Extracting ODRL Digital Right Representations from License Texts using AMR

Malo Revel[1], Aurélien Lamercerie[2], Annie Foret[1] and Zoltan Miklos[1]

[1]IRISA & Univ. Rennes, Campus de Beaulieu, Rennes, France
[2]Tétras Libre, Grenoble, France

Abstract

Licenses of digital resources describe rights and duties for users. If the licenses are expressed in natural language, as it is frequently the case, it is hard to reason and verify the license compatibility to specific uses. We propose an automatic end-to-end workflow for extracting Open Digital Rights Language (ODRL) representations from textual license documents. This process uses AMR semantic representations as an intermediate; it adapts a tool that performs a semantic transduction analysis, using formal rules. This work focuses on deontic modalities expressing the permissions and obligations of the user. We provide a proof of concept and discuss experiments.

Keywords

Content Extraction, Automated Semantic Analysis, Semantic Graph, AMR, License Rights, ODRL

1. Introduction

One may wish to attach a license to a work (e.g. a program, a picture or a dataset) before publishing it on the Web. A license expresses in natural language (NL) the rights and duties that the users of the work must comply with. However, although some licenses such as the licenses provided by Creative Commons[1] are also available in a machine readable format, most licenses are only provided as human readable texts, which are written in NL. Machine readable licenses enable a much easier automatic extraction of the rights expressed by the licenses, which is useful for tasks such as license synthesis [1], or compatibility and compliance inference [2], [1]. Performing such tasks using only NL license texts would be much more arduous and ambiguous.

Overall objectives. The work presented in this paper aims at automatically computing machine-readable versions of NL license texts. More specifically, our target representation format is the Open Digital Rights Language (ODRL [3]), an RDF vocabulary meant to express rights and duties. In order to realize this translation, Abstract Meaning Representation (AMR), a graph-based semantic representation proposed by [4], is used as an intermediary representation.

The information we aim at extracting from the license texts is rather specific. Indeed, what interests us the most are sentences that describe the permissions and obligations of the user. Such sentences typically contain a modal verb such as "may" or "must", and an action that is an object of this modal verb, such as "distribute" or "give credit". Additionally, the action of the sentence refers to a subject (e.g. the user) and a target (e.g. the work offered under the terms of the license). These parts of information all correspond to specific ODRL classes. For instance, in the sentence "You may reproduce the Work", the modal verb "may" expresses a permission, and controls the action "reproduce". The whole sentence means that the action of reproduction appears among the user's permissions.

Thus, our approach aims to extract ODRL representations for specific content, with a focus on deontic modalities expressing authorized actions and unauthorized ones. Minimizing errors in a fully automated workflow is also an important issue.

Contribution. Our main contribution is the proposal of an automatic end-to-end workflow for extracting ODRL representations from textual license documents. This process integrates a pre-existing analysis tool [5], implementing the principles of semantic transduction analysis [6]. This tool automatically produces ontologies from semantic graphs, using formal rules. We adapted it by developing specific rules to deal with deontic modalities and to capture specific semantic content corresponding to the target representation. In addition, an ODRL graph generation step has been added to the workflow.

At this stage, we propose a proof of concept evaluated on a dataset composed of a hundred typical sentences with a moderate complexity. Our experimentation, which is intended to be preliminary, was notably oriented by following ODRL implementation best practices presented

Proceedings of the Sixth Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL 2023), June 23, 2023, Braga, Portugal.

Malo.Revel@irisa.fr (M. Revel);
aurelien.lamercerie@tetras-libre.fr (A. Lamercerie);
Annie.Foret@irisa.fr (A. Foret); Zoltan.Miklos@irisa.fr (Z. Miklos)

[1] https://creativecommons.org/

in the report [7].

Related works are explored in Section 2, while Section 3 introduces the notions and knowledge on which we rely in this paper. The workflow and the implemented methodology are detailed in Section 4. The experimentation carried out is finally presented in Section 5.

2. Related Works

Analyzing licenses written in NL to produce RDF graphs is quite a specific task that has not been extensively addressed yet. A state of the art for this task is provided by Cabrio et al. [8], whose goal is similar to ours, but whose method differs from ours.

In order to extract RDF specifications from NL texts, Cabrio et al. [8] first pre-process the input license by tokenizing, lemmatizing and processing the Parts of Speech of the text, then use classifiers based on Support Vector Machines (SVM) to generate a CC REL [9] or an ODRL [3] graph. CC REL and ODRL are two Resource Description Framework (RDF) vocabularies dedicated to digital rights specification, and particularly license specification. Cabrio et al. [8] evaluated their system on a dataset of 37 licenses that they annotated in RDF. The classifiers achieved an overall precision of approximately 0.77 and a recall of 0.43, varying depending on the classified action.

These results are not totally satisfying and show that the framework of Cabrio et al. [8] must be considered as a first step towards automatic NL license analysis rather than a definitive solution. This article [8] shows that tackling this challenge with the use of SVM is not trivial, and we hope our approach will obtain better results.

Havur et al. [1] present DALICC, which is both a license library that indexes several license texts along with their ODRL representations, and a system to reason on these RDF graphs and for instance compose customized licenses. Moreau et al. [2] propose CaLi, a model that compares and partially orders licenses based on compatibility and compliance relations between the licenses. To do so CaLi arranges formal representations of licenses based on ODRL in lattices over which CaLi reasons. Vu et al. [10] introduce JCivilCode, an AMR dataset that is specific to legal texts, and they evaluate several AMR parsers with this dataset.

3. Background Knowledge

3.1. Semantic Representations, AMR

Semantic parsing is an active area of research, involving several annotation languages or formats to represent semantic information [11]. In this study we use Abstract Meaning Representation (AMR) introduced by [12], in the form proposed by [4].

AMR is a readable expressive form of graph-based semantic representation at the sentence level, for which datasets and (semantic) parsing tools have been developped [13]. This scheme abstracts away from syntax, as a single AMR may correspond to several sentences (paraphrases). AMR relies heavily on the PropBank [2] semantic roles inventory that is verb-oriented.

Basic AMR and example. We consider an AMR in figure 1 for sentence (1) "You may distribute the work". Every AMR has a unique root. It has variables (p, etc.), events and concepts (such as permit-01, determine-01) and roles (such as ARG0, ARG1) that label the edges. In a node, a slash indicates that the variable on the left denotes an instance of the concept on the right. The edge relations follow verb frame descriptions in PropBank, in the case of "determine-01" : ARG0 denotes "distributor", ARG1 denotes "thing distributed".

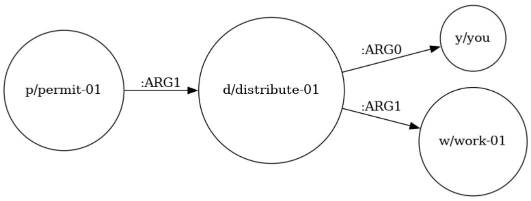

Figure 1: AMR Graph of sentence (1) "You may distribute the work"

Negation is expressed in AMR with a polarity relation. The relation is between the concept marked as negated and the constant "-". For this example (2) "you are not allowed to distribute the work", we get the graph (in PENMAN notation [3]):

```
(v2 / allow-01
  :polarity -
  :ARG1 (v3 / distribute-01
      :ARG1 (v4 / work-01))
  :ARG1 (v1 / you))
```

AMR as an intermediate format. Our hypothesis here is that for texts in the domain of digital licenses, AMR forms are good intermediate representations, from which we can derive the deontic information to be expressed in languages such as ODRL (The Open Digital Rights Language). Obvioulsy at a practical level, a given

[2] see the PropBank sites: https://verbs.colorado.edu/~mpalmer/projects/ace.html and https://propbank.github.io/

[3] described for exemple here https://penman.readthedocs.io/en/latest/notation.html

AMR parser may output errors (a wrong AMR). But importantly, at the level of format and guidelines, the way AMR highlights verbs (with their arguments) and the way AMR expresses and attach modals and negation seems appropriate to extract ODRL-like information. While some general limitations have been noted: AMR does not explicitly deal with universal quantification, and has no tense (future, etc.), we think this should not create problems for the task and texts considered in this study.

3.2. Targeted ontology

ODRL The target representation of our analysis is the Open Digital Rights Language (ODRL). ODRL is an RDF vocabulary dedicated to the representation of statements describing rights and duties, and license texts are made of statements of this kind. An ODRL `Policy` is composed of one or many `Rules`, which are either a `Permission` (what one *may* do), a `Duty` (what one *must* do) or a `Prohibition` (what one *is prohibited to* do). These rules are linked to one or many `Action` which describe the terms of use of an `Asset` - e.g. software or pictures.

For instance, figure 2 shows an ODRL graph (serialized in Turtle syntax) that expresses the information contained in the sentence (1) "You may distribute the Work" (where "You" and "the Work" are terms that are defined within the license text) :

```
@prefix cc: <http://creativecommons.org/ns#>.
@prefix odrl: <http://www.w3.org/ns/odrl/2/>.

:license a odrl:Policy ;
    odrl:permission [
        a odrl:Permission ;
        odrl:target "Work" ;
        odrl:assignee "You" ;
        odrl:action cc:Distribute
    ] .
```

Figure 2: ODRL representation of the sentence (1) "You may distribute the Work"

ODRL actions come from the ODRL main namespace or from other ontologies such as The Creative Commons Rights Expression Language (CCREL), as in the previous example. CCREL is a RDF vocabulary created by Creative Commons and is meant to represent license text information. CCREL is more limited than ODRL.

4. Methodology

Our methodology defines a global processing workflow, starting from the NL license sentences to extract the necessary information for the construction of their ODRL representations. This workflow is composed of several steps (figure 3): (1) document splitting and selection of key sentences, (2) conversion of NL statements into semantic representations (AMR graph), (3) RDF serialization of the resulting representations, (4) transduction parsing to extract the semantic content, and (5) generation of the ODRL representations.

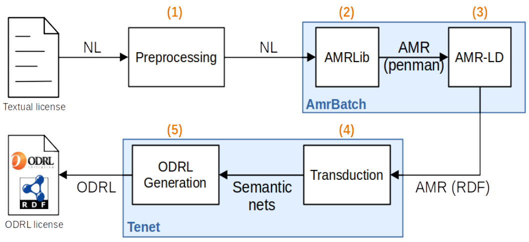

Figure 3: Workflow of our method

4.1. Text Data Preparation / Segment Selection

The first step of our workflow is the **document preprocessing**. It aims at highlighting different sets of sentences from the whole license document, and to select in particular the sentences that correspond to definitions of terminology or expressions of legal rules.

Indeed the only sentences that we want to consider are the ones that describe rights and duties. The first approach would be to consider all the sentences and only keep at the end the ones that describe rights and duties, but this method may result in performance issues, since the semantic parsing takes quite a long time to process. Another approach would be to do a first preprocessing step before the semantic parsing, which filters out some "useless" sentences. This way less sentences are to be parsed, and the remaining "useless" sentences can be ignored in a later step of the analysis.

Moreover, most license texts contain a section (usually the first section of the document) that defines some terms used in the text. Here is an example of a definition of "You" (from [14]) : "**You** means the individual or entity exercising the Licensed Rights under this Public License. Your has a corresponding meaning.". Being able to automatically recognize this section and analyze its definitions would make the extraction of the deontic sentences easier and less ambiguous, because the deontic sentences make use of the defined words.

This preprocessing step is currently purely hypothetic: so far the only computation done on a text before the semantic parsing step is the sentence splitting. As a consequence, in this article we only consider isolated sentences that express rights or duties.

4.2. Semantic Graph Construction

The second phase is the **NL sentence parsing** to construct the corresponding AMR graphs. This step is done by AMRBatch [4], which makes use of two modules:

- AMRLib [5] is a module that among others parses NL sentences to create AMR graphs. This parser may rely on different pretrained models, and we chose the `parse_xfm_bart_large` model, which reached a SMATCH score of 83.7 in 2022.
- AMR-LD [6] is a utilitary module that is used by AMRBatch to convert AMR graphs from the PENMAN format provided by AMRLib into an AMR graph. This RDF version of the AMR is a direct translation, and is still far from the ODRL graph that we aim to reach.

Once the AMR graph is computed, the next step of our method is to extract its patterns that correspond to the ODRL features that we want to produce.

4.3. Strategies on Semantic Graph Patterns

Semantic Transduction Analysis aims at extracting the semantic content to generate the desired representations. This approach, similar to graph transformation techniques, was initiated in the thesis of A. Lamercerie [6]. We take up here the general idea, which we have adapted to an application on licenses in order to extract elements of digital rights.

The notion of **Semantic Net** is introduced to support its implementation. A semantic net is an abstract object that covers graph nodes. It is a way to "capture" a meaning interpretation of nodes in relation. Each net can be typed and associated with different useful semantic data. For example figure 4 shows the nets A and B, which allow to capture respectively a modality and a property.

Transduction Mechanics defines the methodology used to direct the analysis. It aims to bring out semantic nets until one or more nets containing all the data needed to generate the expected ODRL statements are obtained. For its implementation, we define a set of net types, possibly structured as in figure 5. Using this typing, it becomes possible to check if a net has a given type with a simple formula as classNet(x). Similarly, two nets can be related, and we can check this using a formula as arg1(x, y). We can thus check if one or more nets

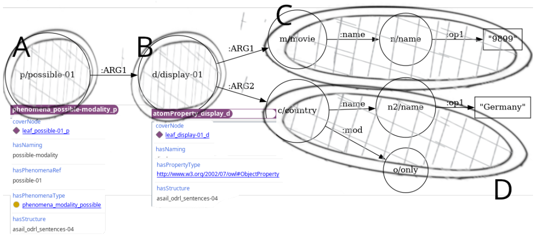

Figure 4: AMR Graph with semantic nets

satisfy certain typing or relation criteria with a formula. By linking such a formula to a constructive method, a mechanism makes possible to generate new nets. These rules are called transduction rules, and can be structured using a schema, introducing recursion if necessary.

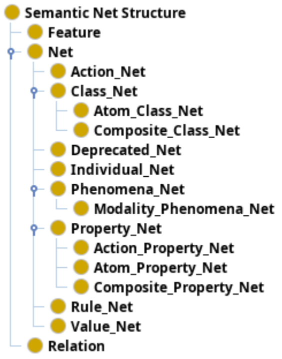

Figure 5: Semantic Net Typing

The **Analysis Procedure** is initialized by analyzing each node of a graph, and by associating an atomic net (a net that covers a single node) to it. The rest of the procedure consists in the analysis of the generated nets, and in the application of different rules permitting to build new nets by composition of different nets. The analysis of the nets A and B of the example of the previous figure (figure 4) reveals the existence of an action (corresponding to the property associated to B). A new net is created, E on the figure 6, covering the node associated with B, and extended by analyzing the relations starting from B (arg0 which points to C, and arg1 which points to D). Finally, A can be combined with this new net to produce another net, covering all the nodes and highlighting the existence of a digital rights rule.

The implementation uses composition rules as described in [6] or [5]. For this work on license rights, rules have been revised or specifically developed. These

[4] https://gitlab.tetras-libre.fr/tetras-mars/amrbatch

[5] https://pypi.org/project/amrlib/

[6] https://github.com/BMKEG/amr-ld

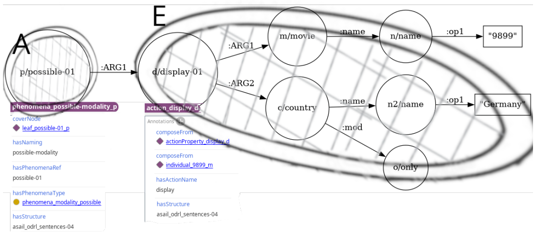

Figure 6: AMR Graph with new semantic nets

have in particular targeted the extraction of modalities and actions, by trying to take into account the variability of the structures to be highlighted, and the relationships between these structures.

Finally, the last step exploits the result of the semantic transduction analysis to **generate the expected ODRL representations**.

4.4. Hypotheses on sentence structures

The sentences that we analyze are expected to follow a certain structure, which corresponds to the semantic graph patterns that are recognized by TENET [7]. This structure is also determined by the ODRL ontology classes and properties. Here are the features that we expect to encounter in sentences that express digital rights.

Modalities Three different deontic operators may appear: permission, obligation and prohibition. The simplest way to express these operators in NL is through the usage of "may" (which is usually translated to the AMR concept `permit-01`) and "must" (translated to `obligation-01`), although deontic modality can be expressed through much more complex structures in license texts (e.g. "the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercize the Licensed Rights in the Licensed Material to" [8] expresses a permission). In addition, negation can be applied to the operator, which changes its meaning. For instance "You are not prohibited to <action>" expresses a permission, while "You are prohibited not to <action>" expresses an obligation.

Entities Some entities can appear in the sentence, and can be actors (`odrl:assignee`) or targets (`odrl:target`) of the action. The main entities are often defined in the definitions section of the text, and they often start with a capital letter.

Actions The actions are the most difficult features to recognize in a sentence, because there are a lot of different actions, and actions may be composed of a lot of AMR concept nodes that are not necessarily close together in the graph. Moreover, some actions have very similar meanings and are hard to differentiate in the AMR and even in the NL sentence. For instance [8], `cc:shareAlike` is textually described as "Redistributions must reproduce the above copyright notice" and `Duty:attachPolicy` as "Redistributions must retain the copyright notice".

Coordinating conjunctions Additionally, a single sentence can express a lot of information (for instance several actions) through coordinating conjunctions.

Aditionnal contraints Finally, elements of the sentence can be nuanced by additional constraints. For example, the assignee may be permitted to do an action only within a certain time interval or a certain country.

Example Here we study a simple sentence that follows the hypotheses above and we show its AMR and ODRL graph. The sentence (3) is the following: "You may not reproduce the Work". In this sentence, the modality is a prohibition, the assignee is "You", the target is "the Work" and the action is "reproduce".

Here is an AMR translation for this sentence (in PENMAN notation):

```
# ::snt You may not reproduce the Work.
(p / permit-01
    :polarity -
    :ARG1 (r / reproduce-01
        :ARG0 (y / you)
        :ARG1 (w / work-12)))
```

And here is an ODRL translation for this sentence (in TURTLE syntax):

```
@prefix cc: <http://creativecommons.org/ns#>.
@prefix odrl: <http://www.w3.org/ns/odrl/2/>.

"License" a odrl:Policy ;
        odrl:prohibition [
                odrl:target "Work" ;
                odrl:assignee "You" ;
                odrl:action cc:Reproduction
        ] .
```

4.5. Sources of error

One important source of error for the whole system is the lack of precision of AMRLib parsing tool for some sentences that we consider. For instance, let us consider the previous sentence (3) again: "You may not reproduce the Work".

[7] https://gitlab.tetras-libre.fr/tetras-mars/tenet
[8] https://creativecommons.org/licenses/by-nc/4.0/legalcode

The AMR described above is manually annotated, and the AMR generated by AMRLib tool with this sentence as input is actually quite different:

```
# ::snt You may not reproduce the Work.
(p / possible-01
      :ARG1 (r / reproduce-01
            :polarity -
            :ARG0 (y / you)
            :ARG1 (w / work-01)))
```

Appart from some concepts having changed, the real issue here is the fact that the negation is misplaced in the AMR. This AMR means "You are authorized not to reproduce the work": the negation is placed on the action, which changes the meaning of the modality.

This remark does not point to a weakness of the AMR format but a weakness of some current parsers, that we hope will improve in the future.

5. Preliminary Experiments

The goal of preliminary experiments is not yet to compete with state-of-the-art methods, but to obtain a proof of concept and to illustrate it. So it was chosen to work with fairly simple sentences at first.

5.1. Dataset

Our dataset is made of a hundred simple sentences expressing deontic policies, inspired by ODRL best practices examples [7]. About 20% of these sentences have been manually crafted in order to convey diverse and interesting linguistic phenomena. The other are automatically crafted using a simple grammar. These sentences must contain a modality (which may be negated), one or several actions that may be expressed in several ways. Actions can be simple or composite, and associated with one or two targets and an assignment. Moreover, 10 additional negative entries are added to the dataset. These 10 sentences are grammatically correct but do not express any deontic policy, and should not be recognized as deontic policies by our system. Indeed, many license text sentences are not relevant for the ODRL extraction and may be ignored during the analysis.

5.2. Evaluating experiments

The performance of our system was studied with the dataset described previously. The results are given by the figure 7. Precision (P) and recall (R) were evaluated for the modality and action extraction task. This dataset is used to show that our approach can work on simple cases. In this context, classification of modalities is decent but not excellent, partly since some negated modalities are incorrectly parsed by AMRLib. However, we have a very good precision and recall on the action classification.

Modalities			Actions		
P	R	F	P	R	F
0.780	0.640	0.703	0.989	0.810	0.891

Figure 7: Precision (P), recall (R) and F-measure (F) of the output of our system on our dataset

A classic F-score evaluation was chosen, because we were primarily interested in evaluating the system's ability to extract the expected semantic content. The results obtained confirm the direction followed. That said, our workflow also covers a formalization task, which it would be more relevant to evaluate with a semantic structure evaluation metric such as Smatch [15]. This measure will be particularly interesting to refine our evaluations in future work. The goal remains the processing of real documents, with complex sentences. Additional experimentation is therefore targeted on broader data, such as the dataset of Cabrio et al [8].

6. Conclusion

In this paper, we present a first step towards a workflow that automatically extracts ODRL representations from natural text licenses, with AMR as an intermediate representation. A formal representation of the rights could enable to automatically verify the licence compability in specific situations. Our first experiment provided a proof of concept, with sufficient results on a simple dataset. In particular, we obtained good performance figures for the extraction of modalities, which express permissions, requirements and prohibitions, and actions, which are what the user can, must or is forbidden to do. Some linguistic phenomena are taken into account, such as negations and coordinating conjunctions. Another outcome of the process is that it provides an indirect evaluation of AMR parsers.

Thus, although our full system is far from being complete, our method is achieving encouraging results so far and we plan to continue working on it. First, the effective implementation of a pre-processing phase (Figure 3) would improve runtime performance and help recognize entities and actions. Also, our rule system is still incomplete and more AMR patterns need to be covered. Some evolutions and alternatives can also be considered, such as the use of another pivot representation as an alternative of AMR or as a complementary information.

Acknowledgments

This work was proposed within the CLARA project (https://project.inria.fr/clara/). We thank the CLARA members and our team colleagues for useful discussions.

References

[1] G. Havur, S. Steyskal, O. Panasiuk, A. Fensel, V. Mireles, T. Pellegrini, T. Thurner, A. Polleres, S. Kirrane, Automatic license compatibility checking, in: M. Alam, R. Usbeck, T. Pellegrini, H. Sack, Y. Sure-Vetter (Eds.), Proceedings of the Posters and Demo Track of the 15th International Conference on Semantic Systems co-located with 15th International Conference on Semantic Systems (SEMANTiCS 2019), Karlsruhe, Germany, September 9th - to - 12th, 2019, volume 2451 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019. URL: https://ceur-ws.org/Vol-2451/paper-13.pdf.

[2] B. Moreau, P. Serrano-Alvarado, M. Perrin, E. Desmontils, Modelling the compatibility of licenses, in: P. Hitzler, M. Fernández, K. Janowicz, A. Zaveri, A. J. Gray, V. Lopez, A. Haller, K. Hammar (Eds.), The Semantic Web, Springer International Publishing, Cham, 2019, pp. 255–269.

[3] W3C Recommendation, ODRL Information Model 2.2, 2018. URL: https://www.w3.org/TR/odrl-model/.

[4] L. Banarescu, C. Bonial, S. Cai, M. Georgescu, K. Griffitt, U. Hermjakob, K. Knight, P. Koehn, M. Palmer, N. Schneider, Abstract Meaning Representation for sembanking, in: Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse, Association for Computational Linguistics, Sofia, Bulgaria, 2013, pp. 178–186. URL: https://aclanthology.org/W13-2322.

[5] A. Lamercerie, D. Rouquet, Construction d'ontologies à partir de textes : démonstration d'une approche basée sur l'analyse de graphes amr, Revue des Nouvelles Technologies de l'Information Extraction et Gestion des Connaissances, RNTI-E-39 (2023) 589–596.

[6] A. Lamercerie, Principe de transduction sémantique pour l'application de théories d'interfaces sur des documents de spécification, Thèse, Université Rennes 1 ; Rennes 1, 2021. URL: https://tel.archives-ouvertes.fr/tel-03366457.

[7] W3C ODRL Community Group, ODRL Implementation Best Practices, 2023. URL: https://w3c.github.io/odrl/bp/.

[8] E. Cabrio, A. P. Aprosio, S. Villata, These are your rights - a natural language processing approach to automated rdf licenses generation, in: Extended Semantic Web Conference, 2014.

[9] W3C Member Submission, ccREL: The Creative Commons Rights Expression Language, 2008. URL: https://www.w3.org/Submission/ccREL/.

[10] S. T. Vu, M. Le Nguyen, K. Satoh, Abstract meaning representation for legal documents: An empirical research on a human-annotated dataset, Artif. Intell. Law 30 (2022) 221–243. URL: https://doi.org/10.1007/s10506-021-09292-6. doi:10.1007/s10506-021-09292-6.

[11] O. Abend, A. Rappoport, The state of the art in semantic representation, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 77–89. URL: https://aclanthology.org/P17-1008. doi:10.18653/v1/P17-1008.

[12] I. Langkilde, K. Knight, Generation that exploits corpus-based statistical knowledge, in: 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1, Association for Computational Linguistics, Montreal, Quebec, Canada, 1998, pp. 704–710. URL: https://aclanthology.org/P98-1116. doi:10.3115/980845.980963.

[13] J. May, J. Priyadarshi, SemEval-2017 task 9: Abstract Meaning Representation parsing and generation, in: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 536–545. URL: https://aclanthology.org/S17-2090. doi:10.18653/v1/S17-2090.

[14] Creative Commons, Attribution-NonCommercial 4.0 International Public License, ???? URL: https://creativecommons.org/licenses/by-nc/4.0/legalcode.

[15] S. Cai, K. Knight, Smatch: an evaluation metric for semantic feature structures, in: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Sofia, Bulgaria, 2013, pp. 748–752.