

***Abstract.** The paper describes a problem of analysis and classification of time and terms characteristics in specialized texts. It must be resulted in practical extraction of such characteristics in legal acts regulating administrative activity. We presented a model and computing environment for extraction of named entities for these characteristics and specified its merits and demerits. The created model of the “time and terms” named entities in this type of documents showed a large difference from the entities described in the literature. On the base of the analysis we developed a structural semantic classification of time and terms characteristics of the administrative activity. This classification allowed to reduce a quantity of variants in its descriptions and to formulate a task of automation for controlling the course of administrative procedures.*

***Keywords:** named entities, time and terms characteristics, structural and semantic classification, Russian legal acts, business process automation*

Satoshi Sekine,
 (Named Entity set),
 1996),
 chy), 150

MUC (Grishman et al.,
 IREX (Sekine et al., 2000),
 (Extended Named Entity hierar-
 chy) (Sekine et al., 2002),
 11

Time: “When (TIME) was that?” Period: “How long (PERIOD) was that?”

1

“ ” (AGE),

: 15 6

ì

ì

1. \hat{i} : (, , , ,) ,

2. “ ” : , , , “ ”

3. : “ ”

4. “ ó ” “ ”

\hat{i} , ,

... *II* 2008 ; ... ; ...
... (“ ”)

... *I* ... ; ...
... 10 ... ; ...

... (“ ”)
... (“ ”)

... 23 ... ; ... (“ ”) ... 15

AND OR

... (+) -
 ... (+) -
 ... +) -
 ... 3 90 - (+)
 ... (+) (+) ..
 ... 20 -
 ... () -
 OR. , , -
 ... () 8 ... - ()
 ... (...) (...)...
 ... (+) 15 10 ... (+)
 ...)... (+ +); ()
 ... 3 3 + +)
 (+ + + +)
 ... 15 (+)
 ... (+ +) ..
 “ ” “ ”

, — 2
 , —
 “ ”
 -
 .
 ,
 .
 ,
 .
 ì , -
 ,
 ,
 : ,
 -
 -
 -
 ,
 -
 -
 -
 ,
 -
 -
 -
 :
 : , -
 :
 - /
 -
 :
 -
 -
 ,
 ,
 ,
 ,
 .
 ,
 -
 ,
 -
 ,
 -

Days – : ,

TimeShifter – , . , -

’ ó .

DocLink – : ,

DayNum – , . -

Hour – . : .

TimerFull – : -

TimerSentence – , . -

FullExtraPointers – , .

ListGroups – .

2 . , -

1. , Period -

3 -

DayNum – ,

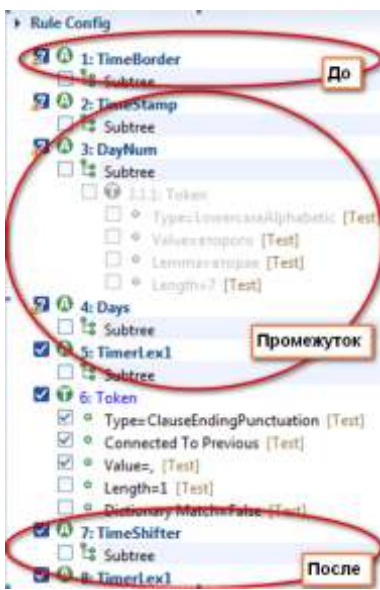
, Token , -
 -
 TimeLex1. -
 -
 (i -
) .

не позднее 12:00 второго рабочего дня, следующего за днем

. 1.



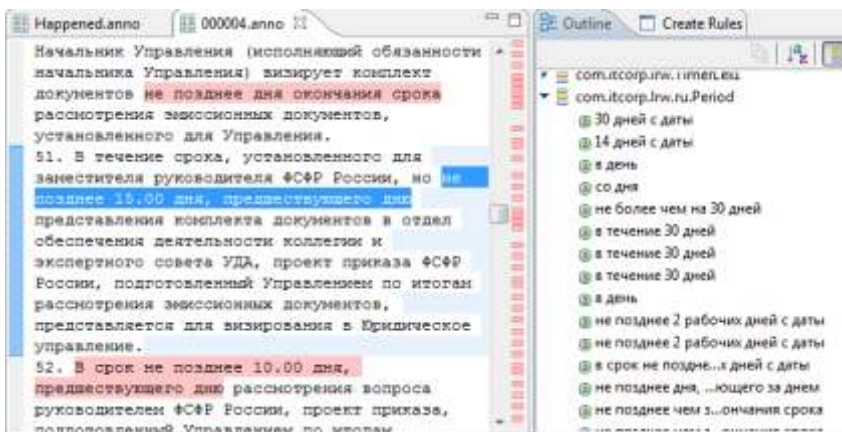
. 2.



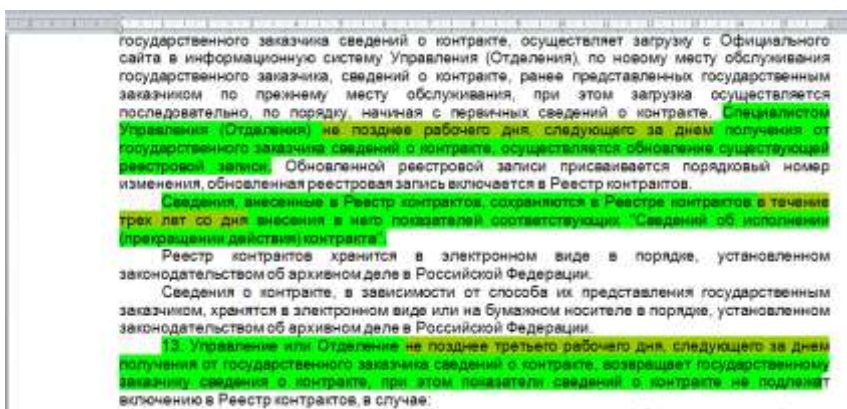
. 3.

UIMA
Content Analytics,
LanguageWare Resource (LRW).
IBM

OASIS
UIMA
Apache.com.
java-
Regex.
LRW,
UIMA,
UIMA.
UIMA,
5.
Period.
UIMA “ ”



. 4.



. 5.

- 1

(1, 2, 3, 5, 6, 7, 10, 12, 30)

(12- ; 10).

1. *Albert Fleischmann*. What Is S-BPM? / S-BPM ONE – Setting the Stage for Subject-Oriented Business Process Management. Communications in Computer and Information Science, 2010, Volume 85, II, 85-106.

2. *Erik F. Tjong Kim Sang* and *Fien De Meulder*. Introduction to the conll-2003 shared task: language-independent named entity recognition. *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003*. Stroudsburg, PA, USA, 2003, Vol. 4, CONLL '03, pp. 142–147.

3. *Grishman R., Sundheim B.* Message understanding conference-6: a brief history. *Proceedings of the 16th conference on Computational linguistics*. Stroudsburg, PA, USA, 1996/Vol. 1, COLING '96, pp. 466–471.

4. *Rau L. F.* Extracting company names from text. *Proceedings of the Seventh Conference on Artificial Intelligence Applications CAIA-92*. Miami Beach, FL, 1991, Vol. I: Technical Papers, pp. 29–32.

5. *Sekine S. and Ishara H.* IREX: IR & IE evaluation project in japanese. *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC 2000)*. 2000.

6. *Sekine S., Sudo K., Nobata Ch.* Extended Named Entity Hierarchy *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)*. 2002.

7. *Stephan H. Sneed*. Exporting Natural Language: Generating NL Sentences Out of S-BPM Process Models. Communications in Computer and Information Science, 2011, Volume 138, Part II, 163-179.

8. () ”,
30.11.1994 N 51- (. 30.11.2011), “ ”,
05.12.1994, N 32, . 3301, “ ”, N 238-239, 08.12.1994.

9. 16 2011 . N 373 “

”, “ ”, 30.05.2011, N 22, . 3169.
10. 31.10.2009 N 879 “

”, “ ”, 09.11.2009, N 45, . 5352.