

GRO Task: Populating the Gene Regulation Ontology with events and relations

**Jung-jae Kim,
Xu Han**

School of Computer Engineering
Nanyang Technological University
Nanyang Avenue, Singapore
jungjae.kim@ntu.edu.sg,
HANX0017@e.ntu.edu.sg

Vivian Lee

European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton, Cambridge,
UK
vivian_clee@yahoo.com

Dietrich Rebholz-Schuhmann

Institute of Computational Linguistics
University of Zurich
Binzmühlestrasse 14
Zurich, Switzerland
rebholz@cl.uzh.ch

Abstract

Semantic querying over the biomedical literature has gained popularity, where a semantic representation of biomedical documents is required. Previous BioNLP Shared Tasks exercised semantic event extraction with a small number of pre-defined event concepts. The GRO task of the BioNLP'13-ST imposes the challenge of dealing with over 100 GRO concepts. Its annotated corpus consists of 300 MEDLINE abstracts, and an analysis of inter-annotator agreement on the annotations by two experts shows Kappa values between 43% and 56%. The results from the only participant are promising with F-scores 22% (events) and 63% (relations), and also lead us to open issues such as the need to consider the ontology structure.

1 Background

As semantic resources in the biomedical domain, including ontologies and linked data, increase, there is a demand for semantic querying over the biomedical literature, instead of the keyword searching supported by conventional search engines (e.g. PubMed). The semantic search requires adapting Semantic Web technologies to the literature, to analyze the complex semantics described in biomedical documents and to represent them with ontology concepts and relations. The ontology-based formal semantics will then form a Semantic Web. The GRO task of the BioNLP Shared Tasks 2013 is to provide a platform to develop and evaluate systems for identi-

fying complex semantic representation of biomedical documents in the domain of gene regulation.

There are solutions for servicing the ontology concepts recognized in the biomedical literature, including TextPresso (Müller *et al.*, 2004) and GoPubMed (Doms and Schroeder, 2005). They utilize term recognition methods to locate the occurrences of ontology terms, together with terminological variations. Systems like EBIMed (Rebholz-Schuhmann *et al.*, 2007) and FACTA (Tsuruoka *et al.*, 2008) go further to collect and display co-occurrences of ontology terms. However, they do not extract events and relations of the semantic types defined in ontologies.

The annotation of those ontology event and relation instances described in text was initiated in the biomedical domain by the GENIA corpus (Kim *et al.*, 2003), and the tasks of the BioNLP Shared Tasks 2009 and 2011 aimed at automatically identifying such ontological annotations. However, the tasks dealt only with a small number of ontology concepts (less than 20 unique concepts in total), considering the thousands of concepts defined in standard biomedical ontologies (e.g. Gene Ontology, anatomy ontologies). The goal of the Gene Regulation Ontology (GRO) task is to confirm if text mining techniques can be scaled up to cover hundreds of (and eventually thousands of) concepts, and thereby to address the complex semantic representation of biomedical documents.

The GRO task is to automatically annotate biomedical documents with the Gene Regulation Ontology (Beisswanger *et al.*, 2008). GRO is a

conceptual model of gene regulation and includes 507 concepts, which are cross-linked to such standard ontologies as Gene Ontology and Sequence Ontology and are integrated into a deep hierarchical structure via is-a and part-of relations. Note that many of the GRO concepts are more specific than those used in the previous BioNLP Shared Tasks. The GRO is one of the first ontological resources that bring together different types of ontology concepts and relations in a coherent structure. It has two top-level categories of concepts, Continuant and Occurrent, where the Occurrent branch has concepts for processes that are related to the regulation of gene expression (e.g. Transcription, RegulatoryProcess), and the Continuant branch has concepts mainly for physical entities that are involved in those processes (e.g. Gene, Protein, Cell). It also defines semantic relations (e.g. hasAgent, locatedIn) that link the instances of the concepts. The GRO task in the BioNLP Shared Task (ST) 2013 assumes that the instances of Continuant concepts are provided and focuses on extracting the instances of the events and relations defined in the GRO.

This paper is organized as follows: We describe the manual construction of the training and test datasets for the task in Section 2 and explain the evaluation criteria and the results in Section 3.

2 Corpus annotation

2.1 Annotation elements

The BioNLP'13-ST GRO task follows the representation and task setting of the ST'09 and ST'11 main tasks. The representation involves three primary categories of annotation elements: entities (i.e. the instances of Continuant concepts), events (i.e. those of Occurrent concepts) and relations. Mentions of entities in text can be either contiguous or discontinuous spans that are assigned the most specific and appropriate Continuant concepts (e.g. TranscriptionFactor, CellularComponent). The event annotation is associated with the mention of a contiguous span

in text (called event trigger) that explicitly suggests the annotated event type (e.g. 'controls' - RegulatoryProcess). If a participant of an event, either an entity or another event, can be explicitly identified with a specific mention in text, the participant is annotated with its role in the event. In this task, we consider only two types of roles (i.e. hasAgent, hasPatient), where an agent of an event is the entity that causes or initiates the event (e.g. a protein that causes a regulation event), and a patient of an event is the entity upon which the event is carried out (e.g. the gene that is expressed in a gene expression event) (Dowty, 1991). The semantic relation annotation is to annotate other semantic relations (e.g. locatedIn, fromSpecies) between entities and/or events, without event triggers. Figure 1 illustrates some of the annotations.

2.2 Document selection

The corpus texts are selected based on the relevance to the topic of gene regulation in humans. Specifically, we first obtained a list of human transcription factors (TFs) and then used PubMed to collect a set of candidate documents. A random subset of 300 documents was then selected for the GRO task from the collection. We annotated entities, events, and relations in them, and divided them into three subsets of 150 (training), 50 (development), and 100 (test) documents. In fact, 100 out of the 200 documents for training and development are from Kim et al. (2011a), though we revised and updated their annotations based on new annotation guidelines, some of which are explained below.

2.3 Annotation guidelines

The first step of annotating ontology concepts in the text is the recognition of a word or a phrase that refers to a concept of the GRO. Such a word or phrase, called mention, is one of the names of the concept, its synonyms, or expressions that are semantically equivalent to or subsumed by the concept. For each mention, we annotate it with the single, most specific and appropriate concept, but not with any general concept. For example, if

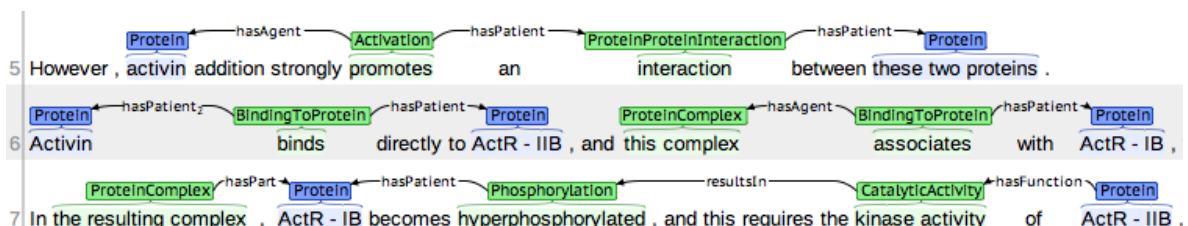


Figure 1. Example annotations of the GRO corpus

a protein is clearly mentioned as a transcription factor in the text, we annotate it with the GRO concept `TranscriptionFactor`, not with `Protein`.

There are many issues in the annotation, and we here introduce our guidelines on two of them about complex noun phrases and overlapping concepts.

1) If a noun phrase refers to an event that corresponds to an Occurrent concept and includes mentions of other concepts, we consider separately annotating the multiple mentions in the phrase with concepts and relations. For example in the phrase “nephric duct formation”, we annotate it as follows:

- “formation”:`CellularProcess` hasPatient “nephric duct”:`Cell`

This means that the phrase indicates an individual of `CellularProcess`, which is an event of forming an entity of `Cell`, which is nephric duct. Another example noun phrase that involves multiple mentions is “Sim-2 mRNA expression”, which is annotated as follows:

- “expression”:`GeneExpression` hasPatient (“mRNA”:`MessengerRNA` encodes “Sim-2”:`Gene`)

However, we do not allow such multi-mention annotation on e.g.

- “mRNA expression”, because this phrase is too generic and frequent so that a multi-mention annotation for it, “expression”:`GeneExpression` hasPatient “mRNA”:`MessengerRNA`, does not encode any ‘useful’ information
- “nuclear factor”, because this factor is not always located in nucleus.

Therefore, we decided that, in general, we avoid annotation of generic information, but consider a thread of information specific only if it involves specific entities like individual gene/protein and cell (e.g. Sim-2, nephric duct). Also, we did not divide a noun phrase to multiple mentions if the relation between the mentions is not always true (cf. “nuclear factor” – “factor”:`Protein` locatedIn “nuclear”:`Nucleus`).

2) As some GRO concepts are overlapping, we made the following guidelines:

(a) When there is ambiguity between `Increase` (`Decrease`), `Activation` (`Inhibition`), and `PositiveRegulation` (`NegativeRegulation`), we annotate

- binary relations with `PositiveRegulation`, ignoring `Activation`

(e.g., “augment”:`PositiveRegulation` hasAgent “Nmi”:`Protein` hasPatient (“recruit-

ment”:`Transport` hasPatient “coactivator protein”: `TranscriptionCoactivator`))

- unary relations with `Increase` (e.g., “enhance”:`Increase` hasPatient “transcription”:`Transcription`)

Note that we cannot exchange the two concepts of `PositiveRegulation` and `Increase` in the two examples due to the arity restriction.

(b) Binding concepts are ambiguous. We annotate as follows:

- For such a GRO concept as “Binding of A to B”, A should be the agent and B the patient.

(For example, when we annotate `BindingOfProteinToDNA` and `BindingOfTFToTFBindingSiteOfProtein`, `Protein` and `TF` will be agents, and `DNA` and `BindingSiteOfProtein` will be patients, respectively.)

- For such a GRO concept as “Binding to A” for binary relation between two entities of the same type, both entities should be patients.

(For example, in the events of binding between proteins with `BindingToProtein` and of binding between RNAs with `BindingToRNA`, the proteins and the RNAs, respectively, will all be patients.)

Other annotation guidelines can be found at the task homepage¹.

2.4 Annotation

Two annotators with biology background annotated the documents with GRO entities, events and relations. They used the Web-based annotation tool `brat` (Stenetorp *et al.*, 2012) for the annotation. Annotator A is the one who annotated the earlier version of the corpus (Kim *et al.*, 2011a). He first revised the earlier version of 100 abstracts (named Set 1) and drafted the new annotation guidelines. Annotator B studied the drafted annotations and guidelines and then further revised them, and the two annotators together updated and made agreements on final versions of the annotations and guidelines. They selected two more sets of 100 abstracts each (named Sets 2 and 3), where Set 2 was combined with Set 1 to become the training and development datasets, and Set 3 became the test dataset. They updated the guidelines after annotating Sets 2 and 3 independently and together combining their annotations.

¹ <http://nlp.sce.ntu.edu.sg/wiki/projects/bionlpst13grotask/>

We estimated the inter-annotator agreement (IAA) between the two annotators for Sets 2 and 3 with Kappa measures as shown in Table 1. The Kappa values between 43% and 56% are moderately acceptable, though not substantial, which is expected with the high degree of the ontology’s complexity and also with the high number of mentions (56 per abstract; see Table 2). Note that the agreement is met, only when the two annotators annotate the same concept on the same mention with the same boundaries and, if any, the same roles/arguments, not considering the generalization criteria used for evaluation (see Section 3 for details). If we relax the boundary restriction (i.e. approximate span matching of (Kim *et al.*, 2009)), the Kappa values for events slightly increase to 47% (Set 2) and 45% (Set 3). Also note that the agreement on relations is higher than those on entities and events.

We analyzed the different annotations by the two annotators as follows: As for the entity annotations, 84% of the differences are boundary mismatches, while the rest are due to mismatch of entity types and to missing by either of the annotators. As for the event annotations, 56% of the differences are also boundary mismatches, and 31% are missed by either of the annotators. The majority (71%) of the differences in relation annotations are due to missing by either annotator, while the rest are mostly due to the differences in the entity annotations.

One negative finding is that the agreement did not always increase from Set 2 to Set 3, which means the two annotators did not improve the alignment of their understanding about the annotation even after making agreements on Set 2 annotations. It may be too early to conclude, and the Kappa value might increase as the annotators examine more examples, since the annotation corpus size in total (Sets 1,2,3 together) is still small compared to the total number of GRO concepts. After examining the IAA, we integrated the independently annotated sets and released the final versions of the three datasets at the task homepage.

Table 1. Inter-annotator agreement results

	Set 2	Set 3
Entities	44.6%	43.8%
Events	45.8%	43.2%
Relations	54.7%	55.9%
All	46.2%	45.3%

2.5 Statistics

Table 2 shows the number of MEDLINE abstracts in each of the three datasets: training, development, and test datasets. It also shows the number of instances for each of the following annotation types: entities (i.e. instances of Continuant concepts), event mentions (i.e. event triggers), event instances (i.e. instances of Occurrent concepts), and relation instances. Note that relation instances are not associated with mentions like event instances. It also shows the number of unique entity/event types (i.e. unique GRO concepts) used in the annotation of each dataset. The total number of unique entity types in the three datasets is 174, and that of unique event types is 126.

Table 2. Number of annotation elements

	Train	Dev.	Test
No. of documents	150	50	100
No. of entity mentions	5902	1910	4007
No. of event mentions	2005	668	2164
No. of event instances	2175	747	2319
No. of event instances with agents	693	251	625
No. of event instances with patients	1214	451	1467
No. of relation instances	1964	581	1287
No. of unique entity types	128	94	147
No. of unique event types	98	72	100

Note that the frequency of event instances in the test dataset (23.2 per document) is much higher than those in the training and development datasets (14.5 and 14.9 per document, respectively). We compared the three datasets and observed that several event types (e.g. GeneticModification), which are popular in the test dataset (e.g. GeneticModification is the 12th frequent type (2.3%)), seldom appear in the other two datasets. It may indicate that the annotators were getting aware of (or familiar with) more GRO concepts as they annotate more documents, where the test dataset is the last annotated. This sudden increase of frequency did not happen for the entity annotations, possibly because the two annotators were provided with candidate entity annotations, though of low quality, from a preliminary dictionary-based entity recognition method and modified them.

Table 3 shows the number of mentions for the most frequent top-level Continuant concepts such as InformationBiopolymer, whose sub-concepts include Gene and Protein, Cell, and

ExperimentalMethod. Please note that these frequent concepts are closely related to the topic of gene regulation, and that this distribution may reflect to some degree the distribution of terms in the sub-domain of gene regulation, but not that in the whole MEDLINE. If you like to see the descendant concepts of those top-level concepts, please refer to the latest version of the GRO².

Table 3. Number of mentions for frequent top-level Continuant concepts

Level 2	Level 3	Level 4	Count
Continuant/PhysicalContinuant			3647
	MolecularEntity		2805
		InformationBiopolymer	2508
		ComplexMolecularEntity	140
		Chemical	127
		Ligand	27
	LivingEntity		584
		Cell	306
		Organism	268
	Tissue		170
	CellComponent		77
Continuant/NonPhysicalContinuant			359
	ExperimentalMethod		123
	Function		111
	MolecularStructure		66
	Locus		25
	Phenotype		11

Table 4 shows the number of event instances for the most frequent top-level Occurrent concepts. Table 5 shows the number of instances for each relation.

Table 4. Number of event instances for frequent top-level Occurrent concepts

Level 3	Level 4	Count
Occurrent/Process/RegulatoryProcess		782
	PositiveRegulation	217
	NegativeRegulation	186
Occurrent/Process/MolecularProcess		422
	IntraCellularProcess	189
Occurrent/Process/PhysiologicalProcess		418
	OrganismalProcess	143
Occurrent/Process/PhysicalInteraction		312
	Binding	296
Occurrent/Process/Mutation		82
Occurrent/Process/Localization		77

² <http://www.ebi.ac.uk/Rebholz-srv/GRO/GRO.html>

	Transport	16
Occurrent/Process/Decrease		73
Occurrent/Process/Affecting		64
	Maintenance	20
Occurrent/Process/ExperimentalIntervention		54
	GeneticModification	54
Occurrent/Process/Increase		49
Occurrent/Process/ResponseProcess		38
	ResponseToChemicalStimulus	13

Table 5. Number of relation instances

Relation	Count	Relation	Count
locatedIn	405	hasPart	403
fromSpecies	274	hasFunction	82
resultsIn	56	encodes	49
precedes	17	hasQuality	1

3 Evaluation

There was one submission for the GRO task of the BioNLP'13-ST, designated as "TEES-2.1" (Björne and Salakoski, 2013). For comparison purposes, the GRO task organizers produced results with a preliminary system by adapting our existing system, designated as OSEE (Kim and Rebholz-Schuhmann, 2011b), for event extraction and developing a simple machine learning model for relation identification. We describe these two systems briefly and compare their results with several criteria.

3.1 System descriptions

TEES-2.1 is based on multi-step SVM classification, which automatically learns event annotation rules to train SVM classifiers and applies the classifiers for 1) locating triggers, 2) identifying event arguments, and 3) selecting candidate events.

OSEE is a pattern matching system that learns language patterns for event extraction from the training dataset and applies them to the test dataset. It performs the three steps of TEES-2.1 in a single step of pattern matching, thus requiring a huge amount of patterns (eventually, a pattern for each combination of the features from the three steps) and failing to consider that many features of a step are independent from other steps and also from event types and can thus be generalized.

We added a simple Naïve Bayes model to the system for identifying (binary) semantic relations between entities, which utilizes such features as

entity strings, the distance between them, and the shortest path between the two entities in the dependency structure of the source sentence, which is identified by Enju parser (Sagae *et al.*, 2007).

3.2 Evaluation criteria

The GRO task follows some of the evaluation criteria of the Genia Event Extraction (GE) task of BioNLP-ST 2009 (Kim *et al.*, 2009), including strict and approximate matching, and also introduce new criteria that consider 1) the hierarchical structure of the GRO and 2) parent and/or grandparent of answer concept. We here explain these new criteria in detail.

1) In this scheme of evaluation, the event results of a participant are classified into the GRO concepts at the third level (see Table 4 for examples), which are ancestors of their labeled classes, and the evaluation results are accumulated for each of those concepts at the third level. This scheme may give us insights on which categories the participant system shows strength or weakness.

2) This scheme is to deal with such a case that the answer class is "GeneExpression", but a participant gives "IntraCellularProcess" or "MolecularProcess", which are the parent and grandparent of the answer class, thus not entirely wrong nor too generic. For example, the scheme "Allowing parents" allows "IntraCellularProcess" to be a correct match to the answer class "GeneExpression", as well as the answer class itself. "Allowing grandparents" accepts the grandparents of answer classes as well as the parents.

3.3 Evaluation results

Table 6 shows the evaluation results of the two systems. Note that all the evaluation results in terms of precision, recall, and F-score in all the tables are percentages. The performance of the TEES-2.1 systems, which is clearly better than the OSEE system, is lower than its performance for other tasks of the BioNLP'13-ST, which is understandable, considering 1) the higher number of GRO concepts than those for the other tasks and 2) the low Kappa value of the inter-annotator agreement.

It also shows that the evaluation scheme that allows the parents/grandparents of answer concepts for acceptance does not greatly help increasing the performance, which may mean that the systems are designed to aim individual concepts, not considering the ontology structure. This issue of considering the structure of the on-

tology in event extraction can be an interesting future work.

Table 6. Evaluation results (percentage)

Evaluation scheme	TEES-2.1			OSEE		
	R	P	F	R	P	F
Strict matching	15	37	22	10	18	13
Approximate boundary matching	16	39	23	10	20	14
Approximate recursive matching	16	39	23	12	20	15
Allowing parents	16	38	23	10	19	13
Allowing grandparents	16	38	23	10	19	13

Table 7 shows the performance of the systems for different event categories in the third level of the GRO. It shows that the systems are good at extracting events of the categories of MolecularProcess (e.g. GeneExpression) and Localization (e.g. Transport), but are, expectedly, poor at extracting events of the categories with small number of training data (e.g. Decrease, ResponseProcess).

Table 7. Evaluation results grouped into 3rd-level GRO concepts (%)

3 rd -level concept	TEES-2.1			OSEE		
	R	P	F	R	P	F
RegulatoryProcess	12	24	16	10	11	11
MolecularProcess	30	60	40	23	51	31
Physiological Process	9	78	17	6	25	9
PhysicalInteraction	18	33	24	3	6	4
Mutation	16	39	23	1	8	2
Localization	21	62	31	16	55	24
Decrease	3	12	4	0	0	0
Affecting	2	50	3	0	0	0
Increase	8	8	8	0	0	0
ResponseProcess	3	8	4	5	50	10

Table 8 shows the performance of the systems for the most frequent concepts and also for some selected infrequent concepts. From the results, we observe that the system performance for an event class does not reflect the number of train-

ing data of the class, and that the performance of the syntactic pattern matching system OSEE is high for the event classes, for which the machine learning system TEES-2.1 also performs well. These observations may indicate that the current approaches to event extraction deal with event types independently, not considering the hierarchical (or semantic) relations between the event types nor relations between entity types.

Table 8. Evaluation results for frequent and infrequent individual concepts (%)

Event class (Count)	TEES-2.1			OSEE		
	R	P	F	R	P	F
RegulatoryProcess (224)	18	23	20	13	13	13
PositiveRegulation (217)	11	22	15	11	9	9
NegativeRegulation (186)	12	23	16	14	10	12
GeneExpression (160)	59	72	65	46	67	55
Disease (143)	0	0	0	1	100	3
Decrease (73)	3	12	4	0	0	0
Localization (61)	16	71	27	20	60	30
DevelopmentalProcess (61)	23	82	36	23	78	35
BindingOfProteinToDNA (55)	13	15	14	0	0	0
GeneticModification (54)	0	0	0	0	0	0

Table 9 shows the performance of the systems for the GRO relations. These results of TEES in the relation identification of the GRO task (F-scores between 50% and 87%) are much higher than the best results of relation identification (40% F-score) in the Bacteria Biotopes (BB) task (Nédellec *et al.*, 2013), which is to extract relations of localization and part-of. Though the two relation identification tasks of GRO and BB cannot be directly compared due to many differences (e.g. entity types, relation types, corpus sources), it may indicate that the GRO task corpus has been annotated consistently enough to train a model with such high performance and that the low performance of event extraction compared to relation identification may be due to the big number of event types and would be resolved as the corpus size increases.

Table 9. Evaluation results for relations (%)

Relation	TEES-2.1			OSEE		
	R	P	F	R	P	F
locatedIn	45	83	58	66	38	48
hasPart	45	81	58	76	22	34
fromSpecies	80	96	87	89	41	56
hasFunction	38	73	50	62	20	30
encodes	49	89	63	45	2	5
Total	49	86	63	72	23	35

4 Conclusion

The main challenge in this task is the increased size of the underlying ontology (i.e. GRO) and the more complex semantic representation in GRO compared to those in other ontologies used for ontology-based event extraction. The complex structure of the GRO enables us to evaluate participant systems at different abstraction/generalization levels. The evaluation results of the participant are quite promising, leading us to open issues in this direction, including the incorporation of ontology structure in event extraction. We plan to extend the corpus semi-automatically by incrementally updating the event extraction system with more training data.

References

- E. Beisswanger, V. Lee, J.-J. Kim, D. Rebholz-Schuhmann, A. Splendiani, O. Dameron, S. Schulz, and U. Hahn, "Gene Regulation Ontology (GRO): design principles and use cases," *Stud Health Technol Inform*, vol. 136, pp. 9–14, 2008.
- Jari Björne, Tapio Salakoski. TEES 2.1: Automated annotation scheme learning in the BioNLP 2013 Shared Task. In proceedings of the workshop of BioNLP 2013 Shared Task, 2013. (submitted)
- A. Doms, M. Schroeder. GoPubMed: exploring PubMed with the gene ontology. *Nucleic Acids Res* 2005; 33:W783–6.
- D. Dowty. Thematic Proto-Roles and Argument Selection. *Language* 67(3):547-619, 1991.
- J.D. Kim, T. Ohta, Y. Tateisi et al. GENIA corpus - a semantically annotated corpus for bio-text mining. *Bioinformatics* 19:i180-i182, 2003.
- J.D. Kim, T. Ohta, S. Pyysalo et al. Overview of BioNLP'09 shared task on event extraction. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing: Shared Task*, Association for Computational Linguistics, pp. 1-9, 2009.

- Jung-Jae Kim, Xu Han and WatsonWei Khong Chua. Annotation of biomedical text with Gene Regulation Ontology: Towards Semantic Web for biomedical literature. In Proceedings of LBM 2011, pp.63–70, 2011a.
- Jung-jae Kim, Dietrich Rebholz-Schuhmann. Improving the extraction of complex regulatory events from scientific text by using ontology-based inference. *Journal of Biomedical Semantics* 2(Suppl 5):S3, 2011b.
- H.M. Müller, E.E. Kenny, P.W. Sternberg. Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biol* 2:e309, 2004.
- Claire Nédellec, Robert Bossy, Jin-Dong Kim, Jung-jae Kim, Tomoko Ohta, Sampo Pyysalo, Pierre Zweigenbaum. Overview of BioNLP Shared Task 2013. *Proc Workshop BioNLP Shared Task 2013, ACL 2013, 2013*. (to appear)
- D. Rebholz-Schuhmann, H. Kirsch, M. Arregui, et al. EBIMed: text crunching to gather facts for proteins from Medline. *Bioinformatics* 23:e237–44, 2007.
- Kenji Sagae, Yusuke Miyao, and Jun'ichi Tsujii. 2007. HPSG Parsing with Shallow Dependency Constraints. In *Proceedings of ACL 2007*, 2007.
- P. Stenetorp, S. Pyysalo, G. Topic, T. Ohta, S. Ananiadou, and J. ichi Tsujii, “brat: a Web-based Tool for NLP-Assisted Text Annotation,” *EACL*. The Association for Computer Linguistics, pp. 102–107, 2012.
- Yoshimasa Tsuruoka, Jun'ichi Tsujii, and Sophia Ananiadou. FACTA: a text search engine for finding associated biomedical concepts. *Bioinformatics* 24(21):2559-2560, 2008.