML: Robust Specification of Event and Temporal Expressions in Text. IWCS-5, Fifth International Workshop on Computational Semantics (2003)

15. Tannier X. WebAnnotator, an Annotation Tool for Web Pages. Proceedings of the 8th International Conference on Language Resources and Evaluation (2012)

16. Widlöcher A., Mathet Y. The Glozz platform: a corpus annotation and mining tool. ACM Symposium on Document Engineering (2012)

17. Wiebe J., & Rilof E., Creating Subjective and Objective Sentence Classiers from Unannotated Texts, Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing) (2005)

18. Yamamoto Y., Takagi T. A sentence classification system for multi biomedical literature summarization. Proceedings of the 21st International Conference on Data Engineering Workshops (2005)

19. Yavchitz A., Boutron I., Bafeta A., Marroun I., Charles P., Mantz J., et al. Misrepresentation of randomized controlled trials in press releases and news coverage: a cohort study. PLoS Med, 9:e1001308 (2012)

20. Yavchitz A., Ravaud P., Altman D.G., Moher D., Hrobjartsson A., Lasserson T., Boutron I. A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. Journal of Clinical Epidemiology, 75, 56-65 (2016)