

გ. გოგინაიშვილი, ს. სერგევი, ლ. ჩხაიძე

**ქართული ენის სინტაქსის არსის ფორმალური სახის აბეზის
ინტელექტუალური სისტემა (სამოქალაქო სარჩელის შინაარსის
ავტომატიზებული კოდირების მახასიათებელი)**

რეზიუმე

სამოქალაქო საქმეთა სასამართლო წარმოების ავტომატიზებული სისტემების შექმნისას დოკუმენტაციის იურიდიულ ხარისხისათვის დიდი მნიშვნელობა აქვს მონაცემთა ბაზის მთლიანობის უზრუნველყოფის საშუალებებს.

მთლიანობის დარღვევის ერთ-ერთ მნიშვნელოვან ფაქტორად, რომელსაც საკმაოდ გავლენა აქვს სისტემის გამოსავალ დოკუმენტებზე, მოგვეჩვენება შეუსაბამობა სარჩელის ტექსტური სახით ფორმულირებული შინაარსის და საქართველოში მიღებული სასამართლო ანგარიშების კლასიფიკატორის ფორმალიზებული კოდების შორის. სასამართლო დოკუმენტაციის ფორმალიზაციაში პერსონალის მონაწილეობა, რომელიც უშუალოდ აგებს პასუხისმგებლობას სასამართლო საქმეთა წარმოებაზე, იწვევს დოკუმენტის შინაარსის დამახინჯების რისკს, რაც დაკავშირებულია ადამიანურ ფაქტორთან.

სტატიაში განხილულია ინტელექტუალური სისტემის არქიტექტურა, რომელიც ახორციელებს ბუნებრივ ქართულ ენაზე დაწერილი სარჩელის ტექსტის იურიდიული შინაარსის კოდის ავტომატიზებულ აგებას.

1. შესავალი

როგორც ცნობილია, მთლიანობის (ე.ი. შინაგანი ლოგიკური წინააღმდეგობის არარსებობის) უზრუნველყოფა წარმოადგენს ერთ-ერთ ძირითად ამოცანას, რომელიც ჩნდება საკმარისად დიდი სირთულის ობიექტების მართვის ავტომატიზებული სისტემის შექმნისას.

როგორც წესია, ამ ამოცანის გადასაწყვეტად ძნელად თუ მოიძებნება უნივერსალური მეთოდი, და ყველა ცალკეულ შემთხვევაში იგი მოიძებნება ორიგინალური შემოქმედებითი წინადადების შემოქმედებითი გამოყენების გზით.

2. ძირითადი ნაწილი

სამოქალაქო საქმეთა სასამართლო წარმოების ავტომატიზებული სისტემა აშკარად ეკუთვნის ამ კლასს, რადგან ამ შემთხვევაში ის შეიცავს მონაცემთა მთლიანობის დარღვევის მთელ რიგ პოტენციურ წყაროს, რომლებსაც შეუძლია უარყოფითი გავლენა ჰქონდეს სისტემის მიერ გამოტანილი დოკუმენტაციის იურიდიულ ხარისხზე.

ეკრძოთ, მთლიანობის დარღვევის ერთ-ერთ ძლიერ ფაქტორად (რომელსაც მნიშვნელოვანი გავლენა აქვს სისტემის გამოსავალ დოკუმენტებზე) მოგვეჩვენება შეუსაბამობა სარჩელის ტექსტური სახით ფორმულირებული შინაარსის და საქართველოში მიღებული სასამართლო ანგარიშების კლასიფიკატორის ფორმალიზებული კოდების შორის.

ეს ხდება იმიტომ, რომ ორივე ხსენებული რეკვიზიტი (რომლებს აქვს დიდი მნიშვნელობა სტატისტიკური ანგარიშის ფორმების უზრუნველყოფის თვალსაზრისით), ივსება სასამართლოს თანამშრომლის მიერ, ე.ი. იმ პირობით, რომელიც უშუალოდ აგებს პასუხისმგებლობას სასამართლო საქმეთა წარმოებაზე.

როგორც კარგადაა ცნობილი, ადამიანის მონაწილეობა ფორმალური დოკუმენტის შექმნის პროცესში აუცილებლად იწვევს ამ პროცესში ადამიანური ფაქტორის ჩართვას. ამის შედეგად მონაცემთა ბაზაში მოთავსებული რეკვიზიტები შეიძლება იყოს განმარტებული (და ბაზაში შეტანილი) ერთმანეთის საწინააღმდეგო მნიშვნელობებით, რაც ამ რეკვიზიტების შემცავ ყველა დოკუმენტის შინაარსის დამახინჯების რისკის ტოლფასია.

სასამართლო დოკუმენტაციის მანქანური სახის ფორმირებაში მონაწილე პერსონალის სპეციალური მომზადების ჩატარება უთუოდ შეამცირებს ავტომატიზებული სისტემის მსგავსი ხარვეზის ალბათობას, მაგრამ, როგორც ადასტურებს პრაქტიკა, მისი მთლიანი გამორიცხვა პრაქტიკულად შეუძლებელია. ამ ამოცანის რადიკალური გადაწყვეტა შესაძლებელია მხოლოდ პროგრამული საშუალებების შექმნის გზით, რომლებიც გარკვეული სიზუსტით ავტომატურად შეასრულებენ ამ ფუნქციას.

აღნიშნულ ამოცანის სირთულე იმაში მდგომარეობს, რომ სასამართლო საქმეთა წარმოების ერთ-ერთი რეკვიზიტი ჩაწერილია არა კოდის სახით, არამედ ბუნებრივი ენის საშუალებით, ჩვენს შემთხვევაში – შეზღუდულ ქართულ ენაზე. ეს ნიშნავს, რომ ამ ამოცანის გადასაწყვეტად შექმნილ სისტემას მნიშვნელოვანი ინტელექტუალური შესაძლებლობა უნდა ჰქონდეს, რომელიც აუცილებელია შესაბამისი ტექსტის შინაარსის მანქანური სახის ავტომატური აგებისათვის [1,2].

3. როგორც ავტომატიზაციის ობიექტის წინასწარმა შესწავლამ გვიჩვენა, მითითებულ რეკვიზიტებში გამოყენებულ ტექსტებს ახასიათებთ ამოცანისათვის მნიშვნელოვანი შემდეგი თვისება:

(1) ტექსტის ლექსიკური შემადგენლობა შეზღუდულია იურიდიულ და მასთან დაკავშირებული პროცესუალური მოქმედებათა ტერმინოლოგიით, მაგალითად:

შრომითი მოვალეობის შესრულებისას ზიანის ანაზღაურება

(2) ტექსტის სინტაქსური სტრუქტურა შეზღუდულია იურიდიულად კორექტული ფორმულირებებით, რომლებსაც აქვთ მხოლოდ სინტაგმის ფორმა, ე.ი. ისინი არ შეიცავენ ზმნის პირიან ფორმებს და შეიცავენ მხოლოდ სახელზმნებს, რომლებს სინტაქსურად ექვემდებარებათ 5-დან 6-მდე არსებითი და ზედსართავი სახელი:

ანაზღაურება -> ზიანის -> შესრულებისას -> მოვალეობის -> შრომითი

(3) ტექსტის მორფოლოგიური ანალიზი დაიყვანება სინტაქსური კავშირების დამყარებისათვის საჭირო ბრუნვების დადგენაზე:

შრომითი\შრომა\INSTRUM

მოვალეობის\მოვალეობა\GENIT

შესრულებისას\შესრულება\BY

ზიანის\ზიანი\GENIT

ანაზღაურება\ანაზღაურება\NOMIN

(4) ტექსტის მიერ აღწერილი სცენის აქტორი წარმოდგენილია სახელზმნით ნომინატივის ბრუნვაში, რომელიც, როგორც წესი, ტექსტის ბოლო პოზიციაში იმყოფება:

ანაზღაურება\NOMIN

(5) აქტორის მოქმედების ობიექტი წარმოდგენილია მასთან უშუალოდ დაქვემდებარებული არსებითი სახელით ნათესაობით ბრუნვაში:

ზიანი\GENIT

(6) აქტორის მოქმედების ატრიბუტებს წარმოადგენენ არსებითი და ზედსართავი სახელები, რომლებიც ექვემდებარებიან აქტორს სხვა ბრუნვების საშუალებით:

შესრულება\BY

მოვალეობა\GENIT

შრომა\INSTRUM

(7) სინტაგმის შიგნით მიზეზობრივი კავშირის დასადგენად გამოიყენება შესაძლო აქტორების და მათთვის დაშვებული კორექტული იურიდიული შინაარსის მქონე ობიექტების მონაცემთა ბაზა:

ანაზღაურება -> ზიანი

(8) სინტაგმის კოდური სახის აგება წარმოებს მონაცემთა ბაზაში განსაზღვრული აქტორის საკლასიფიკაციო კოდიდან:

ანაზღაურება ->>> CODE1

(9) მოცემული აქტორის საკლასიფიკაციო კოდის რამდენიმე ვარიანტის არსებობის შემთხვევაში, კონკრეტული საკლასიფიკაციო მუხლის არჩევა პირველ რიგში განისაზღვრება აქტორთან რელაციურად დაკავშირებული ობიექტით, რომელიც მონაცემთა ბაზიდან შეირჩევა.

ზიანი ->>> CODE2

(10) აქტორის მოქმედების ობიექტის ატრიბუტები შეიძლება იყოს გამოყენებული მხოლოდ სინტაგმის ელემენტების თანადობის უფრო მეტად განვითარებული სტრუქტურის შემთხვევაში.

დასკვნა

4. ზემოთ აღწერილ სპეციფიკაციათა სისტემის რეალიზაცია მოითხოვს სპეციალური პროგრამული სისტემის შექმნას, რომელიც განახორციელებს შემდეგ ალგორითმიულ ფუნქციებს:

(1) სინტაგმაში შემავალი სიტყვაფორმების მორფოლოგიური ანალიზის მოდული;

(2) სინტაგმაში გამოყოფილი ფუძეთა ლექსიკური ანალიზის მოდული;

(3) სინტაგმის აქტორის ძებნის მოდული;

(4) სინტაგმის აქტორის ობიექტის ძებნის მოდული;

(5) სინტაგმის ობიექტის ატრიბუტების ძებნის მოდული;

(6) კლასიფიკაციური კოდის აგების მოდული, რომელიც მოქმედებს მიღებული სინტაგმის გრამატიკული სტრუქტურის მიხედვით და იყენებს სისტემის მონაცემთა ბაზაში აღწერილ შესაძლო სემანტიკურ შეზღუდვებს.

მნიშვნელოვანია, რომ (3)-(5) მოდულების ფუნქციონირება არ შეიძლება იყოს აღწერილი კონტექსტურად თავისუფალი გრამატიკის ტერმინების საშუალებით, რის გამოც მათი რეალიზაცია შესაძლებელია მხოლოდ ეკტრეინგის პრინციპის გამოყენების ბაზაზე.

ლიტერატურა

1. Искусственный интеллект. Справочник в 3 томах. М., 1980.
2. Э.В.Попов, А.В.Преображенский. Особенности реализации ЕЯ-систем. М., 1990.

Г.Г. Гогичайшвили, С. Сергеев, Л.Л. Чхаидзе

**ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПОСТРОЕНИЯ ФОРМАЛЬНОГО
ОБРАЗА СМЫСЛА СИНТАГМЫ ГРУЗИНСКОГО ЯЗЫКА
(на примере автоматизированного кодирования содержания
гражданского иска)**

Резюме

При создании автоматизированных систем ведения судебного делопроизводства для обеспечения высокого юридического качества документации важное значение имеют средства поддержки целостности входящих в базы данных документов.

Одним из существенных источников нарушения целостности, существенно сказывающимся на выходных документах системы, является несоответствие между текстовым образом формулировки содержания иска и ее формализованным кодом для принятого в Грузии классификатора судебной отчетности. Вследствие участия в процессе создания формализованного документа лица, ответственного за судебное делопроизводство, в данном звене функционирования информационной системы неизбежно возникновение риска искажения смысла, обусловленного человеческим фактором.

В работе рассматривается архитектура интеллектуальной системы, предназначенной для автоматизированного построения возможных кодов содержания текста на естественном грузинском языке, описывающего юридическую формулу иска.

G.Gogichaishvili, S.Sergeev, L.Chkhaidze

**INTELLECTUAL SYSTEM OF CONSTRUCTION OF THE GEORGIAN LANGUAGE
SYNTAGMA FORMAL IMAGE SENSE (ON AN EXAMPLE OF THE AUTOMATED
CODING OF THE CONTENT OF THE CIVIL SUIT)**

Summary

At creation of the automated systems of conducting judicial office-work for maintenance of high legal quality documentation, means of integrity support of the documents entering databases are of great value.

One of the main sources of integrity infringement essentially affecting outcome documents of the system is the difference between text image of the claim contents formulation and its formalized code for the classifier of the judicial reporting accepted in Georgia. Owing to the participation of the person responsible for judicial office-work during creation of the formalized document, in the given part of functioning of information system the risk of sense distortion occurrence, which can be caused by the human factor, is inevitable.

In work the architecture of the intellectual system intended for automated construction of possible codes of the contents of the text in natural Georgian language, the claim describing the legal formula is considered.

The work considers the architecture of the intellectual system meant for the automated building of the possible codes of the text contents on the natural Georgian language describing judicial formula of the suit.