

UNIVERZITET U BEOGRADU
RUDARSKO-GEOLOŠKI FAKULTET

Olivera M. Kitanović

ONTOLOŠKI MODEL
UPRAVLJANJA RIZIKOM U RUDARSTVU

Doktorska disertacija

Beograd, 2021

UNIVERSITY OF BELGRADE
FACULTY OF MINING AND GEOLOGY

Olivera M. Kitanović

**AN ONTOLOGY-BASED MODEL FOR RISK
MANAGEMENT IN MINING**

Doctoral Dissertation

Belgrade, 2021

Mentor:

Dr Ranka Stanković, vanredni profesor
Univerzitet u Beogradu, Rudarsko-geološki fakultet
Naučna oblast: Matematika i informatika

Članovi komisije:

Dr Nikola Lilić, redovni profesor
Univerzitet u Beogradu, Rudarsko-geološki fakultet
Naučna oblast: Zaštita na radu i zaštita životne sredine

Dr Aleksandar Cvjetić, redovni profesor
Univerzitet u Beogradu, Rudarsko-geološki fakultet
Naučna oblast: Zaštita na radu i zaštita životne sredine

Dr Igor Miljanović, redovni profesor
Univerzitet u Beogradu, Rudarsko-geološki fakultet
Naučna oblast: Računarstvo i sistemsko inženjerstvo

Dr Branislava Šandrih, docent
Univerzitet u Beogradu, Filološki fakultet
Naučna oblast: Bibliotečka informatika

Datum odbrane:

ONTOLOŠKI MODEL UPRAVLJANJA RIZIKOM U RUDARSTVU

Sažetak

Rudarska proizvodnja obuhvata kompleksne tehnološke sisteme, što nameće potrebu za uspostavljanjem i unapređivanjem sistema upravljanja rizikom. Heterogenost i obim podataka neophodnih za upravljanje rizikom zahtevaju sistem koji ih na fleksibilan način integriše i omogućava njihovo optimalno korišćenje.

Osnovni cilj ove disertacije je razvoj ontologije za domen rudarstva i na njoj zasnovanog modela za upravljanje rizikom. Njegova realizacija podrazumeva i implementaciju algoritama ekstrakcije informacija za popunjavanje ontologije, kao i odgovarajuće softversko rešenje.

Razvoj modela obuhvata i značajno proširenje rudarskog korpusa, kao i kreiranje terminološke baze podataka, realizovano korišćenjem metoda računarske lingvistike i korpusa dokumenata iz oblasti rudarstva (planova, izveštaja, zakona, udžbenika i monografija). Korišćena je i deskriptivna metoda za sistematizaciju podataka, zatim konačni automati i statističke analize za ekstrakciju informacija, kao i komparativna i analitička istraživačka metoda za vrednovanje i interpretaciju dobijenih rezultata.

Za razvoj modela korišćeni su alati informacionih tehnologija: UML za modeliranje koncepata, OWL za razvoj ontologije, SWRL pravila za mehanizam zaključivanja, upitni jezici CQL nad korpusom i SPARQL nad ontologijom.

Rezultati istraživanja pokazuju da je moguće formalizovati informacije i znanje o rizicima u rudarstvu, te razviti model koji će unaprediti efikasnost upravljanja rizikom i pomoći menadžmentu rudnika u donošenju odluka o primeni mera za smanjenje uticaja rizika identifikovanih u rudniku.

Ostvarenjem ciljeva ove disertacije dat je doprinos povećanju efikasnosti u identifikaciji, analizi i reagovanju na rizik kroz izgradnju specifične domenske ontologije za rizike u rudarstvu.

Ključne reči: rudarstvo, rizik, upravljanje rizikom, procena rizika, ontologija, semantička mreža, ekstrakcija informacija, upravljanje znanjem, računarska lingvistika

Naučna oblast: Rudarsko inženjerstvo

Uža naučna oblast: Računarstvo i sistemsko inženjerstvo

UDK broj: 004:004.42/.82:519.8
005.1/.7:005.334:622
519.237/.245:519.76:622
614.8:622:681.3:669
811.163:41'366
81'.32:81'366
81'.32:81'367(043.3)

AN ONTOLOGY-BASED MODEL FOR RISK MANAGEMENT IN MINING

Abstract

Mining production involves complex technological systems, which calls for the need to create and improve risk management systems. The heterogeneity and scope of data necessary for risk management require a system that integrates them in a flexible way and enables their optimal use.

The main goal of this dissertation is to develop an ontology for the mining domain and a risk management model based on it. Its realization includes the implementation of information extraction algorithms for improving the ontology, as well as an appropriate software solution.

The development of the model includes a significant expansion of the mining corpus, as well as the creation of a terminological database, realized using methods of computational linguistics and a corpus of documents from the mining domain (plans, reports, laws, textbooks and monographs). For systematization of data a descriptive method was used, finite automata and statistical analyzes for information extraction, and comparative and analytical research methods for evaluation and interpretation of the obtained results.

Information technology tools were used for model development: UML for concept modeling, OWL for ontology development, SWRL rules for inference mechanism, query languages CQL for corpus and SPARQL for ontology.

The research results show that it is possible to formalize information and knowledge about risks in mining and develop a model that will improve the efficiency of risk management and assist mine management in making decisions on implementing measures to reduce the impact of risks identified in a mine.

Achieving the goals of this dissertation has contributed to increasing efficiency in identification, analysis and response to risk by developing a specific domain ontology for risks in mining.

Key words: mining, risk, risk management, risk assessment, ontology, semantic network, information extraction, knowledge management, computational linguistic

Scientific field: Mining Engineering

Scientific subfield: Computer science and systems engineering

UDC Number:

UDC Number: 004:004.42/.82:519.8

005.1/7:005.334:622

519.237/.245:519.76:622

614.8:622:681.3:669

811.163:41'366

81'.32:81'366

81'.32:81'367(043.3)

Sadržaj

1. Uvod.....	1
1.1. Predmet istraživanja.....	1
1.2. Cilj istraživanja	2
1.3. Polazne hipoteze.....	3
1.4. Naučne metode istraživanja.....	4
1.5. Struktura disertacije sa kratkim pregledom poglavlja.....	5
2. Pregled dosadašnjih istraživanja.....	7
2.1. Rizici u rudarstvu	7
2.2. Upravljanje rizikom.....	15
2.2.1. Uspostavljanje konteksta.....	17
2.2.2. Identifikacija rizika	18
2.2.3. Analiza rizika	19
2.2.4. Procena rizika.....	19
2.2.5. Tretman rizika.....	20
2.3. Softveri za upravljanje rizikom.....	21
2.4. Upravljanje rizikom zasnovano na ontologijama.....	26
2.4.1. Ontološke platforme za upravljanje rizikom.....	26
2.4.2. Klasifikacija teksta u upravljanju rizicima	30
2.5. Ekstrakcija koncepata i njihovih relacija	33
3. Metodologija i resursi za razvoj ontologije.....	35
3.1. Šta je ontologija?	35
3.2. Resursi za ekstrakciju informacija	39
3.2.1. Rečnici iz domena rudarstva.....	39
3.2.2. Korpusi tekstova	40
3.2.3. Morfološki rečnici.....	41
3.2.4. Obrada teksta na prirodnom jeziku - Unitex.....	45
3.2.5. Leximir i Wordnet	47
3.3. Semantička mreža okvira FrameNet.....	50
3.4. Izgradnja ontologije.....	54
3.4.1. Principi razvoja ontologije	54
3.4.2. Automatizacija popunjavanja ontologije.....	55
4. Model sistema za upravljanje rizikom	58
4.1. Terminološki resursi i korpusi	58
4.1.1. Arhitektura sistema za ekstrakciju ontoterminologije	58
4.1.2. Dopuna korpusa RudKorp.....	65
4.1.3. Terminološki resursi.....	66
4.2. Ekstrakcija važnih pojmova.....	67
4.3. Ekstrakcija rudarskih entiteta iz teksta.....	77
4.3.1. Entiteti rudarske opreme	77
4.3.2. Entiteti rudarskih lokacija.....	79
4.3.3. Entiteti otkaza opreme i kvarova.....	81
4.3.4. Entiteti događaja na kopovima	83
4.3.5. Organizovanje pojmova u hijerarhiju	85
4.4. Ekstrakcija relacija između koncepata.....	87
4.4.1. Okidač „kao što je/su“	87
4.4.2. Okidač „je vrsta“	88
4.4.3. Okidač „drugi, ostali, “.....	89
4.4.4. Uzročne veze.....	90
4.5. UML model ontologije	92

4.6. Razvoj ontologije u Protégé.....	97
4.6.1. Koncept razvoja ontologije	97
4.6.2. Izgradnja ontologije.....	99
4.6.3. Mehanizam zaključivanja, upiti i popunjavanje ontologije	103
4.6.4. Predlog interfejsa za unos rizika.....	108
4.7. Publikovanje ontologije alatom VocBench.....	110
5. Evaluacija rezultata na primerima iz prakse	113
5.1. Evaluacija ekstrakcije informacija.....	113
5.1.1. Evaluacija pokrivenosti domena.....	113
5.1.2. Evaluacija prepoznavanja domenskih entiteta	115
5.1.2.1. Preliminarni test.....	115
5.1.2.2. Ispravka transduktora za entitete rudarske opreme na osnovu rezultata evaluacije	116
5.1.2.3. Ispravka transduktora za lokacije na osnovu rezultata evaluacije	117
5.1.2.4. Evaluacija prepoznavanja rudarskih događaja	119
5.2. Evaluacija na primeru upravljanja homogenizacijom uglja.....	119
5.3.	121
5.4. Evaluacija na primeru procene rizika na procesa zaštite na radu	122
6. Zaključak.....	124
6.1. Doprinos	124
6.2. Pravci za dalji rad	125
7. Literatura	127

Spisak slika

Slika 2-1 UML model upravljanja rizikom.....	16
Slika 2-2 Primeri panela softvera iAuditor	22
Slika 2-3 Primeri panela softvera Designsafe za definisanje zadataka, opasnosti i njihovih uzroka.....	23
Slika 2-4 Primeri panela softvera Designsafe za definisanje verovatnoće i mera	23
Slika 2-5 Delovi panela alata PHA za procenu rizika	24
Slika 2-6 Delovi panela softvera ProcenaRizika za procenu rizika	26
Slika 3-1 Od indeksa pojmova do ontologije.....	35
Slika 3-2 Deo panela sa vikipodacima za pojam “upravljanje rizikom”	36
Slika 3-3 Panel alate Leximirka za reč “rizikovati”	43
Slika 3-4 Panel alate Leximirka za reč “odnos uzrok-posledica”	45
Slika 3-5 Panel alata Unitex sa obrađenim rudarskim korpusom	46
Slika 3-6 Panel alata Unitex sa konkordancama za upit „<N+Hum><rizik>“	47
Slika 3-7 Panel alata Leximir sa sinsetima srpskog i engleskog WordNeta.....	48
Slika 3-8 Panel alata Leximir sa stablom nadređenih i podređenih sinseta „rizik“	49
Slika 3-9 Panel alata Leximir sa panelom za ekstrakciju višečlanih reči iz korpusa	49
Slika 3-10 Semantički okvir Rizična_situacija (eng. Risky_situation).....	51
Slika 3-11 Semantički okvir Scenario_rizika (eng. Risk_scenario)	52
Slika 3-12 Semantički okvir Biti_u_opasnosti (eng. Being_at_risk).....	53
Slika 3-13 Semantički okvir Rizikovati (eng. Run_risk)	53
Slika 4-1 Arhitektura Sistema za ekstrakciju i publikovanje terminologije	58
Slika 4-2 Terminologija iz oblasti rudarstva.....	62
Slika 4-3 Terminologija iz oblasti rizika.....	63
Slika 4-4 Primer RDF serijalizacije termina iz oblasti rizika.....	64
Slika 4-5 Veb alat za anotaciju teksta vrstom reči i lemom	66
Slika 4-6 Deo panela sa konkordancama sa prikazanim anotacijama vrstom reči i lemom	66
Slika 4-7 Konačni automat za ekstrakciju strukture pridev-imenica	70
Slika 4-8 Konkordance obrasca pridev imenica za “rizik”	74
Slika 4-9 Skica reči “rizik” u alatu Sketch engine	75
Slika 4-10 Konkordance u alatu Sketch engine.....	75
Slika 4-11 Frekvencije kolokacija za reč rizik u alatu Sketch engine.....	76
Slika 4-12 Skica razlika upotrebe reči rizik i opasnost.....	76
Slika 4-13 Graf za ekstrakciju entiteta rudarske opreme	77
Slika 4-14 Graf za ekstrakciju bagerskih Sistema.....	78
Slika 4-15 Graf za ekstrahovanje pojavljivanja bagera i njihovih naziva.....	78

Slika 4-16 Graf transduktor za prepoznavanje i obeležavanje lokacija	79
Slika 4-17 Podgraf za obeležavanje rudarskih preduzeća (institucija)	79
Slika 4-18 Podgraf za obeležavanje rudarskih basena kao lokacija	80
Slika 4-19 Podgraf za obeležavanje termoelektrana kojima se isporučuje uglj	80
Slika 4-20 Podgraf za obeležavanje rudarskih kopova i izvorišta.....	81
Slika 4-21 Specijalan slučaj - Tamnava i njena polja.....	81
Slika 4-22 Graf za prepoznavanje događaja vezanih za opremu i lokacije	83
Slika 4-23 Tekst obeležen XML etiketama	84
Slika 4-24 Vizuelizacija obeleženog teksta	84
Slika 4-25 Graf koji prepoznaje imeničke fraze	85
Slika 4-26 Graf koji ekstrahuje sinonimske imeničke fraze iz teksta.....	86
Slika 4-27 Graf koji iz teksta ekstrahuje hiperonime sa listama hiponima	86
Slika 4-28 Graf koji ekstrahuje hiponime i njihov hiperonim.....	87
Slika 4-29 Konkordance za upit „<proces> kao što <jesam>“	88
Slika 4-30 Konkordance za upit „rizik ... kao što...je“	88
Slika 4-31 Konkordance za upit „je ... vrsta ...“	89
Slika 4-32 Konkordance za upit „ ... i ili...drugi ostali...“	89
Slika 4-33 Konkordance za upite sa trigerima „uključujući“ i „poznat kao“	90
Slika 4-34 Ključni koncepti upravljanja rizikom i njihove relacije.....	90
Slika 4-35 UML model osnovnih koncepata za identifikaciju rizika.....	93
Slika 4-36 UML model osnovnih koncepata za analizu i procenu rizika	94
Slika 4-37 UML model organizacije rada	96
Slika 4-38 Klasa Rizik povezana sa drugim ontologijama	98
Slika 4-39 Klase matrice rizika	99
Slika 4-40 Potklase uzroka i potklase posledica	100
Slika 4-41 Klase kojima se modeliraju događaj, posledice i preventivne mere	101
Slika 4-42 Dekompozicija rudarskih procesa.....	102
Slika 4-43 Klase organizacije operativnih aktivnosti.....	102
Slika 4-44 Operativni rizici.....	103
Slika 4-45 Primer instanci potklase Uzroci rizika po zdravlje	104
Slika 4-46 Primeri instanci potklase Posledice na zdravlje i bezbednost.....	104
Slika 4-47 Lokacije pre primene pravila zaključivanja	105
Slika 4-48 Lokacije posle primene mehanizma zaključivanja	105
Slika 4-49 Primer SWLR upita za selekciju podataka o opremi.....	106
Slika 4-50 Primer SWLR upita za grupisanje i prebrojavanje podataka	106
Slika 4-51 Primer SPARQL upita za selekciju posledica povezanih sa rizikom	107

Slika 4-52 Primer SPARQL upita za izbor lokacija	107
Slika 4-53 Panel za unos događaja i procenu rizika	108
Slika 4-54 Panel za evidentiranje uzroka i posledice.....	109
Slika 4-55 Panel za unos preventivnih mera	109
Slika 4-56 Panel za definisanje procesa i lokacije događaja	109
Slika 4-57 VocBench sa primerima klasa i primeraka publikovane ontologije	110
Slika 4-58 VocBench sa svojstvima publikovanje ontologije	111
Slika 4-59 VocBench sa svojstvima publikovanje ontologije	112
Slika 5-1 Ispravljeni transduktor za prepoznavanje opreme	116
Slika 5-2 Ispravljeni podgraf za prepoznavanje bagerskih Sistema.....	116
Slika 5-3 Ispravljeni podgraf za prepoznavanje tipova i naziva bagera	117
Slika 5-4 Ispravljeni podgraf za prepoznavanje ogranaka.....	118
Slika 5-5 Ispravljeni podgraf za prepoznavanje kopova i polazišta.....	118
Slika 5-6 Ispravljeni graf za pronalaženje entiteta tamnavskih polja.....	119
Slika 5-7 Pozitivni uticaj uvođenja planirane homogenizacije.....	120
Slika 5-8 Opis elemenata u mreži.....	121
Slika 5-9 Dekompozicija postrojenja za preradu	122
Slika 5-10 Primer vizuelizacije rizika sa njegovim elementima.....	122
Slika 5-11 Primeri instanci vlasnik rizika.....	123
Slika 5-12 Primeri SWRL pravila	123

Spisak tabela

Tabela 2-1 Obrasci za prepoznavanje relacije sinonimije (gde su S1 i S2 ontološki koncepti)	34
Tabela 2-2 Obrasci za prepoznavanje obraza reči	34
Tabela 3-1 Rečnik prostih (jednočlanih) reči	41
Tabela 3-2 Rečnik složenih reči.....	44
Tabela 4-1 Primeri leksičkih funkcija.....	60
Tabela 5-1 Detektovani rizici koji ugrožavaju implementaciju homogenizacije.....	121

1. UVOD

1.1. Predmet istraživanja

Rudarska proizvodnja odvija se u specifičnom prirodnom okruženju i odlikuje je upotreba kompleksnih tehnoloških sistema, što je čini jednom od najrizičnijih industrijskih grana sa velikom količinom neizvesnosti. Pored uobičajenih činilaca poslovanja kao što su tržišna konkurentnost, fluktuacija ljudskih resursa, finansijski resursi, inovativni pristupi, koji su prisutni i u ostalim privrednim granama, rudarstvo dodatno odlikuju ograničeni prirodni resursi, veliki uticaj na životnu sredinu, teški uslovi rada, složena zakonska regulativa i sve veća zainteresovanost šire društvene zajednice. S toga se dodatno nameće potreba za uspostavljanjem i stalnim unapređivanjem sistema upravljanja rizikom, za identifikovanjem rizičnih situacija i blagovremenim reagovanjem na sprečavanje neželjenih događaja, odnosno iskorišćavanje povoljnih prilika.

Kada se pomene rudarstvo i rizik, prva asocijacija jeste bezbednost na radu i to je upravo oblast kojoj je posvećena naročita pažnja i posebna zakonska regulativa, a primena sistema upravljanja rizikom postala je obavezan zahtev. Primena procene rizika bitna je ne samo za obezbeđenje sigurnog radnog okruženja, već predstavlja i zakonsku obavezu u najvećem broju država sveta. Kada je u pitanju upravljanje rizicima u preostalim činionicima rudarskih poslovnih sistema, možemo reći da je nivo prihvaćenosti tog koncepta različit u različitim delovima sveta.

Efikasnost upravljanja rizikom zavisi od procesa procene rizika, čiji rezultati pomažu menadžmentu rudnika u donošenju odluka o merama koje će se primeniti za smanjenje identifikovanih rizika. Imajući u vidu kompleksnost, heterogenost i obim podataka potreban za procenu i upravljanje rizikom, postoji potreba za sistemom koji će, sa jedne strane imati fleksibilan model za integraciju podataka različitih formata, šema i sistema klasifikacije, a sa druge biti robustan i sposoban da obradi veliku količinu podataka i omogući višekorisnički distribuiran rad. Velika količina elektronski dostupnih podataka, nastalih u različitim softverskim okruženjima, često i na različitim jezicima i pismima, dovodi do potrebe za obradom strukturiranih, slabo strukturiranih i nestruktuiranih skupova podataka, uz mogućnost ekstrakcije informacija i njihovog prevođenja u strukturirano znanje (Tserng et al. 2009). Ontologije nude mogućnost za razvoj sistema za upravljanje rizikom kojima se rešavaju pomenuti problemi vezani za neophodne podatke.

Jedan od preduslova za razvoj ontološkog modela za upravljanje rizikom je kreiranje terminološke baze podataka sa pojmovima iz oblasti rudarstva. Savremeni pristupi podrazumevaju izradu kolekcija različitih vrsta tekstova iz oblasti rudarstva i upravljanja rizikom kao podloge za izradu terminoloških baza, kao i kreiranje baze znanja specifične za rudarski domen. U pristupu izloženom u ovoj disertaciji, razvoj ontologije zasniva se na povezivanju terminoloških rečnika iz oblasti rudarstva korišćenjem tehnologija obrade velikih količina podataka (eng. big data), kao i metoda matematičke i računarske lingvistike. Predstavljene su glavne karakteristike ovog pristupa, podaci korišćeni za formiranje terminološke baze podataka, dostupni terminološki izvori – tradicionalni papirni rečnici i digitalni izvori koji se odnose na domen rudarstva, kao i opšti morfološki rečnici neophodni za obradu i razumevanje teksta. Priprema resursa započeta je digitalizacijom rečnika i proširenjem korpusa, što je podrazumevalo i dodavanje termina u opšte morfološke rečnike. Razvoj baze znanja oslanja se na analizu korpusa, čiji su detalji takođe predstavljeni. Primeri upotrebe, kolokacije i sinonimi igraju važnu ulogu u razumevanju pojmova specifičnih za pojedini domen, tako da su korišćene različite metode njihove ekstrakcije.

Metode analize velikih količina podataka otvorile su nove mogućnosti za analizu korpusa obimnih tekstualnih podataka i ekstrakciju podataka i znanja iz njih. Tako, na primer, Chen i njegove kolege (L.-C. Chen, Chang, and Chung 2020) predlažu novi, statistički zasnovan pristup obradi korpusnih tekstova, koji obrađuje jezik u određenom domenu, a koji se zasniva na uspostavljanju liste funkcionalnih reči i njenom ugrađivanju u softversko rešenje, kako bi se proširila lista ključnih reči i fraza značajnih za konkretni domen. Cilj je da se poboljša efikasnost obrade korpusa stručnih tekstova, počev od pripremnih radova, praćenih generisanjem sirovih podataka, optimizacijom procesa pa sve do generisanja željenih podataka. U razvoju sistema prezentiranog u ovoj disertaciji usvojen je pristup zasnovan na podacima, uz oslanjanje na tekstualne, leksičke i terminološke resurse, bez obzira na format njihovog zapisa. Unutar razvoja ovog sistema štampani resursi, papirni rečnici koji pokrivaju terminologiju rudarstva i rizika sistematski su digitalizovani i harmonizovani sa resursima u elektronskom obliku.

U literaturi se susrećemo sa različitim definicijama rizika, u zavisnosti od obima i namene rečnika ili priručnika, tako da se može naći stav da je rizik „funkcija matematičkog očekivanja verovatnosti događaja“ (Bernoulli 1896), potom „opasnost, izlaganje opasnosti, mogućnost stradanja, gubitak...“ (Vujanić et al. 2011), dok se prema standardu ISO 31000:2015 rizik definiše kao „efekat neizvesnosti na ciljeve“. Ako je efekat odstupanje od očekivanog ishoda, neizvesnost stanje nedostatka informacija, a cilj ishod kojem se teži, ova konstatacija se može protumačiti kao da je rizik svaka situacija koja nije izvesna, a koja ima negativan ishod. Rizik je prisutan u situaciji gde jedna aktivnost vodi ka nekoliko različitih, međusobno isključivih ishoda, od kojih svi imaju poznatu verovatnoću nastanka (Nišić 2019). Treba napomenuti da iako je pažnja najviše posvećena prevencijama od negativnih ishoda rizičnih događaja, savremeni pristupi nalažu podjednako tretiranje rizika koji mogu imati pozitivne efekte na organizaciju i tada ih nazivamo prilike.

1.2. Cilj istraživanja

Cilj doktorske disertacije je da razvije model za upravljanje rizikom i odgovarajuće softversko rešenje, pri čemu se model zasniva na ontologijama pojmova vezanih za rizik, čiji je razvoj takođe detaljno opisan. Model podrazumeva i ekstrakciju znanja, koja treba da obuhvati eksplicitno i implicitno znanje, odnosno činjenice i mehanizam zaključivanja, kojim se dolazi do novih znanja.

Ostvarenjem ciljeva ove doktorske disertacije daje se doprinos povećanju efikasnosti u identifikaciji, analizi i reagovanju na rizik, kroz izgradnju specifične domenske ontologije za rizike u rudarstvu. Pored eksplicitnog znanja, ontološki pristup ima za cilj da obuhvati, i implicitna, prećutna (eng. tacit) ekspertska znanja iz različitih heterogenih izvora. Razvoj jedne takve ontologije svakako zahteva znatno vreme, troškove i stručnosti. U eri dostupnosti velike količine, uglavnom tekstualnih podataka (eng. big data) postavlja se pitanje ekstrakovanja informacija i znanja iz narativnog, nestruktuiranog teksta. Rudarstvo, kao i druge oblasti industrije, koristi specifičnu terminologiju, na kojoj počiva izgradnja sistema koji ima mogućnost da obrađuje rudarska dokumenta, od planova, izveštaja, pa do zakona, udžbenika i monografija, i koji iz njih crpi činjenice i znanja važna za rudarstvo i rizike u rudarstvu, a zatim ih strukturira u ontologiju.

Upravljanje rizikom ne znači nužno njegovu minimizaciju, jer se tako mogu izgubiti potencijalne mogućnosti, tako da stepen rizika kom se neka kompanija izlaže uveliko zavisi od načina odlučivanja i stepena pouzdanosti informacija na osnovu kojih se donose odluke (Whittle 1997).

Svrha modela izloženog u ovoj disertaciji je da se postigne potrebni nivo složenosti sa ciljem predviđanja pristupa koji bi pratio životni ciklus rudnika i bio primenljiv na različite grane rudarstva, tako da se omogući i lociranje rizika i predlog mera za njegovo smanjenje.

S tim u vezi, istražena su svetska iskustva u ovoj oblasti, specifičnosti sličnih sistema za srpski jezik i mogućnost primene u uspostavljenom modelu upravljanja rizikom u rudarstvu.

U cilju realizacije postavljenog zadatka u okviru definisane teme bilo je neophodno:

- proučavanje postojećih modela za procenu i upravljanje rizikom,
- proučavanje faktora rizika u rudarstvu,
- proučavanje metoda razvoja ontologija,
- proučavanje metoda analize teksta, pronalaženja i ekstrakcije informacija i znanja relevantnog za rizik,
- uspostavljanje sistema ekstrakcije informacija,
- proučavanje vrsta i struktura ontoloških modela i standarda srodnih domena,
- uspostavljanje sistema opisivanja koncepata relevantnih za oblast istraživanja,
- razvijanje ontološkog modela za upravljanje rizikom u rudarstvu,
- evaluacija rezultata na realnim podacima.

Rizici u rudarstvu analiziraju se na primeru dokumentacije i istraživanja sprovedenih na Rudarsko-geološkom fakultetu Univerziteta u Beogradu za potrebe različitih rudarskih kompanija.

Zaključci ovih istraživanja biće osnova za razvoj jedinstvene metodologije i integralnog modela sistema za upravljanje rizikom u rudarstvu zasnovanog na ontologijama, koji bi se mogao primeniti na rudnike u realnim uslovima i realnom vremenu.

1.3. Polazne hipoteze

Osnovna polazna hipoteza za istraživanje bila je ta da je moguće formalizovati informacije i znanje o rizicima u rudarstvu, te razviti model za upravljanje rizikom koji će unaprediti efikasnost upravljanja rizikom i pomoći menadžmentu rudnika u donošenju odluka o primeni mera za smanjenje uticaja rizika identifikovanih u rudniku. Pošlo se od hipoteze da se na ontologijama može zasnivati model za upravljanje rizikom u rudarstvu, na osnovu kog se dalje može razviti odgovarajuće softversko okruženje, koje će integrisati upravljanje ontologijama relevantnim za rizik i ekstrakciju informacija i znanja. Kroz projektovani model, razvijena je formalna ontologija, značajno dopunjen rudarski korpus i implementirani algoritmi ekstrakcije informacija za popunjavanje ontologije.

Osnovne pretpostavke bile su:

- ontologija rizika u rudarstvu, kao potpuno nova ontologija, treba da bude bazirana na postojećim referentnim ontologijama, primerima dobre prakse i zakonskoj regulativi,
- moguće je identifikovati osnovne grupe uticajnih faktora i njihove činioce, odnosno moguće je odrediti hijerarhijsku strukturu uticajnih faktora,
- za analizu teksta će se koristiti Unitex skup programa i elektronski morfološki rečnici za srpski, a po potrebi će se dopunjavati terminima iz oblasti upravljanja rizikom,
- korišćenjem analize teksta moguće je razviti model za ekstrakciju informacija, eksplicitnog i implicitnog znanja o riziku,
- korišćenjem semantičke analize moguće je izvršiti proveru formalizovanog znanja o rizicima u rudnicima,
- upotreba ontološki zasnovane semantičke analize i automatskog rezonovanja povećavaju efikasnost i robusnost sistema za upravljanje rizikom,

- upotreba ontološki zasnovane analize rizika u rudnicima povećava kvalitet i obezbeđuje adekvatniju promenljivost projektovanog rešenja,
- ontologija i sistem baziran na znanju služe i kao semantički alat za pretraživanje i dalju ekstrakciju informacija,
- moguće je softverski automatizovati ontološki zasnovanu analizu primenom sistema automatskog rezonovanja, što omogućuje efikasnost i pouzdanost rada institucija koje se bave rudarskom problematikom,
- razvijeni sistem može se koristiti u okviru nastavnog procesa, čime se povećava efikasnost nastave u oblasti upravljanja rizikom i informacionih sistema,
- razvijeno rešenje treba da bude bazirano na veb tehnologijama i da obuhvati planiranje upravljanja rizikom, identifikaciju rizika, kvalitativnu analizu, kvantitativnu analizu, planiranje odgovora na rizik i praćenje i kontrolu rizika.

Prikupljanje domenskog znanja iz oblasti upravljanja rizikom zasniva se na intervjuima, interaktivnom i eksplicitnom kodiranju znanja, ali i razvoju algoritama i procedura za automatsku ekstrakciju iz raspoložive dokumentacije, zakonske regulative i standarda, uz ekspertsku verifikaciju znanja, koja je neophodna za pouzdanost ovakvog sistema. Model se projektuje i prezentuje u jedinstvenom jeziku za modeliranje UML¹, potom se razvijaju formalne ontologije, kolekcije resursa i implementiraju algoritmi upravljanja rizikom.

1.4. Naučne metode istraživanja

Za rešavanje problematike u ovom radu se koriste:

- deskriptivna metoda koja uključuje opisivanje, prikupljanje i sređivanje podataka, komparativna i analitička istraživačka metoda koja podrazumeva upoređivanje, vrednovanje i interpretaciju dobijenih rezultata i analizu podataka iz istraživanja drugih autora, kao i metode statističke obrade podataka,
- predloženi razvoj dinamičkog kreiranja ontologije se zasniva na ekstrakciji važnih koncepata rizika, računanjem relevantnosti dokumenata na osnovu vektora ključnih reči i relacija, kombinovanjem vektorskog modela i konačnih automata za istraživanje neizvesnosti, integrisanje neizvesnosti i analizu relacija,
- za postizanje postavljenog cilja, primenjene su kvalitativne i kvantitativne metode za razvoj i upravljanje ontologijama, kao i metode za ekstrakciju znanja. Takođe se koriste savremena softverska razvojna okruženja i platforme, kao i odgovarajući softverski alati i resursi za modeliranje znanja i ekstrakciju podataka iz nestruktuiranog teksta, uz obezbeđivanje visokog kvaliteta izlaznih informacija.

Primenjena metodologija uključuje nekoliko koraka u cilju izgradnje jednog ovakvog sistema: identifikaciju koncepata, instanci koncepata, zatim relacija između koncepata i između instanci. U okviru ovog rada korišćeni su alati za obradu prirodnog jezika i ekspertiska domenska znanja stručnjaka iz oblasti rudarstva i upravljanja rizikom u rudarstvu.

Osnovni koncepti ontologije, hijerarhijska struktura i relacije između koncepata postavljene su uz podršku eksperata, a u skladu sa postojećom metodologijom i standardima u oblasti upravljanja rizikom u rudarstvu. Za pronalaženje relevantnih podataka, instanci koncepata i veza između instanci, korišćene su plitke (eng. shallow) tehnike za obradu prirodnog jezika (NLP): metoda konačnih automata (Gross 1987) i upitni jezik CQL (eng. Corpus Query Language) zasnovan na podudaranju obrazaca u sistemu za upravljanje velikim količinama tekstualnih podataka CQP (eng. Corpus Query Processor) (Evert 2005). Tehnikama obrade prirodnog jezika su ekstrahovani kandidati za sadržaj ontologije, pronalaženjem veza između

¹ UML (engl. Unified Modeling Language) je standardni jezik za vizuelno prikazivanje objektnog modela

imeničkih fraza iz domena rudarstva i njihovih kolokacija. Takođe, metodologija korišćena u radu može doprineti i u proširenju domenske ontologije novim konceptima, ekstrahovanim iz teksta. Podatke dobijene kao rezultat plitkog parsiranja evaluirali su eksperti iz oblasti upravljanja rizikom u rudarstvu.

Istraživanje sprovedeno u okviru ove doktorske disertacije se bavilo i interfejsom između sintakse i semantike različitih rudarskih tekstova, u kontekstu specifičnog rudarskog domena, te su korišćeni specifični domenski obrasci za semantički prikaz sintaksnih struktura. Sistem koristi alate i resurse opisane u odeljku 3, unapređene novim sadržajima, resursima i alatima opisanim u poglavlju 4 i evaluiranim u poglavlju 5. Semantička reprezentacija, razvijena iz sintaksnih struktura i izvora specifičnih za rudarski domen (terminološko - semantički leksikon), unapređena je domenskom ontologijom. Za pronalaženje novih koncepata i ontološkog sadržaja korišćen je korpus koji sadrži različite tipove teksta iz rudarskog domena.

U okviru softverskog rešenja razvijen je modul za ekstrakciju relacija iz rudarskog korpusa, a kao referentni korpus korišćen je SrpKor, koji su razvili D. Vitas i grupa saradnika sa Univerziteta u Beogradu, <http://vvv.korpus.matf.bg.ac.rs/korpus/> (Vitas and Krstev 2012; Utvić 2014). Referentni korpus je korišćen zajedno sa plitkim parsiranjem i leksikonom za kreiranje koncepata iz domena rudarstva i rizika. Korpus koriste domenski eksperti za projektovanje inicijalne ontologije i evaluaciju automatski popunjenih instanci. Parser je korišćen za identifikovanje potencijalnih instanci koncepata između jednostavnih imeničkih fraza i glagola. Relacije između različitih klasa su identifikovane interpretacijom sintaksnih struktura kombinovanjem sa statističkim informacijama.

1.5. Struktura disertacije sa kratkim pregledom poglavlja

Doktorsku disertaciju, pored opštih elemenata, koje čine apstrakt na srpskom i engleskom jeziku, ključne reči, predgovor, sadržaj, spisak skraćenica, slika, tabela i literature, čine i sledeće strukturne celine:

1. Uvod
2. Pregled dosadašnjih istraživanja
3. Metodologija i resursi za razvoj ontologije
4. Model sistema za upravljanje rizikom
5. Evaluacija rezultata na primerima iz prakse
6. Zaključak i preporuke za dalji rad

Prvo poglavlje predstavlja uvod i u njemu je dat osvrt na predmet i cilj istraživanja, opisane su polazne hipoteze, kao i naučne metode koje su korišćene u istraživanju.

Drugo poglavlje daje pregled dosadašnjih istraživanja i literature koja je korišćena u ovom radu. Pregled literature je grupisan u nekoliko odeljaka, dat je osvrt na rizike u rudarstvu, upravljanje rizikom, zatim su prikazani softveri za upravljanje rizikom, nakon čega sledi pregled upravljanja rizikom zasnovanog na ontologiji i ekstrakciji koncepata i njihovih relacija.

Treće poglavlje opisuje korišćene metode i resurse za razvoj ontologije rizika. Opisan je pregled dostupnih izvora, rečnika iz domena rudarstva, zatim korišćenih korpusa tekstova, morfoloških i drugih rečnika. Prezentovana je obrada teksta na prirodnom jeziku alatima Unitex i Leximir i semantičke mreže WordNet i FrameNet. Prikazane su osnove teorije ontologije i baza znanja, principi razvoja i automatskog popunjavanja podataka, praćeni konkretnim primerima.

Četvrto poglavlje je centralno i opisuje srž ontološkog modela sistema za upravljanje rizikom.

Najpre je prikazana arhitektura sistema, sa svim korišćenim komponentama. Zatim sledi priprema resursa, koja uključuje digitalizaciju papirnih rečnika, proširenje korpusa, dodavanje domenskih pojmova morfološkim elektronskim rečnicima opšte namene. Sledeći korak je proces ekstrakcije rudarskih entiteta i sastavljanje terminološke baze znanja. Za inicijalni razvoj ontologije korišćen je UML, nakon čega je data detaljna specifikacija implementacije ontologije i demonstriran prikaz u alatu VocBench.

Peto poglavlje bavi se evaluacijom ekstrakcije informacija, nakon čega slede primeri na konkretnim primerima upravljanja rizikom u rudarstvu.

Poslednje poglavlje nudi diskusiju, završne napomene i obrise budućih planova za unapređenje sistema i primenu u drugim oblastima.

2. PREGLED DOSADAŠNJIH ISTRAŽIVANJA

Najuticajnije podjelu znanja dao je (Polanyi 1964) filozof koji se bavi individualnim znanjem. Prema Polanijevoj tipologiji, postoje dve vrste znanja: implicitno i eksplicitno znanje. Po Tacitu (prećutno) znanje je znanje koje postoji u umu pojedinaca, teško ga je formalizovati i deliti sa drugima putem pisanih uputstava ili jednostavnim saopštavanjem. To je “unutrašnje” znanje koje se ponekad uzima zdravo za gotovo (i može se činiti krajnje jednostavnim), ali ga je teško opisati. Jezik se, na primer, ne može savladati pukim iščitavanjem rečnika i učenjem gramatike, zanat se teško može preneti samo objašnjavanjem.

Za razliku od prećutnog znanja, eksplicitno znanje se definiše kao znanje, koje se može jasno kodifikovati i formalizovati, i zbog toga se lako prenosi i deli putem različitih mehanizama komunikacije i čuva na nekom medijumu.

2.1. Rizici u rudarstvu

Postoje dve vrste pristupa za identifikaciju opasnosti: (i) neformalni pristup zasnovan na prethodnim podacima i istoriji (ii) formalni pristup zasnovan na tehnikama identifikacije opasnosti (Henley and Kumamoto 1996).

Budući da se rudarstvo smatra za visokorizičnu delatnost, za procenu rudnog odnosno mineralnog bogatstva potrebno je izvršiti uspešno istraživanje na tri koloseka: eksploatacija, ocena mineralnih sirovina i procena mineralnih rezervi. Tipičan životni ciklus rudnika mora biti predvidiv da bi se utvrdila održivost projekta i da bi bila moguća procena rizika projekta. (Rendu 2017).

Ključne komponente koje utiču na izazove u proceni rizika u rudarskoj industriji, i čine problematiku posebno kompleksnom su:

- proizvodni procesi uključuju puno povezanih delova čija analiza zahteva širok spektar stručnjaka,
- postoji visok nivo neizvesnosti vezan za trenutne i buduće vrednosti ulaznih promenljivih,
- uslovi i rizični faktori, koji nisu predvidljivi, menjaju se tokom života rudnika.

Glavni rizici sa kojima se suočava rudarstvo i metalska industrija u 2019 i 2020 godini ukazuju na novu eru promena kako u okviru sektora tako i izvan njega. Rudarska industrija se sve više suočava sa problemima kao što su dozvole za rad, složena zakonska regulativa, društvena očekivanja, digitalna transformacija i izazovi u vezi sa odlukama o poslovanju i kapitalnim ulaganjima.

Skala najrizičnijih aspekata poslovanja se značajno menja poslednjih godina, što ukazuje na to da su promene na globalnom planu već prisutne. Društvene promene, nove tehnologije i trka za transformacijom poslovnih modela pokreću čitav niz izazova za rudarske i metalske kompanije. Pritisak tehnoloških i automobilskih kompanija u trci da osiguraju snabdevanje tržišta svojim proizvodima otvara nove izazove, a istraživanja pokazuju da tehnološke kompanije imaju potencijal da značajno utiču na rudarski sektor, da imaju pristup kapitalu i ulažu u inovacije i tehnologije koje zahtevaju efikasnije rudarske operacije.

Prvih deset poslovnih rizika s kojima se suočavaju rudarstvo i metalska industrija danas u svetu su²:

1. Licence za rad se u istraživanju poslovnih rizika nalaze se na vrhu liste poslednjih nekoliko godina. Rudarski sektor se nalazi sve više pod pritiskom javnosti koja zahteva sve veću etičku transparentnost i niže stope emitovanja štetnih materija, kao i pod pritiskom akcionara, naročito rudnika uglja, koji zahtevaju preoblikovanje portfolija rekonfigurisanjem postojećih operacija ili čak i zatvaranje rudnika. Dosadašnji pristup ovoj oblasti mora biti prevaziđen u svetlu sve većeg broja dostupnih informacija i većeg broja zainteresovanih strana. Kompanije moraju ugraditi u svoj poslovni sistem potrebne procedure i politike kako bi uzele u obzir potrebe društvene zajednice, države, zaposlenih i životne sredine u okviru i izvan života rudnika.
2. Budućnost ljudskih resursa u svetlu digitalnih i tehnoloških inovacija postaje iz godine u godinu sve uticajni faktor. Inovacije svakako donose povećanje produktivnosti, konkurentnosti i sigurnosti, ali sa druge strane nose i neizvesnost u pogledu znanja i veština koje će u svetlu novih tehnologija biti potrebne. Migracija ljudskih resursa često je ogledalo kretanja tržišnih uslova, tako da u periodima kriza mnogi radnici odlaze u druge sektore i kasnije se ne vraćaju u rudnike. Rudarske kompanije moraju imati strategiju da privuku ljude sa potrebnim kompetencijama za nove načine rada, a paralelno sa tim da rade na prekvalifikaciji zaposlenih obučavajući ih za sticanje novih znanja i veština koje će u budućnosti biti neophodne.
3. Digitalizacija i optimizacija podataka, uključujući i digitalnu efikasnost su ključni za postizanje konkurentne prednosti. Istraživanja pokazuju da mali broj menadžera ima znanja iz oblasti digitalnih tehnologija i da je digitalizacija u rudarstvu na niskom nivou, svodeći se na pojedinačna rešenja. Tek sa primenom softverskih rešenja u celom lancu vrednosti mogu se stvoriti digitalni rudnici koji će doneti transformaciju i novu dominaciju na tržištu. Poseban naglasak je na tome kako što bolje upravljati podacima da bi se iz njih izvukla maksimalna vrednost.
4. Redukcija emisije štetnih gasova je jedan od pritisaka na rudarstvo pri čemu se sa jedne strane zahteva smanjenje emisije štetnih gasova, a sa druge strane potrošnja energije je u stalnom porastu. Tranzicija na čiste energije pruža ogromne mogućnosti za rudarske kompanije koje se bave eksploatacijom litijuma, kobalta, bakra, aluminijuma, nikla i drugih minerala.
5. Rizici sa visokim faktorima uticaja zahtevaju naročitu pažnju, tako da većina rudarskih kompanija jasno iskazuje i upravlja kritičnim rizicima u svom poslovanju. Međutim, ovi kritični rizici često su statični i ostaju u gotovo istom formatu godinama. Kompanije bi trebalo da vrše procenu šta bi se dogodilo usled nekih većih tehnoloških promena, npr. porast upotrebe plastike i drugih naprednih materijala.
6. Maksimizacija povraćaja uložених sredstava predstavlja dugoročni cilj da se poveća povraćaj kapitala uz istovremeno smanjenje nestabilnosti dobiti. Rudari trebaju da usvoje nove pristupe rizicima u odnosu na tradicionalne kada je u pitanju alokacija kapitala, kako bi izbegli propuštanje novih prilika i mogućnosti. Neke od oblasti razmatranja su promene u načinu finansiranja, preispitivanje poslovanja i zajedničke investicije.

² https://www.ey.com/en_gl/mining-metals/10-business-risks-facing-mining-and-metals#, pristupljeno avgusta 2020.

7. Informacione tehnologije u oblaku i deljeni podaci su ključni za povećanje produktivnosti i minimiziranje troškova jer omogućavaju automatizaciju odluka uz minimalizaciju gubitaka u celom lancu vrednosti. Međutim, paralelno sa većom digitalizacijom povećava se i rizik od sajber napada kroz fizičku imovinu, digitalnu infrastrukturu i poslovne procese. Kompanije treba da promovišu mere sigurnosti i kulturu odgovornosti kroz kontinuirano unapređivanje.
8. Izmena proizvodnje može dovesti, u narednih 10–15 godina, do značajne krize snabdevanja. Manji budžeti za istraživanje, manji broj otkrivenih nalazišta i manje procene postojećih ležišta su zabrinjavajuće kada se ima u vidu rast potražnje mineralnih resursa kao rezultat globalnog rasta. Strategije za rast snabdevanja rudnika ogledaju se kroz pristup kapitalu, finansiranje postojećih projekata ili novih rudnika, jačanje investicija za nalaženje novih resursa i izvlačenje veće dobiti iz postojećih resursa.
9. Inovacije i porast produktivnosti mogu se postići preispitivanjem postojećeg načina rada i spremnošću za unapređenje. Inovacije mogu biti ključne za rešavanje bitnih strukturnih pitanja u rudarstvu, kao što su: padajuća cena rude, porast eksploatacije ruda na udaljenim i nepristupačnim lokacijama, pristup i troškovi energije i infrastrukture, rast kompleksnosti operative, poboljšanje upravljanja vodama.
10. Rastući troškovi se moraju strogo kontrolisati, pri čemu smanjenje troškova treba držati pod kontrolom, kroz inovacije i partnerstva, ugovore sa dobavljačima koji na duži rok mogu smanjiti troškove, strateška zajednička ulaganja.

Kao stalna dilema postavlja se pitanje kako smanjiti troškove i ostati konkurentan. Troškovi u rudarskom sektoru su visoko podložni inflatornim pritiscima, pri čemu se tokom perioda većih cena troškovi kao što su zarade, potrošni materijal, goriva i energija, često povećavaju većom stopom od opšte inflacije. Dodatno, inkrementalne promene u načinu rada rudnika rezultiraju rastućim troškovima, što uključuje sve veću kompleksnost rudnika, porast upotrebe tehnologije, promene u radnoj snazi i sve veće investicije u dozvole za rad. Da bi ovo nadoknadile rudarske kompanije moraju: da se fokusiraju na programe održivog smanjenja troškova; da preispitaju ulaganja kapitala u unapređenja razvoja i zaliha; uvedu automatizaciju u različitim segmentima poslovanja ; sprovode blagovremen otkup zemljišta.

Troškovi energije predstavljaju i do 30% ukupnih troškova kompanije, a pored troškova, na izbor energetske izvora utiču i društveni i reputacijski razlozi, održivost i dostupnost izvora energije. Da bi smanjile ove rizike, kompanije se odlučuju za kombinaciju izvora energije - fosilna goriva, hidroelektrane i obnovljivi izvori energije, zamenu opreme na dizelski pogon električnom. Krajnje rešenje leži u razvoju niza novih visoko-efikasnih energetske sistema za rudarske operacije, koji ih čine zelenijim, jeftinijim i održivijim.

Izazovi koji se nameću mogu biti nepoželjni, ali ujedno mogu predstavljati i prilike za transformaciju. Promene postoje unutar rudarskog sektora, npr. automatizacija izaziva promenu poslovnih procesa, ali dolaze i van sektora, npr. tehnološke kompanije mogu postati investitori sa namerom da skrate snabdevanje mineralima, naročito kada su u pitanju kobalt i litijum, takođe, države mogu sa sopstvenim kapitalom da postanu glavni akcionari radi obezbeđivanja snabdevanja za nacionalnu industriju i zaštitu radnih mesta, trgovci takođe traže mogućnosti za što bolje pozicioniranje u snabdevanju ključnim proizvodima. Odgovor na ove rizike moguće je naći u investiranju u digitalne tehnologije i inovacije, ulaganje u nove mogućnosti i jačanje veza sa svim zainteresovanim stranama.

Prevare i korupcija se takođe navode kao jedan od rizika, tako da mora postojati jača kontrola poslovanja, izvođača i dobavljača. Prevare postaju sve sofisticiranije sa rastućom međusobnom povezanošću, pri čemu mediji imaju značajnu ulogu u njihovom razotkrivanju. Rizik od prevare je povezan sa rizikom za ugled kompanije i dozvolama za rad.

Proizvodi novog doba, poput baterija, pametnih telefona, laptopova pa sve do naprednih odbrambenih sistema dominiraju sve više u oblasti visokih i zelenih tehnologija. Rudarstvo mora usvojiti veći nivo fleksibilnosti u svojim poslovnim modelima kako bi bili prilagodljivi za promene u tražnji, a odluke o tome gde ulagati ili plasirati kapital se moraju donositi mnogo unapred.

Badri predlaže analitički pristup za integraciji bezbednosti i zdravlja na radu u upravljanju projektnim rizicima u industriji, gde bi se uključili istraživači i zajedno sa ljudima iz prakse kreirali novi pristup koji bi se sastojao od dva okvira. Prvi okvir je teorijski i na njemu bi radili istraživači, a uključuje definisanje osnovnih načela, istraživanje literature, opis izazova i izrada konceptualnog modela. Drugi, praktični okvir na kome rade ljudi iz prakse, bi prema kreiranom konceptualnom modelu sadrži prikupljanje i analizu podataka, primenu modela, opis rezultata i promena. Zaključci i preporuke iz drugog okvira bili bi ulaz za prvi okvir i smernice na osnovu kojih bi istraživači radili na unapređenju teorijskog konceptualnog modela. Primena pristupa dala je merljive rezultate u dvema rudarskim kompanijama u Kvebeku (Kanada) (Badri, Gbodossou, and Nadeau 2012)

Događaj koji ima pridruženu verovatnoću da će prouzrokovati štetu, povredu, odgovornost, gubitak ili bilo koji drugi negativan uticaj naziva se rizikom (Dasgupta et al. 2016; Slywotzky and Drzik 2005; Beasley, Clune, and Hermanson 2005; Lu et al. 2009). Organizacije su uvek u potrazi za informacijama vezanim za takve događaje izazvane unutrašnjom i spoljašnjom ranjivošću, kako bi se mogli izbeći mogući negativni uticaji preventivnim delovanjem. Izvori rizika mogu biti mnogi. Poteškoća u prepoznavanju rizika proizilazi iz raznolikosti izvora. Rizici mogu nastati iz neizvesnosti na finansijskim tržištima (Leidner and Schilder 2010; Ykhlef and Algawiaz 2014), industrijskih procesa ili zbog neuspeha projekta. Neočekivani događaji poput prirodnih katastrofa, zakonodavnih akata, namernih napada ili određenih poteza od strane konkurenta mogu dovesti do situacije koja može uticati na organizaciju i stoga se može nazvati rizikom. Uopšteno, rizik ima sledeće karakteristike: tip rizika koji ukazuje na prirodu neželjenog potencijala, uzrok rizika koji predstavlja događaj koji može izazvati specificirani rizik i uticaj (posledicu) koji se bavi težinom nastale štete.

Tuffano (1996) se fokusira na finansijski rizik i uspeh na berzi, sugeriše da praksa upravljanja rizikom u rudarstvu treba da se oslanja na uspostavljanje balansa između učinka kompanije i prioriteta menadžmenta. Menadžeri koji imaju na raspolaganju više opcija upravljaju sa manjim rizikom, oni koji poseduju više deonica u preduzeću upravljaju sa većim rizikom, menadžeri koji su novi u svom poslu takođe preferiraju veće rizike. Kompanije sa nižim finansijskim mogućnostima sklone su većim rizikovanjem sa cenama, kompanije u kojima su spoljni akcionari većinski vlasnici teže ka manjim rizicima (Tuffano 1996). To ukazuje da ne postoji čvrsta veza između upravljanja rizikom i karaktera firme. Pokazuje se da su, kada je dobit menadžera povezana sa upravljanjem rizikom, rizici umereniji na korporativnom nivou. To je naročito bitno kada je upravljanje rizikom skupo i tada se mora osigurati da su korporativni resursi namenjeni maksimiziranju vrednosti, a ne smanjenju rizika od strane menadžera. Mnogi autori tvrde da povezivanjem zarada menadžmenta sa vrednošću akcija, kompanije mogu dobro da usklade stimulanse menadžera sa interesima spoljnih akcionara.

Dowd, P.A. (1994) istražuje procenu rizika u oceni rezervi i planiranju površinskih kopova navodeći da analiza rizika u rudarskim poduhvatima treba da uključi efekte neizvesnosti na sve relevantne promenljive. U praksi, rezerve rude su obično najznačajniji faktor. Neizvesnost o rudnim rezervama ne mogu se razmatrati u analizama na tradicionalan način specificiranjem histograma ili raspodelom verovatnoće vrednosti zato što pokazatelji rezervi (kvalitet i količine) su u funkciji geološke na kojoj su izmereni. Geostatistička simulacija nudi metodu predstavljanja nesigurnosti u pogledu kvaliteta i količine u analizi rizika na način koji uzima u obzir trodimenzionalnu prirodu njihove pojave (Dowd 1994).

Steffen (1997) se bavi metodologijom uvođenja rizika u izradu rudničkih planova za površinske kopove. Preporučuje postizanje balansiranja rizika između procene mineralnih resursa i projektovanja kosina. Proces projektovanja identifikuje kriterijume za različite faze planiranja, tako da je poslovni rizik odvojen od geološkog, odnosno geotehničkog rizika, pa se predlažu odgovarajuće procedure u okviru izrade rudničkih planova. Autor zaključuje da formalni proces projektovanja poboljšava komunikaciju između ključnih donosioca odluka, tako da daje dodatnu vrednost akcionarima i investiranju (Steffen 1997).

Integracija tri nivoa planiranja od životnog veka rudnika, preko dugoročnog do kratkoročnog planiranja pruža robusnu proceduru planiranja, koja pruža dobru komunikaciju između zaposlenih na svim nivoima, od inženjera planiranja preko menadžmenta do odeljenja za marketing. Kriterijumi smanjenja troškova za planiranje su praktični, i predstavljaju mnogo korisnije sredstvo od procene cene ili pristupa smanjenja kvaliteta; tako da se koncentrišu napori i pospešuju veštine tima za planiranje i fokusiranje na ona pitanja kojima mogu najviše doprineti poboljšanju planiranja rada rudnika.

Primena dugoročnog planiranja zasnovanog na troškovima znači da se rizici mogu pravilno proceniti u smislu perioda poslovnog rizika i najboljeg rešenja dobijenog kao balans između dobiti investicije i rizika poslovanja. Krajnji rezultat procesa planiranja je maksimalno iskorišćena rudna rezerva za trošak proizvodnje. Razlika između predložene procedure i onoga što se uobičajeno koristi je da je proces planiranja vođen konkurentnošću rudnih rezervi na svetskom tržištu. Menadžment aktivno učestvuje u donošenju odluka i definisanju kriterijuma u planiranju. Oni se odnose na očekivane stope prihoda za akcionare, nivo rizika koji su prihvatljivi za kompaniju, fleksibilnost potrebnu u radu i prihvatljive troškove poslovanja. Treba doneti odluke o fundamentalnim parametrima, kao što je period rizika definisan kao period poslovnog rizika, isto kao i veličina odstupanja. Iskustvo je pokazalo da je unapređenje u komunikaciji i razumevanju donošenja odluka između menadžmenta i inženjera planiranja značajno doprinelo rastu vrednosti akcija.

Godoy i Dimitrakopoulos (2004) bave se dugoročnim planiranjem proizvodnje u rudnicima i pronalaženjem optimalnog plana dinamike otkopavanja rude i otpada kroz sve faze u životnom veku jednog rudnika. U svom pristupu uključuju upravljanje rizicima koji su povezani sa neizvesnošću prostiranja rudnog tela i kvaliteta rude. Proces optimizacije uključuje nekoliko faza i bazira se na integraciji više jednako mogućih reprezentacija rudnog tela, ekonomskoj analizi, specifičnostima rudnika i proizvodne opreme. Kvantifikovana procena neizvesnosti u vezi sa kvalitetom rudnog tela i tempom eksploatacije rude, otkopavanja i odlaganja otpada, dovodi do toga da proizvodnja u svim fazama života rudnika prolazi kroz zone u kojima se rizik od ne dostizanja ciljeva svodi na minimum, u kom su planovi otporni na rizike i u kojim je povećana neto vrednost projekta u odnosu na tradicionalne pristupe dugoročnom planiranju. (Godoy and Dimitrakopoulos 2004)

Tripathy i Ala (2018) ističu identifikaciju opasnosti kao ključnu fazu u procesu procene rizika, nakon koje treba sprovesti analizu i procenu rizika u cilju povećanja bezbednosti na radu. Na osnovu više od 7000 izveštaja o nesrećama u podzemnim rudnicima uglja u Indiji identifikovali su 172 opasnosti. Opasnosti su prvo kategorizovane u šest grupa: geomehaničke (kretanje stenskih masa), mehaničke (užad, sistem transportnih traka, damperi), hemijske (miniranje i eksplozija), električne (struja), geohemijske (prašina, gas i drugi zapaljivi materijali) i ekološke (poplave, zagađenje voda). Zatim je izvršena dodatna klasifikacija prema faktoru na: ljudski, mašina/alat, metoda rada/proceduralni, i radno okruženje/upravljački. Za svaki opisan događaj navedeni su rizici koje oni nose i predložene mere za uklanjanje i kontrolu u vidu inženjeringa, administrativne, lične zaštitne opreme ili ponašanja zaposlenih. Opsežna lista data u ovom radu ograničena je samo na opasnosti po bezbednost, pa autori preporučuju za buduća istraživanja identifikaciju opasnosti vezane za zdravlje, zaštitu životne sredine, složenost operativnog okruženja, organizaciju i ljudske resurse. Autori daju sistematizaciju opasnosti, u kojoj navode tip, opasnost, rizik, kao i metode za identifikaciju, koje su potom podeljene na neformalne i formalne (Tripathy and Ala 2018).

Olofsson (2020) kombinuje jedan novi pristup baziran na „zamišljenoj budućnosti“ (eng. „imagined futures“) i budućem radu sa konceptualnim tehnikama za pronalaženje i procenu ležišta i njegovu budućnost. Praćenjem procesa istraživanja od početka do kraja, arhivski i novi podaci opisuju liniju budućih radova u kojoj se sve veći broj stručnjaka i profesionalaca okuplja kako bi projektovao buduće stanje koje postaje sve bogatije i složenije kako proces napreduje. Počevši od geološke studije za smanjenje neizvesnosti procene ležišta, veličine i geometrije ležišta, procena kvaliteta mineralne sirovine se zasniva na podacima projektovanim za budućnost, odnosno za trenutak u kom će da se eksploatišu (Olofsson 2020).

Istraživači se oslanjaju na podatke, teoriju (npr. geologiju, geohemiju, geofiziku i ekonomiju), iskustvo i metodologiju modeliranja kako bi kreirali verne slike budućnosti rudnika. Upotreba industrijskih standarda daje oblik trajnosti predviđanjima datim u procesu istraživanja. Međutim, iako se rizik može definisati i smanjiti kroz prikupljanje podataka, analizu i modeliranje, svaka zamišljena budućnost ili virtuelna realnost još uvek je relativno nesigurna (Wynne 1988).

Sagledavanjem procesa istraživanja u celini, otkriva se da je ono što je započelo kao materijalni predmet mineralna anomalija u zemlji je definisana i redefinisana u četiri glavne faze. Prva faza podrazumeva konceptualne studije projektnih ideja za identifikaciju potencijalnih lokacija na osnovu njihovih geoloških karakteristika. Zatim se istražuje vrednost potencijalne lokacije primenom niza geoloških, geofizičkih i geohemijskih metoda. Lokacija se zatim redefiniše kao mineralni resurs sa ekonomskim potencijalom. U završnoj fazi, mineralni resurs se prevodi u standardizovani entitet - rezerve, koji vrednovanje kvaliteta definiše iz ugla ekonomske, geološke, tehničke, ekološke, pravne i društvene perspektive, obezbeđujući da će se uspešno eksploatisati.

Wynne (1988) uvodi termin „neukrotiva tehnologija“ (eng. unruly technology) i navodi diskusiju: praktična pravila, nepraktični diskusi i razumevanje javnosti. Konačno, ontološka neizvesnost buduće eksploatacije može da proizvede nepredviđene okolnosti u kojima će se predviđeno buduće stanje pokazati kao manje trajno nego se mislio. Dalje, ako nisu svi rizici redukovani na planirani način i neizvesnost može ostati uprkos naporima da se prikupe informacije koje mogu osigurati ekonomično eksploatisanje resursa. Spremnost da se odustane od projekta, smatrajući da je ležište minerala nerentabilno, važan je deo upravljanja

materijalnim i vremenskim izazovima koji se odnose na istraživanje mineralnih resursa. Socijalna istraživanja o uticaju rudarstva i istraživanja mineralnih resursa dugo su se fokusirala na tehnološki aspekt rudarstva, ali se pristup sada menja i sve veći uticaj imaju društveni, ekološki faktori, i procena uticaja na životnu sredinu.

Miljanović i kolege navode sledeće kategorija rizika: 1) rizici povezani sa ljudskim resursima (zdravlje i bezbednost na radu, povrede), 2) gubici postrojenja (incidenti, prirodne katastrofe), 3) rizici po životnu sredinu (zagađenje vode, vazduha, kontaminacija tla, opasan otpad), 4) rizici po principu odgovornosti (greške, pravna pitanja, manjkavost usluga i proizvoda, ...), 5) rizici poslovnih procesa (prekid rada opreme, gubici, porast troškova,...), 6) Projektni rizici (probijanje rokova i budžeta, politički rizici, problem u finansiranju, treća lica...) (Miljanović, Milutinović, and Tokalić 2016). Postoje različite klasifikacije faktora rizika, u zavisnosti od aspekta posmatranja: interni ili eksterni; potom tehnički, organizacioni ili pravni,... U eksterne faktore spadaju: ekonomski i finansijski, tržišni uslovi, zakonska regulativa, eksproprijacija, geopolitički, socijalni, regionalno partnerstvo i prirodni faktori (Milutinović, Miljanović, and Panić 2013).

Nisu retki slučajevi kada nije jednostavno razgraničiti kategorije pojedinih rizika: kod upravljanja tehnološkim sistemima, otkaz sistema može biti prouzrokovan otkazom hardvera, softvera, potom organizacioni problem ili ljudski faktor, ali može biti prouzrokovan i prirodnim dešavanjima ili drugim eksternim incidentima. Neki od čestih problema su: ignorisanje ili previd defekta, prekid komunikacije, prekid signalizacije, konflikti zaposlenih i menadžmenta, nesprovođenje odluka, problem nabavke rezervnih delova, potom geološki: rasedi, zone sa manjim kvalitetom mineralne sirovine,... (Kirin et al. 2020). Ljudski faktor postaje sve aktuelnija tema u studijama o riziku i sigurnosti, gde se prikupljaju i analiziraju informacije o ljudskim sposobnostima, njihovim ograničenjima i drugim karakteristikama u kontekstu posla koji obavljaju, njihove međusobne interakcije i interakcije sa mašinama, sistemima i okruženjem u cilju da se realizuje bezbedan radni proces (Kirin et al. 2015).

Rizici se mogu klasifikovati i u odnosu na subjekte: 1) hardver, u koji spadaju: oprema, cevovodi, instrumenti, konstrukcije, materijali, ali i projekti; 2) softver, pod kojim se podrazumevaju: procedure, instrukcije, specifikacije; 3) ljudski resursi, za koje se vezuju: upravljanje, operateri, održavanje; 4) spoljašnji: usluge (recimo para, električna energija), prirodni faktori (kiša, mraz), sabotaža.

Uključivanje ekspertskog znanja odnosno iskustvom stečenog znanja je od naročitog značaja, recimo rizik otkaza sistema povećava se ukoliko se znanje izvan opsega posmatrane discipline ne uključi u analizu.

Plan sprovođenja procesa procene rizika u rudniku, između ostalog, uključuje organizaciju i koordinaciju sprovođenja procesa procene rizika, kreiranje tima stručnjaka za procenu rizika i definisanje termin plana, načine prikupljanja dokumentacije neophodne za procenu rizika, načine prikupljanja informacija za procenu rizika od zaposlenih (ankete na radnom mestu), konsultovanje sa predstavnicima zaposlenih i informisanje o rezultatima procene rizika i preduzetim merama, druge aktivnosti potrebne za sprovođenje, izmenu i dodavanje u proces procene rizika (Kirin et al. 2013). Tokom identifikacije opasnosti i štete uzimaju se u obzir: rutinske i nerutinske aktivnosti celokupnog osoblja koje ima pristup radnom mestu (uključujući pregovarače i posetioce), ljudski faktor (ponašanje, sposobnosti i drugi faktori), rad opreme, identifikovanih opasnosti i štetnosti koje potiču izvan radnog mesta, a mogu imati nepovoljan uticaj na zdravlje i sigurnost na radu, opasnosti i štete koje se javljaju u blizini

radnog mesta usled radnih aktivnosti ili materijala, promena sistema upravljanja, uključujući privremene, i njihov uticaj na operacije / aktivnosti, procese, proizvode / usluge i radno mesto, obaveze u vezi sa procenom rizika i primenom neophodnog upravljanja, podaci o prethodnim incidentima, nesrećama i vanrednim situacijama u vezi sa operacijama / aktivnostima, procesima, proizvodi / usluge i radno mesto, uključujući njihovo prilagođavanje ljudskim sposobnostima.

Kao prvi korak u identifikaciji hazarda, odnosno opasnosti vrši se identifikacija svih procesa u rudniku za koje se pretpostavlja da imaju visok nivo rizika. Vrednovanje samog nivoa rizika se vrši tako što se za specifične korake u procesu rada primenjuju sledeći kriterijumi: verovatnoća pojavljivanja, kriterijum zdravlja i bezbednosti na radu, poštovanje zakona i propisa, kriterijum odnosa sa okruženjem, kriterijum izloženosti riziku (Kirin et al. 2015). Za rudarsku proizvodnju su karakteristični sledeći hazardi: prisustvo visokog pritiska ili temperature u sistemu, čin pušenja u određenim zonama, eksplozivna svojstva materijala, neodgovarajuće ponašanje zaposlenih, skladištenje toksičnih supstanci, rudarski radovi, odlagališta i flotacija u blizini naseljenih oblasti, otpadne vode, sabotaza, loše održavanje, neobučen kadar, nedisciplinovan kadar, dotrajalost postrojenja, pogrešno obeležavanje instrumenata (ventila, tastera,...) i sl.

U praksi se pokazalo da prekid poslovnog procesa se preklapa sa drugim vrstama rizika: 1) ispadanje značajnog uređaja ili dela opreme iz operativnog toka; 2) gubitak vlasništva usled vatre ili eksplozije; 3) pitanja odgovornosti koja prouzrokuju privremeni prestanak rada; 4) rizik od nemogućnosti pokrivanja troškova; 5) smanjenje obima internih resursa; 6) značajan porast troškova.

Preduslov za mogućnost smanjenja rizika je poznavanje izvora nekih rizika koji se mogu pojaviti u poslovanju rudnika, što uključuje predvidive i nepredvidive rizike, ranije pomenute eksterne i interne rizike, rizike prema fazi nastanka, rizike kod uvođenja novog projekta/proizvoda. Prema stepenu uočljivosti rizici poslovanja mogu se podeliti na: predvidive rizike, delimično predvidive rizike i nepredvidive rizike.

Rizici eksternog karaktera dolaze spolja, iz tržišnog ambijenta u kome se odvija poslovanje rudnika, na primer: rizici od nastupa i načina rada konkurencije; rizici od nedostupnosti ili ne zainteresovanosti kupaca; rizik od neplanirane pojave novih tehnologija; rizik u poslovanju sa bankom, dobavljačima, distributerima, špediterima i drugim komitentima; rizici uzrokovani ponašanjem državnih organa; rizici proizašli zbog institucionalne regulative; rizici od nepovoljnih promena u ambijentu (cenovnih odnosa, valutnog kursa, stope inflacije i dr.); rizici od pojave novih trendova u potrošnji.

U rizike internog karaktera spadaju³: rizik od zastarevanja poslovnih i razvojnih faktora; rizik od nedostataka ideja i inovacija; rizik od neracionalnog i šablonskog rada; rizik od neodgovornog ponašanja osoblja; rizik od nezakonitog i nemoralnog ponašanja; rizik od krađa, provala, pronevera i dr.; rizik od požara, poplava i drugih elementarnih nepogoda; rizik od ekoloških akcidenata; rizici od kala, rastura, loma, kvarenja, demoliranja, gubljenje upotrebljivosti i dr.

Rizici prema fazi i mestu nastanka, mogu se razvrstati u dve grupe: rizici u izboru poslovne koncepcije i rizici u promeni poslovne koncepcije. Rizici u primeni poslovne koncepcije su: rizik

³ Dušan Stevanović, Upravljanje rizikom u preduzeću NIS
<http://docshare01.docshare.tips/files/24332/243323228.pdf>, pristupljeno 1.2.2021.

od pogrešne primene dobrih poslovnih strategija, rizik od pogrešnog operativnog menadžmenta, rizik od lošeg funkcionisanja kanala distribucije i dr.

Razvoj novog proizvoda izuzetno je rizičan i postoje tri aspekta posmatranja: tehnološki rizik, rizik performansi, ekonomski rizik.

Štete koje mogu nastati ostvarivanjem određenih rizika, mogu se ispoljiti kroz: gubitak imovine, gubitak zbog odgovornosti, gubitak poslovnog ugleda i osvojene ekonomske i društvene pozicije, gubitak važnih saradnika i druge gubitke.

2.2. Upravljanje rizikom

Rizik se definiše kao efekat neizvesnosti na ciljeve, prema ISO 31000 standardu. Osnovni rečnik iz domena upravljanja rizikom ili menadžmenta rizikom⁴ kako definiše Institut za standardizaciju Srbije uvodi definicije osnovnih pojmova u sklopu dokumenta „Terminology and guidelines for risk prevention ISO CD31000“⁵ (Terminologija i smernice za sprečavanje rizika). Ovaj vodič pruža osnovni rečnik za razumevanje pojmova o upravljanju rizicima u različitim organizacijama i poslovnim okruženjima, može se koristiti i kao podloga različitim aplikacijama, a može se primeniti u svim sektorima poslovanja.

Vodič pruža definicije pojmova koji se odnose na upravljanje rizikom i ima za cilj standardizaciju jedinstvene terminologije upravljanja rizikom u procesima i okvirima koji se bave upravljanjem rizikom. Namenjen je onima koji se bave upravljanjem rizicima, onima koji su uključeni u aktivnosti ISO-a i IEC-a, i programerima nacionalnih ili sektorskih standarda, vodiča, postupaka i kodeksa prakse koji se odnose na upravljanje rizikom. Faze procesa upravljanja rizikom su:

- uspostavljanje konteksta organizacije,
- identifikacija rizika,
- analiza rizika,
- procena rizika,
- tretman rizika,
- praćenje i merenje,
- komunikacija i konsultacije, koja prožima sve prethodne faze.

Principi i smernice o upravljanju rizikom obrađuje Standard ISO 31000:2018(en) Risk management — Guidelines⁶. Upravljanje rizikom sastoji se od četiri faze: identifikacija, analiza, procena i tretiranje rizika. Osnovne smernice za upravljanje rizikom jesu pronalaženje odgovora na pitanja:

- Šta bi se moglo dogoditi, gde i kada?
- Zašto i kako se to može dogoditi?
- Kakve bi mogle biti posledice ako se to desi?
- Koje su kontrole postavljene da bi se poboljšala dobit i sprečile ili smanjile štetne posledice?
- Koliko su efikasne ove kontrole?
- Koliki je nivo rizika?

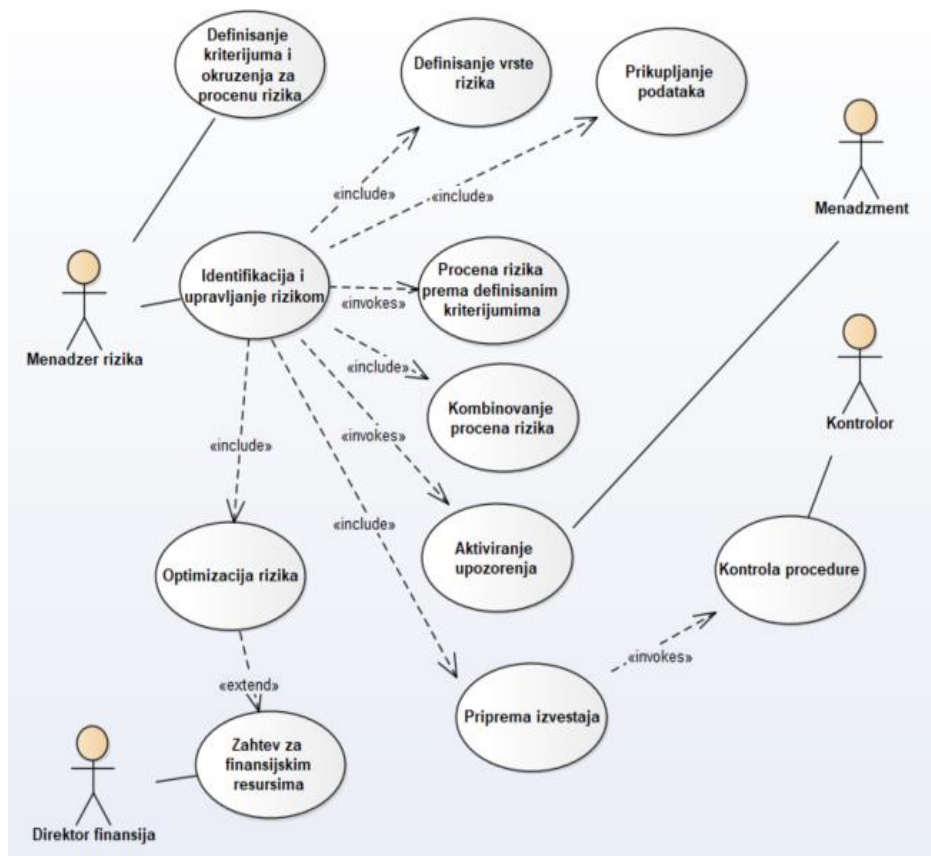
⁴ <https://iss.rs/sr/Cyrl/term/info/16489> menadžment rizikom (eng. risk management): koordinisane aktivnosti radi usmeravanja i kontrolisanja neke organizacije (2.57) u vezi sa rizikom

⁵ <https://www.iso.org/obp/ui/#iso:std:iso:guide:73:ed-1:v1:en>

⁶ <https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en>

- Kako najbolje tretiramo rizik?

Slika 2-1 prikazuje tipične aktivnosti u procesu identifikacije i upravljanja rizikom, a to su: identifikacija kroz prikupljanje podataka i definisanje vrste rizika, procena prema unapred definisanim kriterijumima, kombinovanje različitih kriterijuma kada se radi o senzitivnim rizicima, optimizacija rizika kroz procenu visine štete i koštanja prevencije, priprema izveštaja, aktiviranje upozorenja i primena kontrolnih mera.



Slika 2-1 UML model upravljanja rizikom

Može se reći da ne postoji jedinstven metod za procenu rizika i njegovo tretiranje. Opšte pravilo je da vrsta i strogost postupka procene rizika zavisi od ozbiljnosti potencijalnih posledica i njihove verovatnoće. Za potencijalne događaje sa značajnim posledicama ili visokom stopom izvesnosti neophodne su stroge mere u proceni rizika. Sa druge strane, kada su posledice prihvatljive ili malo verovatne moguće je upotrebiti jednostavnije tehnike.

Rizik je definisan kao verovatnoća da će se neki događaj dogoditi i konsekvence koje bi taj događaj imao kada bi se desio. Posledice mogu biti pozitivne i negativne. Uobičajena praksa je da se rizici povezuju za događaje sa negativnim ishodom, dok se događaji sa pozitivnim posledicama tretiraju kao mogućnosti ili prilike.

Projektni rizik je funkcija raspona vrednosti koje projekat može imati, i verovatnoća da će se ove vrednosti materijalizovati u zavisnosti od distribucije verovatnoće ulaznih parametara. Da bi bio uspešan i održiv, projekat mora zadovoljiti finansijske, ekološke i socijalne ciljeve koje su definisali menadžment, akcionari i druge zainteresovane strane.

Neophodno je razumeti neizvesnosti povezane sa varijablama koje određuju vrednost projekta, kako ta neizvesnost stvara rizik i kako upravljati rizikom. Termin faktori rizika ili varijable rizika koriste se da definišu promenljive koje stvaraju neizvesnosti u evaluaciji projekta. Da bi se razumelo upravljanje rizikom vezano za evaluaciju projekta moraju se ispuniti sledeći uslovi:

- razumevanje životnog ciklusa uspešnog plana, životni ciklus se može podeliti u nekoliko faza, a odluke se moraju donositi nakon svake faze,
- uspostavljanje potrebnih informacija za donošenje zdravih odluka na kraju svake faze, informacije se opisuju kao ulazne varijable, čija je vrednost neizvesna i mogu se menjati u prostoru ili kroz vreme,
- odrediti opseg vrednosti koje ulazne varijable mogu imati i verovatnoću da će se te vrednosti desiti,
- kvantifikovanje posledica koje promena ulazne varijable može imati na ostvarivanje ciljeva,
- identifikovanje načina za smanjenje verovatnoće negativnih posledica uz povećanje verovatnoće za postizanje ili premašivanje zadatih ciljeva.

Proces upravljanja rizikom, uz stalnu komunikaciju i konsultacije, uspostavlja kontekst, identifikuje rizike, analizira rizike, procenjuje rizike, tretira rizike i prati i preispituje.

- uspostavljanje konteksta u upravljanju rizicima znači identifikovanje ciljeva organizacije i eksternih i internih faktora koji mogu biti izvor neizvesnosti, kako bi se rizici lakše identifikovali,
- identifikovanje rizika otkriva šta, gde, kada, kako i zašto se nešto može dogoditi ili pokrenuti, kao i opseg mogućih efekata na ciljeve,
- analiza rizika istražuje informacije o rizicima, efikasnost i pouzdanost kontrole, kontekst, statističke podatke, rezultate modeliranja predviđanja ili stručne procene, kriterijume rizika razvijene tokom uspostavljanja konteksta sa ciljem donošenja odluke da li je potrebno tretirati rizik,
- evaluacija/procena rizika donosi odluke da li svaki rizik spada u organizacione rizike i da li zahteva tretman, obično specificira potrebne akcije od strane menadžera na svakom nivou rizika i vreme potrebno za njihovo izvršenje, specificira se koji nivoi menadžmenta treba da prihvate kontinuirano izlaganje i toleranciju određenog nivoa rizika.
- tretiranje rizika uključuje postupke modifikacije rizika menjanjem posledica koje mogu nastati ili njihove verovatnoće. Ovaj proces zahteva kreativno razmatranje opcija i detaljno projektovanje da bi se pronašao i izabrao najbolji tretman rizika.

Upravljanje znanjem u oblasti rizika prikazuje Slika 2-1, putem UML⁷ dijagrama slučajeva upotrebe gde se mogu videti ključni akteri i procesi upravljanja rizikom.

2.2.1. Uspostavljanje konteksta

Kriterijumi za definisanje i vrednovanje rizika zasnivaju se na ciljevima organizacije, kao i na definisanju eksternog i internog konteksta. Kontekst rizika su spoljni i unutrašnji parametri ili razmatranja koja se moraju uzeti u obzir prilikom upravljanja rizikom i postavljanja obima i kriterijuma rizika.

⁷ Unified Modelling Language – jedinstveni jezik modeliranja

Eksterni kontekst organizacije čine kulturno, socijalno, političko, pravno, regulatorno, finansijsko, tehnološko, ekonomsko, prirodno i konkurentno okruženje, bilo međunarodno, nacionalno, regionalno ili lokalno; ključni pokretači i trendovi koji imaju uticaj na ciljeve organizacije; odnosi sa eksternim zainteresovanim stranama i njihove percepcije i vrednosti.

Prepoznavanje značajnih faktora u unutrašnjem okruženju koji nose neizvesnost su upravljanje, organizaciona struktura, uloge i odgovornosti; politike, ciljevi i strategije koje postoje za njihovo postizanje; sposobnosti u smislu resursa i znanja (npr. kapital, vreme, ljudi, procesi, sistemi i tehnologije); informacioni sistemi, tokovi informacija i procesi donošenja odluka (formalni i neformalni); odnosi sa internim zainteresovanim stranama i njihove percepcije i vrednosti; kultura organizacije; standardi, smernice i modeli koje je usvojila kompanija; i oblik i obim ugovornih odnosa.

Kroz uspostavljanje konteksta definišu se slabosti i snage preduzeća kao interni faktori i prilike i opasnosti kao eksterni faktori. Cilj je sagledati trenutni položaj preduzeća, šta su dobre karakteristike, a šta bi moglo da se poboljša, kako otkloniti ili umanjiti slabosti, iskoristiti prilike, a opasnosti izbeći ili ih pretvoriti u prilike.

Postavljanje opsega i granica za procenu rizika se realizuje specifikacijom prirode odluka koje se moraju doneti na bazi rezultata procene rizika; definisanje specifičnih kriterijuma koji će se koristiti za evaluaciju rizika; definisanje okvira promena, aktivnosti ili funkcija u odnosu na vreme i lokaciju; identifikovanje potrebnih studija i njihovog obima, ciljeva i potrebnih resursa; definisanje dubine, širine i strogosti procene rizika, i sl.

2.2.2. Identifikacija rizika

Identifikacija rizika uključuje prepoznavanje šta, zašto, gde, kada i kako može da se dogodi i da izvrši uticaj na sprečavanje ili povećanje sposobnost organizacije u postizanju postavljenih ciljeva. Sveobuhvatna i sistematska analiza je ključna za lociranje svih potencijalnih rizika, jer neidentifikovani rizici ostaju van dalje analize i tretiranja. Identifikacija treba da uključi sve rizike, bez obzira da li su oni pod kontrolom organizacije ili ne. Metode za identifikaciju rizika su: strukturirano mozganje, bombardovanje idejama (eng. structured brainstorming); HAZOP (Hazard and Operability Study) i FMEA (Failure Modes and Effects Analysis).

HAZOP je kvalitativna, induktivna, analitička tehnika za sistematsko ispitivanje dobro definisanih (postojećih ili planiranih) procesa ili operacija. Počeci primene HAZOP analiza, kao induktivno analitičke tehnike, se vezuju za potrebe procesne industrije u Engleskoj. Njena praktična primena počela je šezdesetih godina prošlog veka kao rezultat poboljšane verzije analize "šta-ako", pa se zato ponekad naziva i "šta-ako" analiza (what-if).

Autori Hazop-a su Rohm i Haas, koji su je prvi put primenili 1977. godine u pilot studiji koja se odnosila na ulogu hemijske industrije u poljoprivredi. Metodologija sprovođenja HAZOP studije propisana je standardom: IEC 61882:2016, a u okviru standarda IEC 31010:2009 je preporučena. Metoda je posebno pogodna za tehnološke procese u industriji i primenljiva je u okviru procesa upravljanja rizicima prema standardu ISO 31000.

Analizom se prvo odgovara na pitanje: „Šta bi se moglo dogoditi, gde i kada“. Potom se pravi lista događaja, situacija ili okolnosti koje mogu imati uticaj na postizanje svakog od relevantnih ciljeva, bez obzira da li oni mogu sprečiti, umanjiti, odložiti ili se pak može poboljšati postizanje cilja. Sledeći korak je pravljenje liste svih mogućih uzroka koji mogu pokrenuti događaj ili dovesti do situacije. Ove informacije se beleže u obrazac registra rizika. Konkretno, prvo se definišu elementi (parametri), delovi sistema, zatim njihove potencijalne promene stanja.

Nakon toga se identifikuju devijacije tih promena stanja i preko iscrpne liste ključnih reči (osnovnih, proširenih i izvedenih) konstatuje da li je određena transformacija odnosno promena stanja rizična ili ne (Crawley and Tyler 2015; Crawley, Preston, and Tyler 2000).

Analiza načina otkaza i efekata (FMEA; često se piše sa „režimima otkaza“ u množini) je postupak pregleda što većeg broja komponenata, sklopova i podsistema kako bi se identifikovale potencijalne situacije otkaza u sistemu, kao i njihovi uzroci i posledice. Za svaku komponentu, načini otkaza i njihovi rezultujući efekti na ostatak sistema evidentiraju se u određenom radnom listu FMEA. Postoje brojne varijacije takvih radnih listova. FMEA može biti kvalitativna analiza (Rausand and Hoyland 2003), ali se može iskazati i kvantitativno kada se koriste matematički modeli koeficijenata otkaza (Tay and Lim 2008) kombinuju sa statističkom podacima otkazima. FMEA je jedna od prvih strukturiranih, sistematskih tehnika za analizu otkaza, a i danas je često prvi korak studije pouzdanosti sistema.

2.2.3. Analiza rizika

Analiza rizika odnosi se na razumevanje rizika i doprinosi odlukama da li je potrebno kontrolisati rizik i koje su mere delovanja najprihvatljivije i najisplativije. Analiza rizika obuhvata razmatranje pozitivnih i negativnih posledica i verovatnoće da se te posledice mogu dogoditi, uzimajući u obzir već primenjene kontrolne mere.

U inicijalnoj fazi koriste se kvalitativne metode za analizu rizika. Kvantitativni pristup se koristi kada je potrebna preciznija i stroža definicija, generalno kada je najverovatnija posledica velika, kada postoje kvantitativni podaci i kada je donosiocima odluka potreban veći nivo preciznosti.

Analiza rizika može se sprovoditi paralelno u fazi identifikacije ili naknadno, u analizi se uvek uzima u obzir trenutno poslovanje organizacije, uzimajući u obzir postojeće kontrole i njihovu efikasnost. Takođe, razmatra se i maksimalni mogući uticaj rizika na organizaciju bez obzira na kontrolu, tj. procenjuju se posledice koje mogu da nastanu i u slučaju da postojeće kontrole izostanu ili budu neefikasne. Potencijalna izloženost se koristi za identifikaciju ključnih kontrola koje bi trebale garantovati uspeh i njihovo kontinualno nadgledanje radi praćenja stepena efikasnosti.

Rezultat analize rizika čini specifikacija: 1) prioriteta rizika za tretiranje, 2) rizika koji treba da budu u nadzoru višeg nivoa, naročito u pogledu napredovanja tretiranja rizika; 3) rizika koji treba da budu predmet osiguranja, posebno kontinuirano nadgledani i za koje se mora raditi periodična revizija.

2.2.4. Procena rizika

Proces upoređivanja rezultata analize rizika sa kriterijumima rizika sprovodi se kao dodatan postupak u proceni rizika, onda kada je to potrebno, da bi se dodatno proverilo da li je rizik i / ili njegova veličina prihvatljiva ili podnošljiva. (Ayyub 2014)

Model procene rizika otkaza rudarske opreme zasnovan na fazi logici analizira indikatore rizika, težinu, ozbiljnost pojave i mogućnost otkrivanja (Petrović et al. 2014), gde su pokazatelji rizika dati kao lingvističke promenljive. Predstavljeni model je primenjen za procenu nivoa rizika od otkaza elemenata tračnih transportera koji rade u teškim uslovima u rudniku uglja. Autori pokazuju prednosti prezentovanog modela u poređenju sa standardnim postupkom izračunavanja u FMEA metodi procene rizika.

Model za procenu rizika na deponijama industrijskog otpada mineralnog porekla (Nišić 2019) omogućava analizu svih parametara rizika, i to: scenarija potencijalnih udesa i njihovih verovatnoća, potom posledica koje proističu iz udesa uključujući njihovu značajnost. Procena rizika povezanih sa radom skladišta jalovine Veliki Krivelj, Bor (Nišić, Knežević, and Lilić 2018) primenjuje polukvantitativnu procenu rizika i koristi upotrebu matrice rizika dimenzije 4×4.

Matrice rizika predstavljaju koristan alat za predviđanje nivoa rizika ili za rangiranje rizika na osnovu procenjene verovatnoće i značajnosti posledica (Elms 1992). Iako se inicijalno ubrajaju u kvalitativne tehnike, mogu imati i polu-kvantitativnu formu, ukoliko se varijablama dodele i numeričke, pored opisnih vrednosti. Mogu biti različitih dimenzija, ali najčešće su u upotrebi matrice formata 3x3, 4x4, 5x5 i 7x7, pri čemu su najpraktičnije za upotrebu one sa neparnim brojem ćelija zbog lakšeg izdvajanja srednjih vrednosti (Nišić 2019). Matrice ne moraju biti kvadratne, dimenzije mogu biti i 4x6, ali se u rudarskoj praksi najčešće koriste dimenzije 5x5 (Nišić et al. 2016).

Identifikacija opasnosti sa procenom rizika u rudarskoj industriji, usmerena na primenu AI metoda u analizi bezbednosti u rudarskom okruženju koju predlažu Lilić, Obradović i Cvjetić (2010) predstavlja formalizaciju ekspertskog znanja. Glavni cilj sistema je procena opšteg stanja bezbednosti u rudniku zasnovana na hijerarhijskoj strukturi potciljeva gde se svaki potcilj može posmatrati kao procena skupa parametara (gas, prašina, klima, buka, vibracije, osvetljenost, geotehnička opasnost) koji određuju opšte stanje i kategoriju bezbednosti u rudniku od opasnosti koje se mogu javiti u rudarskom okruženju (Lilić, Obradović, and Cvjetić 2010), (Lilić, Ćurčić, and Radosavljević 2009). Procena rizika od zagađivanja vodnih resursa pod uticajem napuštenih rudarskih radova na prostoru Srbije je bio predmet multidisciplinarnog istraživanja (Atanacković 2018).

2.2.5. Tretman rizika

Neophodno je izabrati prioritete i implementirati najadekvatniju kombinaciju tretmana rizika, to je obično kombinacija opcija, izabranih na osnovu troškova i koristi, efikasnosti i drugih kriterijuma od značaja za organizaciju. Faktori poput pravnih, društvenih, političkih i ekonomskih takođe trebaju biti uzeti u obzir. Za identifikovanje mogućih mera za tretiranje rizika koristi se Bow-Tie analiza. Utvrđivanje isplativosti za tretman rizika uključuje primenu „cost benefit“ analize.

Pejčić i kolege (2009) navode da je upravljanje rizikom prema modelu ISO 31000 svojevrsna šansa za poštanske kompanije jer za razliku od tradicionalnog pristupa prema kom je rizik loš i povezan sa hazardom i nešto što treba prebaciti na druge, ISO 31000 upravljanje rizikom tumači kao stvaranje vrednosti i to od neizvesnosti, koja u velikoj meri karakteriše proces pružanja poštanskih usluga. Kao ključne pouke za one koji će implementirati ISO 31000 navode da ISO 31000:2009 obezbeđuje principe i generičke smernice za upravljanje rizikom, da ga mogu koristiti preduzeća, udruženja, grupe ili pojedinci i da nije specifično projektovan za određenu delatnost, sektor ili službu. Standard se može primeniti u bilo kom periodu života kompanije za različite aktivnosti i za bilo koju vrstu rizika, bez obzira na prirodu rizika, i pozitivne ili negativne posledice. Iako standard obezbeđuje generičke smernice, ne promoviše uniformnost upravljanja rizikom u organizaciji, tako da nije „recept“ ili vežba usaglašavanja sa nečim (Tarle, Petrović, and Bojković 2009).

Karović i Komazec (2010) prikazuju upravljanje rizikom na sistemskim osnovama u vojsci i navode da je preduzimanje mera za ublažavanje rizika proces izbora najbolje strategije za upravljanje rizikom, u skladu sa postojećim ograničenjima i postavljenim ciljevima. Konkretno, formira se plan izbornih strategija za upravljanje rizikom: za kontrolu rizika, smanjenje rizika, zadržavanje rizika, prebacivanje ili transfer rizika. Kontrola rizika ne pokušava da eliminiše rizik već se nalaze načini za njegov monitoring, što znači određivanje njegove verovatnoće i posledica po sistem. Smanjenje rizika podrazumeva da se vrše određene izmene u sistemu, izmene koncepta, zahteva, specifikacija i performansi (smanjenje maksimalne brzine, smanjenje radne temperature, uvođenje novih tehnologija, primena mera zaštite ...) radi smanjenja mogućnosti ostvarenja rizičnog događaja i njegovog uticaja na sistem i okolinu. Na taj način eliminišu se izvori neprihvatljivog i prihvatljivog rizika i zamenjuju se sa rešenjima koja nose mali (marginalan) rizik (Karović and Komazec 2010). Strategija zadržavanje rizika podrazumeva prihvatanje rizika ako se pojavi u sistemu, poželjna je u slučaju pojave malih rizika, ali ne i kod velikih rizika. Prebacivanje ili transfer rizika predstavlja prenos na drugu stranu, najčešće ugovornu stranu, tako da se vrši relokacija ili distribucija rizika kojim se smanjuje ukupan rizik kompanije.

Upravljačke akcije tima se organizuju kroz akcioni plan, čiji su elementi: izvor rizika; odabrane strategije i aktivnosti; projektovani datum završetka; odgovorna osoba; stvarni datum završetka, i mere performansi. Kontrola sprovođenja definisanih mera podrazumeva sistematsko praćenje i ponovnu ocenu rezultata primenjenih strategija i njihovo upoređivanje sa postavljenim kriterijumima. Aktivnosti kontrole su u direktnoj vezi sa sistemom indikatora stanja koje obezbeđuju pravovremenu i tačnu sliku stanja sistema.

SGS Minerals Services nudi rešenja koja smanjuju rizik, a jedan od primera implementacije je uradio SGS Beograd 2015 godine za JP EPS TENT⁸ kada je u rad pušten mehanički sistem za uzorkovanje i brze (online) analize sadržaja vlage i pepela u toku istovara iz vozova na osnovu kojih se izračunava toplotna moć primljenog uglja i na osnovu tih vrednosti se može dalje tretirati u proizvodnji. Cilj je bio smanjenje rizika od velikih oscilacija kvaliteta uglja koji dovode do brojnih problema u proizvodnji električne energije.

Svrha razvijenog modela za procenu rizika na deponijama industrijskog otpada mineralnog porekla (Nišić 2019) je da se u skladu sa procenjenim rizikom i mogućim udesnim situacijama preduzimaju mere za njihovo sprečavanje i uspostavi efikasan plan delovanja u takvim uslovima, te da se rizik eksploatacije deponije drži pod kontrolom.

2.3. Softveri za upravljanje rizikom

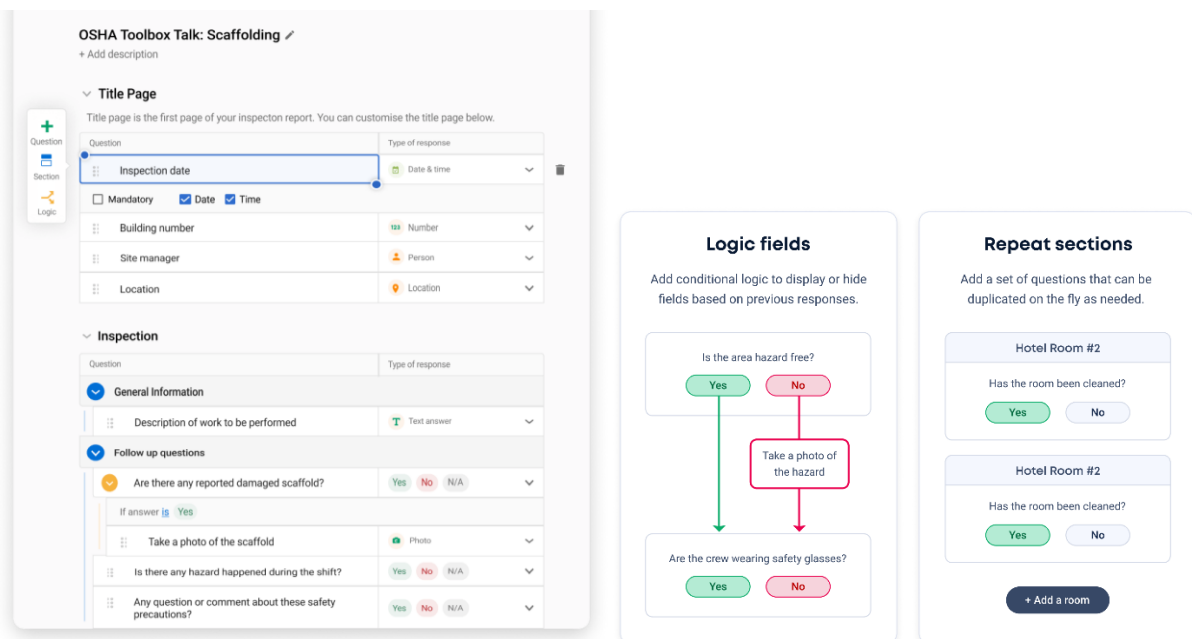
Softver za upravljanje rizikom pomaže organizacijama da smanje izloženost operativnim i drugim rizicima, poboljšavajući kvalitet i minimizirajući gubitke boljim upravljanjem podacima. Mnogi softveri za upravljanje rizikom primenjuju standardne postupke ili metode kako bi pomogli korisnicima da se izbore sa neizvesnošću i nepovoljnim uticajima. Prvo, pomažu u identifikuju od koga ili odakle mogu doći pretnje. Drugo, pružaju podršku za procenu stepena ugroženosti imovine od prethodno identifikovanih pretnji. Treće, izračunavaju ili utvrđuju verovatnoću da rizik postane problem i stepen ili obim uticaja koji će stvoriti. Zatim, softver predlaže načine za smanjenje rizika i pomaže u utvrđivanju prioriteta koje mere za

⁸ <https://www.sgs.rs/sr-latn-cs/news/2015/09/coal-sampling-system-at-tent> pristupljeno 1.3.2021

smanjenje rizika bi trebale biti prave. U nastavku je dat pregled nekoliko najzastupljenijih softvera iz oblasti upravljanja rizikom.

Kompanija *Capterra*, koja je deo *Gartner Digital Markets*, bavi se istraživanjem, testiranjem i ocenjivanjem softverskih rešenja različitih kategorija, pa među njima i softverskih rešenja za upravljanjem rizikom⁹.

iAuditor¹⁰ se koristi za sprovođenje preko 2 miliona inspekcija mesečno u svim industrijama radi bezbednosti i kontrole kvaliteta i poslovanja. Koristi ga preko 25.000 preduzeća širom sveta za obavljanje više od 600 miliona provera godišnje radi poboljšanja sigurnosti i kvaliteta na njihovim radnim mestima. Svi podaci o inspekciji prikupljaju se u realnom vremenu, omogućavajući korisnicima trenutnu detekciju propuštenih inspekcija i neuspelih akcija i brzo ublažavanje rizika. Inicijalno se kreiraju obrasci kontrolnih listi, koje mogu da se štampaju ili pretvore u Ekcel obrasce za inspekciju ili da se koriste na mobilnom uređaju (Slika 2-2). Dalje se dodaje logika koja pomaže timu da brže završi inspekcije, tako što se prikazuju i sakrivaju polja u zavisnosti od odgovora koji su izabrani. Izveštaji se generišu odmah nakon završetka inspekcije i dele sa timom, menadžerima, klijentima ili kupcima.



Slika 2-2 Primeri panela softvera iAuditor

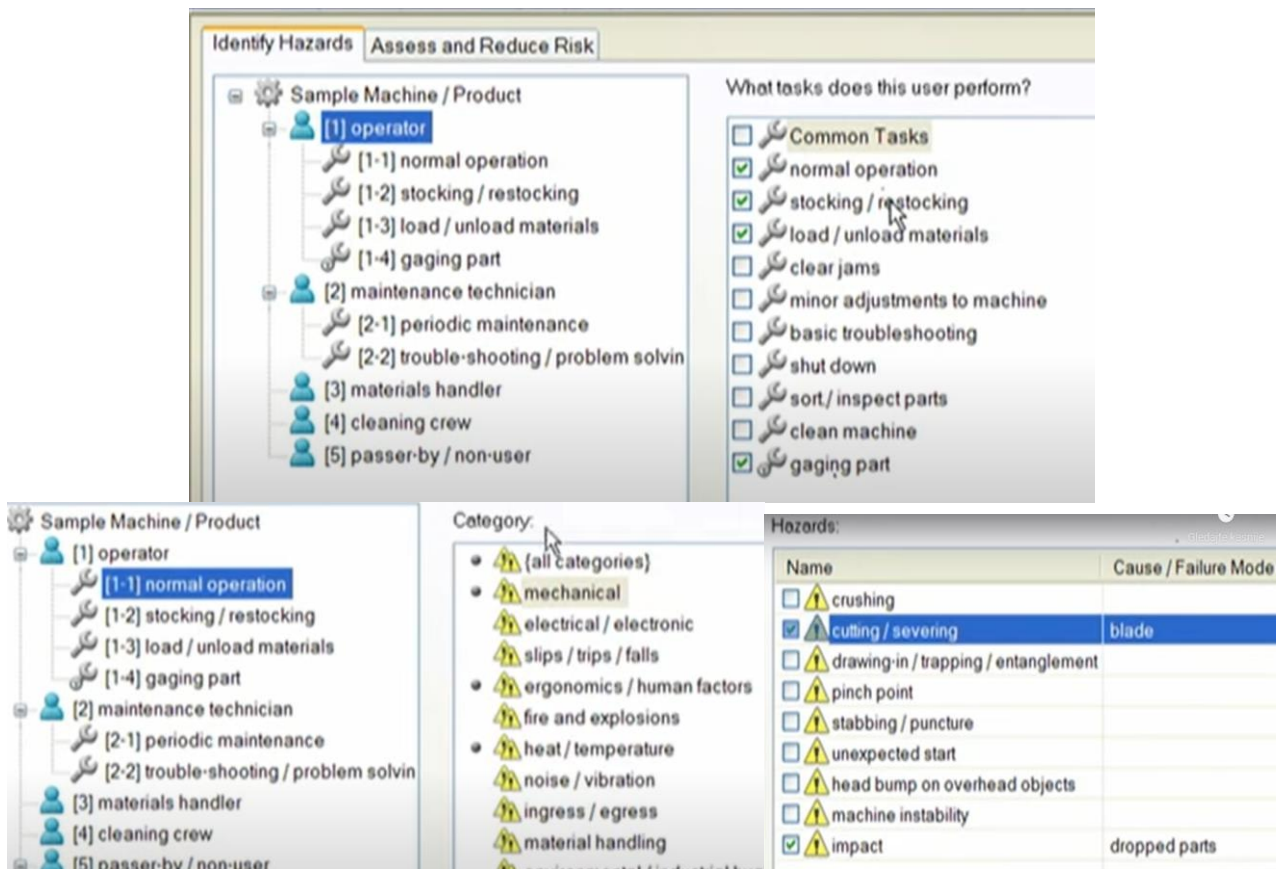
Designsafe¹¹ je alat za procenu rizika zasnovanu na zadacima, koji implementira tehnike za uklanjanje i kontrolu opasnosti i projektovanje sigurnosti. Projektantima Designsafe može poslužiti kao brz i lak alat za procenu opasnosti i rizika, kompanijama pomaže da identifikuju potencijalne opasnosti i pronađu metode za njihovo eliminisanje, sprečavanje nesreće, smanjenje troškova, poboljšavanje produktivnosti i smanjenje odgovornosti, a projektnim aktivnostima povezanim sa rizikom daje prioritet prema rangu. Prednost korišćenja alata je i pomoć korisniku da prepozna opasnost koja bi se mogla prevideti, a pruža i pomoć u pronalaženju hitnih akcija smanjenja rizika za postojeće opasnosti.

⁹ <https://www.capterra.com/risk-management-software/s/free/>

¹⁰ <https://safetyculture.com/iauditor/>

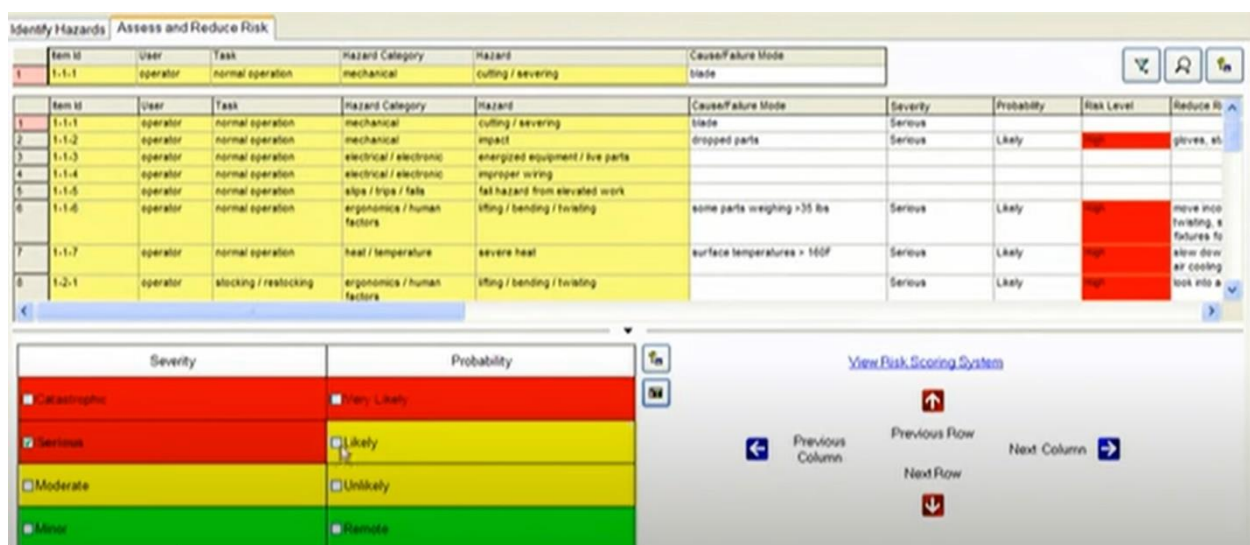
¹¹ <https://www.designsafe.net/softwareabout>

Definisanje zadataka zaposlenih je prvi korak u proceni rizika, nakon čega sledi identifikacija opasnosti za svaki pojedinačan zadatak. Konkretno, prvo se izabere kategorija, a zatim konkretna opasnost i konačno uzrok (Slika 2-3).



Slika 2-3 Primeri panela softvera Designsafe za definisanje zadataka, opasnosti i njihovih uzroka

Za svaki zadatak i identifikovanu opasnost definiše se kritičnost i verovatnoća, nakon čega se bira metoda za smanjenje rizika, odnosno mera. Dalje se opet procenjuje verovatnoća i važnost, kako bi se procenio eventualni rizik nakon primenjenih mera i videlo da li je prihvatljiv (Slika 2-4). Može da se zabeleži i ko je odgovoran, frekvencija inspekcija i slično.



Slika 2-4 Primeri panela softvera Designsafe za definisanje verovatnoće i mera

Quentic¹² je veb platforma prilagođena zadacima koja objedinjuje brojne funkcije uz kombinovanje softverskih modula koje se mogu prilagoditi potrebama: dokumentacija, organizacija, evaluacija, precizan pregled podataka, bilo da se radi o globalnim analizama ili detaljnim informacijama, specifično za oblast zaštite na radu.

ManageEngine Log360¹³ je integrisano rešenje koje upravlja dnevnicima i mrežnom bezbednošću u realnom vremenu i pomaže u zaštiti poverljivih podataka, sprečavanju unutrašnjih bezbednosnih pretnji i borbi protiv spoljnih napada.

Wrike¹⁴ je platforma za upravljanje rizikom koju koristi više od dva miliona ljudi u 140 zemalja. Daje podršku upravljanju finansijama, uz smanjenje projektnog rizika i povećanje profitabilnosti. Koriste se alati za vremensko praćenje kako bi se optimizovali resursi i osigurao okvir budžeta.

Alanen i kolege definišu model podataka¹⁵ sa ključnim činiocima sistemskog inženjerstva i sigurnosne procese sistema upravljanja mašinama. Model obuhvata zahteve i njihovu validaciju, procenu rizika, ponašanje (slučajevi upotrebe sistema i funkcionalne specifikacije), strukturu sistema i dokumentaciju. Naglasak je na definisanju veze između ovih činioca, polazeći od procene rizika do bezbednosnih zahteva, funkcionalnih specifikacija, dizajna, primene, verifikacije i validacije (Alanen et al. 2011). Slika 2-5 prikazuje komponentu sistema za preliminarnu analizu opasnosti: alat PHA (eng. Preliminary Hazard Analysis)

The screenshot shows the PHA Tool interface with the following fields and values:

- Hazard code:** PHA_1
- Risk assessment code:** RISK_ASS_1_
- Hazard type or group:** Mechanical
- Hazard zone:** Platform, Scissors service area, machine near vicinity
- Task / Operation:** Any task
- Life cycle phase:** Any life cycle phase
- Accident scenario:** While persons in duty are on the platform or there are persons in vicinity or maintenance man is in scissors service are the platform moves horizontally by itself and causes squeezing, cutting, falling, flinging etc of the person in duty or person in vicinity
- Hazardous situation:** Any situation where people are on the machine or close to the machine
- Hazardous event:** Platform moves horizontally by itself
- Hazard:** Squeezing, cutting, falling, flinging etc of the person in duty or person in vicinity
- Known safety measures:** User interface devices return automatically to the neutral position (see safety req SAFETY736)
- Recommended new safety measures:** Movement enabling is done with 2 channels (one fault will not cause an unintended movement) at least partially; communication channel is built to SIL 2 level (see IEC 61784-3 part 5.6.2; also analyze other faults than just the "corruption" faults)
- Harm severity:** 4: death, losing an eye or arm
- Probability:** 5: Very high
- Frequency:** 2: > 2 weeks to <= 1 year (duration <= 10 min)
- Avoidance:** 3: Rarely
- Risk: PL and SIL:** d 2
- Remarks:** (empty field)

Slika 2-5 Delovi panela alata PHA za procenu rizika

¹² <https://www.quentic.com/health-safety-and-environmental-management-software>

¹³ <https://www.manageengine.com/log-management/siem-solution-log360.html>

¹⁴ <https://www.wrike.com/use-cases/project-risk-management/>

¹⁵ <https://www.vtresearch.com/sites/default/files/pdf/tiedotteet/2011/T2583.pdf>

Centar za analizu rizika i upravljanje krizama (CARUK)¹⁶ izradio je onlajn veb aplikaciju za procenu rizika koja prati zahteve i strukturu standarda SRPS A.L2.003:2017 „Procena rizika – bezbednost i otpornost“. Softver je dostupan na adresi <https://procenarizika.rs>, licence za korišćenje softvera izdaju se na godišnjem nivou, a u ponudi su tri opcije: osnovna (jedan korisnik, jedan klijent, jedna procena) za korisnike koji će procenjivati rizike za svoju organizaciju, napredna (jedan korisnik, jedan klijent, više procena) – za korisnike koji procenjuju rizike za veće sisteme za koje se može raditi više odvojenih procena i zlatna (jedan korisnik, više klijenata, više procena) – za korisnike koji vrše usluge procene rizika za druga pravna društva. Sistem se suštinski iznajmljuje jer korisnici nakon registrovanja, a u zavisnosti od ugovora dobijaju pristup platformi, dakle nemaju lokalnu bazu sopstvenih podataka. Za potrebe ovog istraživanja korišćena je demo licenca koja traje 5 dana.

Za opis konkretne opasnosti predviđeno je polje za unos slobodnog teksta, dok se za ostale stavke izbor između ponuđenih alternativa vrši klikom na kućicu pored opisa.

Softver prati unos Akta o proceni rizika, koji se sastoji se od sedam delova: 1) Uvodni podaci (zakonska osnova, osnovni podaci o klijentu i o licima koja su uključena u postupak procene rizika; 2) Uspostavljanje konteksta (makrolokacija, mikrolokacija, konkurencija, istorijat štetnih događaja, veličina organizacije, način organizovanja, način i stepen zaštite u delovanju zaposlenih; 3) Procena rizika (za jedanaest grupa rizika predviđenih standardom SRPS.A.L2.003:2017 daje se opis opasnosti, određuje vrednost ugroženosti identifikovane opasnosti, stepen izloženosti ili učestalosti, ranjivosti, štete i kritičnosti, a nivo verovatnoće i posledica, nivo rizika i njegova prihvatljivost se automatski određuju); 4) Postupanje sa rizicima; 5) Nadgledanje i preispitivanje (za svaku identifikovanu opasnost daje se dodatan opis; 6) Komunikacija u vezi sa rizicima (opis); 7) Zaključak. Aplikacija automatski izračunava i nivo rizika po grupama rizika:

1. Rizici opšte poslovnih aktivnosti
2. Rizici po bezbednost i zdravlje na radu
3. Pravni rizici
4. Rizici od protivpravnog delovanja
5. Rizici od požara
6. Rizici od elementarnih nepogoda i drugih nesreća
7. Rizici od eksplozija
8. Rizici po životnu sredinu
9. Rizici od neusaglašenosti sa standardima
10. Rizici u upravljanju ljudskim resursima
11. Rizici u oblasti informaciono-komunikaciono-telekomunikacionih sistema

Slika 2-6 prikazuje karakteristične delove panela na kojima se prikazuju i unose podaci.

¹⁶ <http://www.caruk.rs/>

ZAHTEVI ZA PROCENU RIZIKA		Velicina opasnosti	Izlozenost	Ranjivost	Verovatnoća	Šteta	Kritičnost	Posledice	Nivo rizika	Kategorija rizika	Prihvatljivost
1	2	3	4	5	6	7	8	9	10	11	12
3.1	Postojanje pravne, organizacijske i planske regulative iz oblasti zaštite životne sredine.										
3.2	Opremljenost odgovarajućom opremom za reagovanje u vanrednim situacijama ili ekološkim incidentima i osposobljenost ljudskih resursa za reagovanje u njima.										
3.3	Odgovorna osoba za reagovanje u vanrednim situacijama i ekološkim incidentima.										
3.4	Biohazardni potencijal korišćene tehnologije.										
3.5	Postojanje eksterne organizacije koja vrši nadzor i direktne komunikacije sa odgovornim osobama u državnim službama za reagovanje u vanrednim situacijama ili ekološkim incidentima.										

Izloženost / Učestalost

Zanemarljiva
 Povremena
 Duga
 Pretežna
 Trajna

Vrlo retko
 Povremeno
 Često
 Pretežno
 Veoma često



◀ Prethodna

Sledeća ▶

Odustani ✕

Velicina opasnosti	Izlozenost	Ranjivost	Verovatnoća	Šteta	Kritičnost	Posledice	Nivo rizika	Kategorija rizika	Prihvatljivost
3	3	3	3	3	3	3	9	3	Neprihvatljiv

◀ Prethodna

Odustani ✕

Sačuvaj 💾

Slika 2-6 Delovi panela softvera ProcenaRizika za procenu rizika

2.4. Upravljanje rizikom zasnovano na ontologijama

2.4.1. Ontološke platforme za upravljanje rizikom

Razvoj i implementacija ontološki zasnovanih platformi (okvira) za upravljanje rizikom (eng. Ontology-based risk management (ORM) frameworks) su povećali efikasnost projekata vezanih za upravljanje rizikom tako što su poboljšali proces rada omogućavajući ponovnu upotrebu znanja (eng. knowledge reuse), uključujući identifikaciju, analizu i odgovor na rizike projekata, kao i validaciju kroz demonstraciju slučajeva upotrebe (Tserng et al. 2009).

Primer takvog okvira za upravljanje rizikom je IRIS portal¹⁷ (Meditkos et al. 2012; Angelides et al. 2013), koji je zasnovan na tehnikama i alatima semantičkog veba, kao što su OWL 2 jezik¹⁸ i OWLIM¹⁹ semantički repozitorijum. Arhitektura i funkcionalnosti semantičkog portala IRIS omogućavaju upravljanje rizicima u industriji korišćenjem vrhunskih (eng. state-of-the-art) semantičkih tehnologija. Portal omogućava veb orjentisano upravljanje slučajevima rizika koji su modelirani kao ontologije rizika, pomažući ekspertima za pojedine domene da obavljaju operativne poslove. Portal pruža mogućnost dodavanja, brisanja i ažuriranja slučajeva rizika,

¹⁷ <http://www.vce.at/iris/frameset/frame-dissemination.html>

¹⁸ OWL 2 Web Ontology Language, <https://www.w3.org/TR/owl2-overview/>, pristupljeno 1.3.2021

¹⁹ GraphDB (ranije OWLIM) najskalabilniji semantički repozitorijum <https://www.w3.org/2001/sw/wiki/Owlim>, pristupljeno 1.2.2021

omogućava funkcionalnosti krajnjim korisnicima za pretraživanje i pregled modeliranih slučajeva rizika i njihovih karakteristika, bazirano na semantičkim relacijama koje proizilaze iz ontološkog modela nakon procedure rasuđivanja (eng. reasoning procedure).

IRIS pruža novi okvir za konzistentnu procenu i harmonizovanu kvantifikaciju rizika. Angelides i ostali (2013) uspostavljaju integrisani evropski sistem za smanjenje industrijskog rizika kako bi se omogućila: veća sinergija između različitih grana industrije, smanjenje proizvodnih troškova evropskog industrijskog tržišta i obezbedili napredni bezbedni uslovi i standardi za proizvodnju, održavanje i upotrebu industrijskih proizvoda.

Paradigma rizika platforme IRIS zamišljena je kao ideja čiji je cilj da odgovori na zahtev holističkog pristupa upravljanju rizicima u evropskoj industriji, kao sveobuhvatno teorijsko i aplikativno sredstvo koje bi pomoglo menadžerima rizika u njihovom poslu u svim fazama procesa upravljanja rizikom. Paradigma rizika se kontinuirano razvija kroz rekurzivni proces u skladu sa svojim osnovnim komponentama, a to su metod identifikacije rizika, alat za procenu rizika, sistem za nadgledanje i kontrolu rizika, strategija za prevenciju i ublažavanje rizika. Sve ove komponente, kao samostalni alati i najbolja praksa za odgovarajuće procese u upravljanju rizikom (tj. identifikacija, procena, nadgledanje i reagovanje) objedinjene su na portalu rizika IRIS, koji je efikasan veb alat. Portalu IRIS direktno pristupaju praktičari i pojedinci koji se suočavaju sa stvarnim, složenim i zahtevnim problemima rizika u različitim sektorima industrije.

Ontologije u IRIS projektu formalno predstavljaju sve varijable povezane sa procesom procene rizika, naime ulaz i izlaz u alate za procenu rizika, promenljive prethodnih slučajeva ili modela, i razne druge alate koji se koristi za procenu rizika. Ontologije su potrebne da bi se olakšala integracija prakse procene rizika iz različitih domena, da bi se uklonila nerazumevanja između tih domena i alata u vezi sa korišćenim terminim, da bi se omogućili različiti načini za opis jednog pojma (sinonimi, prevodi itd.) i na kraju, da bi se omogućilo zaključivanje na višem nivou apstrakcije, tj. opšta pravila koja se primenjuju na grupu specifičnih slučajeva. IRIS ontologija sadrži 28 klasa i 32 svojstva za reprezentaciju četiri osnovna domenska koncepta, to su Case Study (Studija slučaja), Risk Case (Slučaj rizika), Risk (Rizik) i Risk Variable (Varijable rizika), i njihove relacije.

Galatesku et al. (2010) opisuje upotrebu ontologija i ontološki zasnovanih modela u sistemu obuke za prevenciju profesionalnih rizika. Semantički baziran pristup prevenciji rizika fokusiran je na personalizovanu obuku (za određeni kontekst, npr. određena aktivnost, radno mesto, tip operatera, radna mašina, itd.), koji koristi automatsko pronalaženje preventivnih dokumenata i aktivnosti koje odgovaraju zahtevima zadate obuke (Galatescu et al. 2010). Sistem integriše tri nivoa reprezentacije profesionalnih rizika i kontekst za njihovo pojavljivanje i sprečavanje: semantički okvir, modeliranje i izvršni sloj. Semantički sloj sastoji se od referentne ontologije i domenskih ontologija koje daju osnovni vokabular za domene sa potencijalnim rizicima. Ontologije domena popunjavaju stručnjaci za domene (projektanti rizika) pomoću ontološkog editora. Oni su predstavljeni taksonomijama specifičnim za domen, tj. hijerarhijskom strukturom konceptata, potom relacijama kao što su: specijalizacija, sinonimija, kompozicija (deo od) ili relacije tipa liste, kao i atributa i ograničenja pojmova i relacija u svakoj ontologiji. Atributi koncepta u ontologiji mogu se referisati na eksternu ontologiju.

Semantički sloj za prevenciju profesionalnog rizika ima korenske ontologije koje pomažu u identifikaciji i klasifikaciji faktora rizika, posledica i preventivnih mera, opasnih aktivnosti i

procesa kojima pripadaju itd. Glavni koncepti su: rizik, izvršitelj (starter ili operator), proces, aktivnost, događaj, radno mesto, posledica, radni instrument, radni objekat, dokument, preventivna akcija.

Lykourantzou je sa kolegama (2011) razvio platformu za upravljanje operativnim rizikom zasnovanu na ontologiji, koja omogućava i razmenu informacija pohranjenih u ontologijama rizika između organizacionih jedinica i kompjutersko zaključivanje preko heterogenih ORM aplikacija u organizaciji (Lykourantzou et al. 2011). Identifikacija rizika uključuje identifikaciju glavnih rizičnih događaja koji mogu negativno uticati na performanse procesa i organizacije uopšte. Svi prepoznati rizici moraju biti ispravno dokumentovani sa: kratkim opisom rizika, jasnom definicijom njihovih glavnih (korenskih) uzroka, dostupnim informacijama o stvarnim incidentima, kao i pregledom kontrolnog okruženja ili spiskom postojećih provera. Postupak procene rizika na platformi podrazumeva da se za svaki rizik koji je identifikovan u prethodnom koraku sprovedu sledeće aktivnosti: 1) procena efikasnosti postojećih kontrola / mera. 2) procena i dokumentovanje verovatnoće da se desi događaj operativnog rizika, koristeći unapred definisanu skalu ocena verovatnoće realizacije. 3) procena i dokumentovanje uticaja rizičnog događaja, korišćenjem unapred definisane skale za ocenu uticaja, uz definisanje da li rizik treba tretirati ili prihvatiti.

Osnovni koncepti za upravljanje operativnim rizikom i relacije između njih:

- belongsToProcess: ProcessRisk → Process (DP, funkcionalno).
- hasRisk: Process → ProcessRisk (OP)
- hasImpact: ProcessRisk → Impact, (OP, funkcionalno)
- hasLikelihood: ProcessRisk → Likelihood, (OP, funkcionalno).
- isResultOf: ProcessRisk → RiskEvent, (OP, funkcionalno).
- hasRootCause: ProcessRisk → RiskRootCause, (OP, funkcionalno)
- hasTreatmentPlan: ProcessRisk → TreatmentPlan (OP)
- hasIncident: ProcessRisk → RiskEvent, (OP).

Gde DP označava „data property“, odnosno svojstvo podatka, dok OP je „object property“, odnosno svojstvo objekta. To što je nešto funkcionalno svojstvo znači da primerak, odnosno instanca za to svojstvo može imati najviše jednu vrednost. U poglavlju 4 će biti više reči o razvoju ontologija za oblast upravljanja rizikom.

Alan sa kolegama (2007) je implemetirao upravljanje rizikom zasnovano na ontološki baziranoj, fleksibilnoj i slabo integrisanoj arhitekturi. Predloženo rešenje bazira se na dva suštinska koncepta: 1) modeliranje, iskopavanje²⁰ (eng. data mining) i ponovna upotreba domenskog znanja (eng. knowledge mining and reuse) upotrebom ontologije, a potom njegova integracija i 2) upotreba polu-formalnih modela rizika za iskopavanje neintuitivnih operativnih i procesnih uvida (eng. insights). Oni koriste STRATrisk²¹ za dobijanje skrivenog, teško dokučivog, implicitnog znanja i kontekstualnog, iskustvenog znanja u okviru radnih procesa gde se procedure mapiranja konceptata dodatno poboljšavaju korišćenjem usmerenih mapa koncepta (eng. directed concept maps). Metodologija obuhvata sledeće aktivnosti: 1) kreiranje zajedničkog rečnika rizika i infrastrukture za razmenu znanja; 2) stvaranje održive arhitekture

²⁰ Termin „iskopavanje podataka/informacija/znanja“ se odnosi na prevod „data/information/knowledge mining“, koji još uvek nemaju opšteprihvaćene prevode, može se pronaći i pod nazivom „istraživanje podataka“, što ipak nije adekvetan prevod, koriste se još i traženje podataka, prekopavanje podataka, rudarenje podataka, dejta majning. U ovom radu se podrazumeva proces otkrivanja šablona u velikim skupovima podataka, korišćenjem metoda mašinskog učenja, statistike i baza podataka sa ciljem izvlačenje informacija iz skupova podataka i transformacija tih informacija u mašinski razumljivu strukturu radi dalje upotrebe.

²¹ <https://stratrisk.wordpress.com/2011/02/27/risk-management-processes/>

za upravljanje znanjem i iskopavanje rizika (eng. risk mining); 3) razvoj kvalitativnih modela za detaljnu analizu i planiranje scenarija (Allan et al. 2007).

Vizija deljenja ontologija može se postići samo ako se koristi čvrsta metodologija koja garantuje kolaborativni inženjerski proces. Ključne prednosti razvoja i korišćenja deljenih ontologija se ogledaju u omogućavanju razmene znanja, njegove ponovne upotrebe i mašinske čitljivosti i razumljivosti. Ontologija STRATrisk razvijena je korišćenjem alata Protégé²². Generalizovani model znanja u obliku ontologije je obezbedio ubrzanje i povećani opseg raspoloživog specifičnog znanja za konkretan domen. Strateško odlučivanje i analiza se zasnivalo na sveobuhvatnijoj, deljenoj, zajedničkoj bazi znanja. Takođe, dodatna prednost je u vizuelizaciji koja je realizovana uz pomoć ontološkog pregledača OntoViz²³, koji kao svoju bazu koristi Graphviz²⁴, a prikazuje klase grupisane po njihovim svojstvima i informacije o tim svojstvima i grupisane instance sa listama njihovih svojstava. Ove grupe su povezane linijama koje predstavljaju uspostavljene relacije između objekata. Još jedan alat za vizuelizaciju u okviru platforme Protégé je Jambalaya²⁵, koji prikazuje informacije na sličan način, pri čemu su instance i klase grupisane po svojim svojstvima.

Opšta motivacija za razvoj informacionih sistema zasnovanih na ontologijama je sposobnost da se integrišu znanja iz različitih domena u jedan okvir, u jednu platformu. Korišćenje ontologija kao osnova za aplikacije za simulaciju rizika doprinosi konceptualnom zahtevu ponovne upotrebe, upotpunjavanja i nadogradnje znanja. Ponovna upotreba se postiže inkapsulacijom određenih delova modela rizika i njihovom razmenom pomoću unutrašnjih ili eksternih skladišta podataka, najčešće korišćenjem koncepta povezanih otvorenih podataka (eng. linked open data).

Ontologija rizika sastoji se od ontologije meta rizika, domenskog modela rizika i baze znanja o riziku. Manuelan, pojedinačnu unos podataka i specifikacija procesa, iako su delimično potrebni, mogu se zameniti automatizovanim procedurama iskopavanja, specijalnim alatima.

Nota sa kolegama (2010) uvodi metamodel koji može predstavljati osnovnu strukturu iz koje mogu biti izvedeni distribuirani modeli upravljanja rizikom u odnosu na nekoliko domena primene. Njihov pristup upravljanju rizikom zasnovan je na ontologijama rizika, specijalizovanim za predstavljanje i deljenje znanja o riziku u datom aplikativnom (eng. application) domenu. Promenom bazne ontologije, metamodel se može prilagoditi novom aplikativnom domenu, tako da se logika za upravljanje rizikom može ponovo upotrebiti uz razuman napor prilagođavanja (Nota, Aiello, and Di Gregorio 2010).

Da bi se odredilo ponašanje sistema za upravljanje rizikom za reagovanje na događaje sa mogućim negativnim uticajem na ispravno izvršavanje aktivnosti koristi se formalizam za specifikaciju zahteva RSF (eng. Requirement Specification Formalism), koji je zasnovan na pravilima. Pristup izabran za upravljanje rizicima u virtuelnom preduzeću (VE, Virtual Enterprise) fokusiran je na projektne aktivnosti. Svakoj aktivnosti pridružena je jedna ili više matrica rizika koje predstavljaju podatke o trenutnoj vrednosti izloženosti riziku. Ta vrednost se može promeniti kada se manifestuje neželjeni događaj sa uticajem na posmatranu aktivnost. Razmatrane su tri vrste matrice rizika: raspored (eng. schedule), kvalitet artefakta (eng. artifact quality) i troškovi određene aktivnosti (eng. cost of a given activity).

²² <https://protege.stanford.edu/>

²³ <https://sourceforge.net/projects/ontoviz/>

²⁴ <https://graphviz.org/>

²⁵ <https://thechiselgroup.org/jambalaya/>

Razvijena ontologija (Nota, Aiello, and Di Gregorio 2010) ima nekoliko klasa: *Activity* koja predstavlja aktivnosti planirane za sprovođenje projekta, koje u jednom trenutku mogu imati samo jedan status (novi, započeti itd.); *Organization* označava virtuelno preduzeće i pridružuju mu se aktivnosti. Aktivnost može zavistiti od jedne ili više drugih aktivnosti, a rizik može zavistiti od jednog ili više rizika. Klasifikacija daje kvalitativno određivanje izloženosti riziku (klasa: *Riskexposure* čiji primerici mogu imati vrednosti svojstava: niska, srednja, visoka ili ekstremna); klasa *Phase* se odnosi na funkcije (identifikovanje, analiziranje, planiranje, praćenje, kontrolisanje, komuniciranje). Za svaki rizik su definisane odgovornosti kako bi se upravljanje rizikom dodelilo osobi sa potrebnim specifičnim veštinama.

U tradicionalnom istraživanju kontrole finansijskog rizika često se naglašava linearna uzročno-posledična veza između rizika i rizičnog događaja (Yang and Wang 2019). Međutim, uzročni mehanizam događaja rizika finansiranja aktivnosti je složena funkcija interakcije, integracije i nadogradnje rizika. Da bi se bolje promovisao zdrav razvoj finansiranja i umanjila učestalost pojave rizičnih događaja, potrebno je proučiti karakteristike i modele korelacije između rizika. Yang i Wanka (2019) su predložili okvir za ontologiju finansijskog rizika zasnovan na karakteristikama događaja finansijskog rizika i u tom okviru su izgradili model ontologije finansijskog rizika.

2.4.2. Klasifikacija teksta u upravljanju rizicima

Lu i kolege (2009) su razvili okvir za obeležavanje specifičnih iskaza o riziku, oslanjajući se na prethodna istraživanja u oblasti ekonomije, menadžmenta, lingvistike i obrade prirodnog jezika (Lu et al. 2009). Projektovali su i implementirali automatizovani sistem za identifikaciju rizika zasnovan na okviru za obeležavanje. Za verifikaciju okvira sprovedena je studija ručnog anotiranja izjava o riziku u novinskim člancima časopisa Wall Street Journal, a evaluacija je dala obećavajuće rezultate za automatsku identifikaciju rizika.

Tim je dalje razvio okvir za kvantifikaciju rizika baziran na karakteristikama rizika definisanim u literaturi iz domena ekonomije i menadžmenta, kao i na različitim semantičkim kategorijama navedenim u prethodnim studijama iz oblasti računarske lingvistike. Rečenice u vestima su analizirali pomoću sledećih pet dimenzija: vreme, epistemički modalitet (eng. epistemic modality, odnosi se na znanje i na stepen njegove validacije), dokazivost, apstraktne ili činjenične informacije i polaritet. Da bi izvukli podatke povezane sa rizikom, oni su se fokusirali na potencijalni uticaj tri ključna aspekta: budući vremenski okvir, neizvesnost i vrednost kompanije. Iskaz vezan za rizik može se definisati kao rečenica koja uključuje buduće vreme, na neizvesnost i podrazumeva negativan uticaj na vrednost kompanije.

Kodak, based in Rochester, N.Y., said it **expects** net earnings of \$1.15 to \$1.30 cents a share for the full year. (1)

Mercedes **will** invite potential buyers to hot restaurants and special concerts where it will let them test-drive a C240 or C320. (2)

Analysts had **forecast** the company would report earnings of 90 cents a share. (3)

Buduće vreme ukazuje na iskaz koji može da sadrži nadolazeći događaj ili stanje, npr. “expects” (prev. očekuje) i “will” (prev. će). Treća rečenica ne sadrži buduće vreme, odnosi se na raniju zaradu kompanije. Za klasifikovanje pronađenih izraza vezanih za vremensku dimenziju korišćena je binarna klasifikacija (yes/no).

Although Univision has emerged as the **likely** buyer, Disney and Tribune also have expressed interest. (4)

National Semiconductor Corp. cut its revenue **forecast** for the fiscal second quarter, citing inventory and backlog adjustments from customers and distributors. (5)

Neizvesnost su iskazivali kao epistemički modalitet ili dokazivost, ili oba. Izraz "likely" (prev. verovatno) u (4) je primer neizvesnosti koja dolazi od epistemičkog modaliteta. U (5), "revenue forecast" (prev. prognoza prihoda) ukazuje na potencijalnu nesigurnost koja je povezana sa izvorom informacija. Iz izraza se mogu zaključiti različiti nivoi nesigurnosti. Kada je u pitanju kategorija izvesnost-neizvesnost, postoje istraživanja koja je dele u četiri ili pet kategorija (Rubin, Kando, and Liddy 2004), mada se obično kao preliminarna prvo izvrši binarna klasifikacija gde je rečenica klasifikovana kao „jeste neizvesnost“ ili „nije neizvesnost“.

Disappointing ICD sales were offset by a 7% increase to \$248 million in sales of pacemakers. (6)

Identifying Firm-Specific Risk Statements in News Articles 47

Brown boosted its rating on the Denver telecommunications services provider. (7) Boeing fell 2.55 to 58.02. (8)

Dimenzija vrednost kompanije ukazuje na to kako informacije utiču na očekivanje tržišta o vrednosti kompanije. I apstraktne i činjenične informacije mogu doprineti ovoj dimenziji. Polaritet rečenice takođe može nagoveštavati pravac uticaja. Treba istaći da se ova dimenzija ne može tretirati kao trosmerna klasifikacija (dobro, loše, nema efekta), kao rezultat fokusiranja na informacije na nivou rečenice, jer jedna rečenica može imati i dobre i loše implikacije. Na primer, prva polovina u (6) ima negativne implikacije, dok druga polovina ima pozitivne implikacije. Jedna od mogućnosti za rešavanje ovog problema je fokusiranje na neto vrednost kompanije. Međutim, značenje izvorne poruke može se iskriviti ako se pretpostavi da je važan samo neto efekat, stoga su Lu i kolege razmatrali vrednost kompanije kroz dve odvojene dimenzije: dobre i loše vesti. Kada rečenica pripada obema kategorijama, ukoliko osnovna poruka ima pozitivan uticaj na vrednost kompanije rečenica je obeležena kao dobra vest, i obrnuto. Druga komplikacija nastala je iz složene prirode procene vrednosti kompanije. Procene se mogu razlikovati jer anotatori²⁶ imaju različita znanja o svetu. Nerealno je pretpostaviti da anotator (pa i računarski algoritam) može imati savršeno znanje o svetu. Praktično rešenje je pretpostaviti da anotator ima osnovna poslovna znanja, ali ne poseduje detaljno znanje o određenom događaju. Pod ovom pretpostavkom, (7) je dobra vest, jer je rejting kompanije dobar pokazatelj za ukupni finansijski status kompanije; (8) su loše vesti, jer cena akcija odražava trenutnu tržišnu vrednost kompanije.

Na osnovu okvira za kvantifikaciju rizika i rezultata ručne anotacije Lu i kolege su implementirali sistem za identifikaciju rizika, koji je izdvajao, na nivou rečenice, informacije povezane sa rizikom, koristeći četiri predložene dimenzije. Svaka dimenzija je odvojeno modelirana. Za svaki ulazni članak, sistem je prvo delio ceo tekst na pojedinačne rečenice. Zatim je vršio ekstrakciju termina i svaka rečenica se pretvarala u „vreću reči“²⁷ (eng. bag of words). Četiri binarna klasifikatora su korišćena za identifikaciju informacija vezanih za buduće vreme, neizvesnost, dobre vesti i loše vesti. Ocene četiri klasifikatora su dalje korišćene za izračunavanje procene rizika rečenice.

Osnovno predstavljanje vreće reči je obuhvatalo i etikete vrsta reči (eng. part of speech, POS tags). Porterov stemmer²⁸ su koristili za pronalaženje stemova pojedinačnih reči, kako bi se

²⁶ Osobe koje obeležavaju tekst, dokumenta, dodajući informacije koje će mašine kasnije lako interpretirati i na osnovu njih obučavati modele ili vršite neke druge statističke analize

²⁷ Model vreće reči predstavlja pojednostavljujuću predstavu koja se koristi u obradi prirodnog jezika i pronalaženju informacija. U ovom modelu tekst (kao što je rečenica ili dokument) predstavljen je kao vreća njegovih reči, ne uzimajući u obzir gramatiku i čak redosled reči, ali zadržavajući višestrukost.

²⁸ stemovanje je takođe poznato i kao korenovanje ili podsecanje, a predstavlja postupak kojim se određuje koren reči, na primer za kopa za kopati, tako da zamenjuje u tekstu sve oblike: kopam, kopaš, kopaju, itd.

obezbedila normalizacija teksta (Dechow and Dichev 2002). Razmatrane su četiri transformacije oblika reči: 1) vreća reči bez stemovanja; 2) vreća reči sa stemovanjem; 3) vreća reči bez stemovanja sa POS etiketama; 4) vreća reči sa stemovanjem i sa POS etiketama. Koristili su dva popularna pristupa mašinskom učenju : model maksimalne entropije (Maxent) (matematički ekvivalent modelu diskretnog izbora) i metod potpornih vektora (eng. linear support vector machine, SVM) klasifikator sa linearnim jezgrom (Abarbanell, Lanen, and Verrecchia 1995). Evaluacija razvijenog modela nije potvrdila hipotezu i autori su zaključili da je potrebno postići bolju ekspresivnost rezultata.

Kogan i kolege (2009) procenjuju rizik na osnovu finansijskih izveštaja primenom regresije, gde za zadati deo teksta, procenjuju vrednost koja odgovara slici realne situacije izvedene iz značenja teksta (Kogan et al. 2009). Oni primenjuju regresione tehnike na obiman korpus finansijskih izveštaja, konstruišući regresione modele nestabilnosti za period nakon izveštaja. Njihovi modeli uspešno predviđaju ciljne varijable, naročito nakon uvođenja standardizacije u formi finansijskih izveštaja.

Rojabanu i Alagarsamy (2012) su predložili pristup zasnovan na tehnikama iskopavanja informacija iz teksta (eng. text mining) koje se primenjuju za proaktivno upravljanje rizikom (Rojabanu and Alagarsamy 2012). Pristup započinje analizom rizika nakon čega sledi izgradnja i popunjavanje repozitorijuma rizika. Popunjavanje se vrši na više načina: intervjuisanjem; pregledom arhive prethodnih sličnih projekata; analizom ekspertskih izveštaja, strateških izveštaja, zahteva za promenom i proširenjem obima projekta, učešćem zainteresovanih strana, formalnim sesijama identifikacije rizika, uzroka i učinka, tokova procesa i dijagrama uticaja. Preduslov za upravljanje rizikom je postojanje repozitorijuma kao resursa. Za odabrane kategorije identifikuju se rizici i pronalazi učestalost njihove pojave primenom tehnika iskopavanja iz teksta, nakon čega sledi klasifikacija koja omogućava kategorizaciju rizika u cilju proaktivnog upravljanja rizikom. Implementirani prototip alata koristi algoritam CART (Classification And Regression Trees for Machine Learning) za iskopavanje iz teksta, kojim se identifikuju sličnosti rizika u aktuelnom projektu koji se analizira, sa rizicima u prethodnim projektima. Registar rizika sastoji se od opisa rizika, verovatnoće, uticaja i drugih relevantnih detalja. Identifikovani rizik se pohranjuje u arhivu rizika čime se repozitorijum kontinualno dopunjuje. Pravila za klasifikaciju vrše grupisanje po preciznosti i tačnosti, na osnovu uticaja rizika, verovatnoće rizika, matrice rizika (izračunava se u tabeli registra rizika nakon što se unesu uticaj i verovatnoća), prioriteta rizika i kvalitativnog uticaja (opisni komentari o potencijalnom uticaju rizika).

Dasgupta i Dey (2016), su razvili platformu za analizu rizika zasnovanu na tehnikama obrade prirodnog jezika, koja obrađuje prijavljene iskaze o riziku i analizira narativne opise rizika preduzeća i klasifikuje ih u validne i nevalidne kategorije. Analitičkim postupcima su izdvajali informacije iz teksta koje se odnose na različite kategorije rizika i njihov mogući uzrok i uticaj. Za evaluaciju platforme sprovedena je studija ručnog obeležavanja u kojoj su menadžeri, stručni za datu oblast, anotirali prikupljene opise rizika. Evaluacija je ukazala na obećavajuće rezultate automatizovane analize i identifikacije rizika (Dasgupta and Dey 2016).

Primećeno je da pojava određenih jezičkih obrazaca vrsta reči igra važnu ulogu u određivanju faktora rizika i samih rizika. Automatizovani sistematski pristup analizi tekstualnih dokumenata su dodatno unapredili na osnovu povratnih informacija menadžera, tako da kontinualna izgradnja registra rizika može značajno da koristiti analitičarima rizika pri smanjenju lažnih alarma, a takođe može otkriti i nedoslednosti ili nedostatke izveštavanja o riziku.

Registar rizika je čuvao samo validne rizike verifikovane na osnovu povratnih informacija, tako da su time unapredili dalju ekstrakciju i kompilaciju uzroka, kao i njihovog uticaja na izvršenje projekta. Na primer, opis rizika „Neovlašćeni pristup osetljivim podacima na poštanskom

sandučiću PL ili GLs može rezultirati gubitkom kritičnih podataka o projektu ili učiniti podatke nedostupnim“ može spadati u vrstu rizika: „Curenje informacija“; Uzrok: „neovlašće pristup“; i uticaj: „gubitak podataka“.

Morfološki analizator su koristili za identifikovanje vremena, aspekta i modalitet glagola, na primer „očekuje, biće, ...“, potom identifikuju elemente za neizvesnost, na primer „verovatno, moguće,... koristeći rečnik epistemskih modalnih izraza, pri čemu računaju i frekvencije pojavljivanja identifikovanih elemenata. Kreiraju se kandidati fraza, odnosno obrazaca sekvenci dužine do 10, koje se pojavljuju bar dva puta, pri čemu se koristi algoritam PrefixSpan²⁹ (Jian Pei et al. 2001) da kandidata promovise za relevantan jezički obrazac. Isti PrefixSpan algoritam koristi se i za ekstrakciju pravila za utvrđivanje različitih uzroka i njihovih potencijalnih uticaja na rizik, koristeći označeni niz uzroka i uticaja. Jezičkim pravilima pridružene su vrednosti koje se koriste za određivanje redosleda u kojem će se pravila aktivirati u slučaju višestrukog podudaranja pravila.

Zlatni standardni skup podataka je činio korpus od 39000 opisa rizika dužine 3 - 250 reči uglavnom od rukovodilaca preduzeća koji izveštavaju o performansama svoje kompanije i opštim situacijama. Podgrupa od 8400 opisa ručno je anotirana informacijama: da li sadrži validnu izjavu o riziku ili ne, pri čemu je za validan opis rizika identifikovan i uzrok rizika i potencijalni uticaj. Pronašli su 3300 opisa koji odgovaraju validnim izjavama o riziku, dok 5100 opisa pripada kategoriji ne-rizika. Od 8400 anotiranih podataka, 2000 je korišćeno za identifikovanje frekvencije jezičkih obrazaca. Ostalih 6400 opisa rizika su zatim dati kao ulaz u sistem za identifikaciju rizika za evaluaciju, koja je pokazala da je F mera za klasifikaciju rizika 0.86, za uzrok i uticaj je 0.72.

2.5. Ekstrakcija koncepata i njihovih relacija

Ekstrakcija koncepata se može posmatrati kroz prepoznavanje i obeležavanje imeničkih fraza, koje se u sledećem koraku obrade povezuju relacijama. Istraživači se najčešće bave ekstrakcijom relacije sinonimije, hiperonimije (nadređeni) i hiponimije (podređeni). U ovom odeljku su predstavljeni obrasci, koji su bili inspiracija za primenu sličnih rešenja u ekstrakciji termina iz resursa na srpskom jeziku.

Jedan od prvih radova na temu ekstrakcije relacija opisuje metod za automatsko dobijanje leksičke relacije hiponimije iz narativnog teksta (Hearst 1992), gde se identifikuje skup leksiko-sintaksičkih obrazaca koji su lako prepoznatljivi i koji ukazuju na leksički odnos od interesa. Kasnije su brojni istraživači koristili³⁰ i dopunjavali inicijalni skup tako da su izgrađena brojna praktična rešenja, među kojima i ona koje koriste društvene mreže, na primer Fejsbuk (Roller, Kiela, and Nickel 2018).

Tabela 2-1 prikazuje obrasce koji su se pokazali produktivni u praksi za pronalaženje sinonima (Tovar et al. 2018).

Tabela 2-1 Obrasce za prepoznavanje relacije sinonimije (gde su S1 i S2 ontološki koncepti)

#	Obrazac za engleski jezik	Obrazac za srpski jezik
1	S1 (S2)	S1 (S2)
2	S1 is often referred to as S2	S1 često se naziva i S2
3	S1 is referred to as S2	S1 se naziva S2
4	S1 alias S2 S1 aka S2	S1 alijas S2 S1 aka S2
5	S1 as known as S2	S1 poznat kao S2

²⁹ <https://pypi.org/project/prefixspan/>

³⁰ Prema sajtu Google Akademik <https://scholar.google.com/> ima 4351 citat (na dan 1.2.2021)

6	S1 frequently abbreviated as S2	S1 često skraćeno kao S2
7	S1 usually called S2	S1 se obično naziva S2
8	S1 also called S2	S1 takođe nazvan S2
9	S1 called as S2	S1 nazvan kao S2
10	S1 is called S2	S1 se naziva S2
11	S1 are called S2	S1 se zove S2
12	S1 sometimes called S2	S1 se ponekad naziva i S2
13	S1 known as S2	S1 poznat kao S2
14	S1 also referred to as S2	S1 se takođe naziva S2
15	S1 often described S2	S1 je često opisivao S2
16	S1 commonly known as S2	S1 obično poznat kao S2

Tabela 2-2 prikazuje modifikovane Herstove obrasce koje koristi istraživački tim Fejsbuka (Roller, Kiela, and Nickel 2018).

Tabela 2-2 Obrasci za prepoznavanje obraza reči

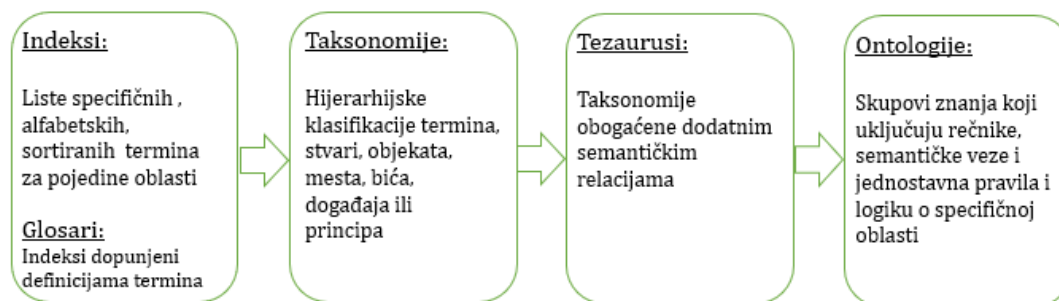
Obrazac za engleski jezik	Obrazac za srpski jezik
X which is a (example class kind ...) of Y	X koji je (primer klasa vrsta ...) Y
X (and or) (any some) other Y	X (i ili) (bilo koji neki) drugi Y
X which is called Y	X koji se naziva Y
X is JJS (most)? Y	X je superlativ Y
X a special case of Y	X poseban slučaj Y
X is an Y that	X je Y koji
X is a !(member part given) Y	X je !(član deo dat kao) Y
!(features properties) Y such as X1, X2,...	!(atributi svojstva) Y kao što su X1, X2, ...
(Unlike like) (most all any other) Y, X	(za razliku od kao) (najviše svih bilo kog drugog) Y, X
Y including X1, X2, ..	Y uključujući X1, X2, ..

Primer hiperonimije i hiponimije jeste sastavnica opreme. U strukturi rotorni bager ->radni točak->strela radnog točka, za radni točak hiperonim je rotorni bager, a hiponim strela radnog točka.

3. METODOLOGIJA I RESURSI ZA RAZVOJ ONTOLOGIJE

3.1. Šta je ontologija?

Termin „ontologija“ potiče iz filozofije gde se odnosi na granu metafizike koja se bavi postojanjem bića, dok u informatici i veštačkoj inteligenciji se odnosi na „eksplicitnu specifikaciju konceptualizacije“ (Gruber 1995). U ovom radu se pod pojmom ontologija podrazumeva isključivo značenje u informatičkom domenu. Gruber pod konceptualizacijom podrazumeva: „objekte, koncepte i druge entitete za koje se pretpostavlja da postoje u nekoj oblasti interesovanja i veze koje postoje među njima“. Kostur ontologije čine klasa, okviri ili koncepti (eng. classes, frames, concepts), za koje se vezuju njihovi primerci, odnosno instance (eng. instances, individuals). Relacije između klasa se uspostavljaju na različite načine, najčešće nasleđivanjem, formalnim pravilima, odnosno aksiomama ili, svojstvima (eng. properties, attributes). Primeri klasa su: *Uzrok*, *Posledica*, *RangUticaja*, dok bi recimo primerci za *Rang_uticaja* bili: *beznačajan*, *mali*, *umeren*, *veliki*, *ogroman*.



Slika 3-1 Od indeksa pojmova do ontologije

Slika 3-1 prikazuje poziciju ontologije u klasifikaciji terminoloških resursa sa aspekta njihovog sadržaja, strukture i načina organizacije, iako ove granice često nisu najjasnije postavljene. Svaka sledeća kategorija poseduje osobine prethodne i dobija nove odlike. Indeks predstavlja jednostavnu listu sortiranih pojmova, glosar sadrži i definicije pojmova, a taksonomija je dodatno podržana hijerarhijskom klasifikacijom elemenata. Tezaurus je pored hijerarhijskih veza obogaćen i drugim semantičkim relacijama između pojmova, a ontologije su nadograđene pravilima zaključivanja.

Osnovna namena ontologija je deljenje i ponovna upotreba znanja, kako između ljudi, tako i između softverskih aplikacija i servisa na internetu. Ukoliko kompanije ili istraživački timovi u nekoj oblasti koriste isti model podataka, iste klasifikacije i nomenklature, razmena podataka, informacija i znanja je jednostavnija i efikasnija. Da bi se postigla mašinska razumljivost podataka, potrebno je da oni budu predstavljeni u vidu standardizovanog jezika, pa se smatra da „ontologije pružaju infrastrukturu za transformaciju veb informacija i podataka u veb znanja – semantički veb“ (Gašević, Đurić, and Devedžić 2006).

Razvoj ontologija se vezuje za veštačku inteligenciju i semantički veb, ali se primena može naći u mnogim oblastima, poput geologije (Sen and Duffy 2005), rudarstva (Cox 2004; Kolonja et al. 2016), građevinarstva (Tserng et al. 2009), itd. Ontologijama se uspostavlja zajednički rečnik za aplikacije i servise semantičkog veba, na način koji je opisan standardima konzorcijuma W3C (World Wide Web Consortium)³¹. Široko prihvaćen jezik za ontologije na vebu je OWL (Web Ontology Language)³², koji se koristi u ovoj disertaciji. Još jedan važan standard koji se koristi

³¹ <https://www.w3.org/>

³² <https://www.w3.org/OWL/>

je okvir za opisivanje resursa na webu: RDF (Resource Description Framework)³³ koji osim opisa daje i semantičke veze između elektronskih resursa na webu pohranjivanjem „uređenih trojki“ (eng. triples): subjekat, predikat i objekat.

Jedna od široko korišćenih opštih baza znanja, na čije digitalne resurse se vezuju brojne domenske ontologije su i Vikipodaci³⁴ (Wikidata), u kojoj npr. koncept rudnik ima kod Q820477 i dostupan je na webu na putanji <https://www.wikidata.org/wiki/Q820477>, a definiše se kao „mesto sa koga se eksploatišu razne vrste ruda“, dok recimo rizik ima kod Q104493 i dostupan je na putanji <https://www.wikidata.org/wiki/Q104493>, a definiše se kao „potencijal nekontrolisanog gubitka zbog osnovne ranjivosti ili neizvesnosti“. Subjekti i objekti su na webu predstavljeni kodovima koji počinju sa Q na putanju <https://www.wikidata.org/wiki/>.

Na slici je prikazan Q189447 koji predstavlja pojam „upravljanje rizikom“, gde se vidi da je koncept leksikalizovan u različitim jezicima, a nivo pridruživanja definicije i sinonima zavisi i od aktivnosti zajednice konkretnog govornog područja.

управљање ризиком (Q189447)

skup mera za sistematsku identifikaciju, analizu, procenu, praćenje i kontrolu rizika
menadžment rizikom ✎ [измени](#)

✦ На другим језицима
Конфигуриши

Језик	Ознака	Опис	Псеудоними
српски / srpski	управљање ризиком	skup mera za sistematsku identifikaciju, analizu, procenu, praćenje i kontrolu rizika	menadžment rizikom
енглески	risk management	set of measures for the systematic identification, analysis, assessment, monitoring and control of risks	
шпански	gestión de riesgos	enfoque estructurado para manejar la incertidumbre relativa a una amenaza	gestión de riesgo manejo de riesgos risk management
енглески (Велика Британија)	Ознака није дефинисана	Опис није дефинисан	
арапски	إدارة المخاطر	Опис није дефинисан	مخاطر المؤسسات إدارة المخاطر ادارة المخاطر
асамски	সংশয় ব্যবস্থাপনা	Опис није дефинисан	
азербејџански	risk idaresi	Опис није дефинисан	
белоруски	кіраванне рызыкамі	галіна кіравання па мінімізацыі рызык	
Belarusian (Taraškievica orthography)	кіраваньне рызыкамі	комплекс мераў па сыстэматычнай ідэнтыфікацыі, аналізе, ацэнцы, маніторынгу і кантролю рызыкамі	
бугарски	Управление на риска	Опис није дефинисан	
бенгалски	ঝুঁকি ব্যবস্থাপনা	Опис није дефинисан	
каталонски	gestió de risc	Опис није дефинисан	
чешки	Risk management	Опис није дефинисан	
велшки	rheoli risg	Опис није дефинисан	
дански	risikostyring	Опис није дефинисан	
немачки	Risikomanagement	Gesamtheit der Maßnahmen zur systematischen Erkennung, Analyse, Bewertung, Überwachung und Kontrolle von Risiken	Risk-Management Riskmanagement Risk management Risikobetrachtung Risiko-Management Kumul-Risiko Risikomanagementsystem

Википедија (45 уноса) ✎ [измени](#)

- ar إدارة المخاطر
- as সংশয় ব্যবস্থাপনা
- bg Управление на риска
- bn ঝুঁকি ব্যবস্থাপনা
- cs Risk management
- da Risikostyring
- de Risikomanagement
- el Διαχείριση Κινδύνου
- en Risk management
- es Gestión de riesgos
- et Riskijuhtimine
- fa مدیریت ریسک
- fi Riskienhallinta
- fr Gestion des risques
- he ניהול סיכונים
- hi जोखिम प्रबंधन
- hu Kockázatkezelés
- id Manajemen risiko
- is Áhættustýring
- it Gestione del rischio
- ja リスクマネジメント
- kk Төуекел менеджменті
- kn ಅಪಾಯ ನಿರ್ವಹಣೆ
- ko 위험관리
- lv Risk-menēdžments
- lt Rizikos valdymas
- ms Pengurusan risiko
- ne जोखिम व्यवस्थापन
- nl Risicobeheer
- no Risikostyring
- pl Zarządzanie ryzykiem
- pt Gerenciamento de risco
- ro Managementul riscului
- ru Управление рисками
- si අවදානම් කළමනාකරණය

Slika 3-2 Deo panela sa vikipodacima za pojam “upravljanje rizikom”

Navedimo dalje primere svojstava, recimo svojstvo „proučava“ je P2579i nalazi se na adresi <https://www.wikidata.org/wiki/Property:P2579>, a „sastoji se od“ je P527 i pronalazi se na

³³ <https://www.w3.org/RDF/>

³⁴ <https://www.wikidata.org/>

adresi: <https://www.wikidata.org/wiki/Property:P527>, dakle za svojstva kod počinje na P i dopisuje se na putanju <https://www.wikidata.org/wiki/Property:>.

Navedimo primer RDF rečenice:

```
Q189447 P527 Q5687675,  
P527 Q11679892,  
P527 Q1058438`.
```

Prethodna RDF rečenica bi u prirodnom jeziku glasila „upravljanje rizikom sastoji se od identifikacije hazarda, analize rizika, procene rizika“, na srpskom ili „risk management consists of hazard identification, risk analysis, risk assessment“ na engleskom. Dakle ideja višejezičnih baza znanja je da se koncept definiše jedinstvenim kodom, a da se potom za svaki jezik evidentira naziv, definicija i eventualno lista sinonima kojima se leksikalizuje u konkretnom jeziku.

U ovom radu se ograničavamo na korišćenje domenske ontologije, za domen rudarstva, tačnije za upravljanje rizicima u rudarstvu. Pored domenskih, postoje i opšte ontologije (eng. upper ontologies) koje obuhvataju zajedničke koncepte za veliki broj domena, i ontologije zadataka (eng. task ontologies) sa formalnim pravilima za izvršavanje nekog konkretnog zadataka. Koncepti ontologije koja se razvija u okviru disertacije, biće povezani sa konceptima opštih ontologija i baza znanja o čemu će više reči biti u sledećem poglavlju.

Često korišćena opšta ontologija je SUMO³⁵ (eng. Suggested Upper Merged Ontology), koja je nastala objedinjavanjem i harmonizacijom više ontologija, tako da sada sadrži 13.256 termina, 154.549 formalnih pravila (aksioma) i 5.926 pravila. Navedimo samo karakteristične isečke za opisivanje nivoa rizika:

```
riskLevel  
appearance as argument number 1  
(documentation riskLevel EnglishLanguage "Relates an instance of Investing  
to the level of risk associated with the investment.")  
appearance as argument number 2  
(format EnglishLanguage riskLevel "the risk level of %1 is %2")  
consequent  
(=>  
  (and  
    (riskTolerance ?Agent ?Level)  
    (instance ?Investment Investment)  
    (possesses ?Agent ?Investment))  
    (riskLevel ?Investment ?Level))  
consequent (druga notacija)  
riskTolerance an Agent and a RiskAttribute instance an Object and  
Investment possesses Agent and Object  
riskLevel Object and RiskAttribute
```

Još jedna opšta ontologija na koju se često referišu istraživači je Dablinsko jezgro (eng. Dublin Core Ontology)³⁶, koju održava organizacija: Inicijativa za metapodatke ili „DCMI“ i publikuje u različitim formatima. U disertaciji će se koristiti RDF šema za reprezentaciju znanja (DCMI Metadata expressed in RDF Schema Language)³⁷. Navedimo samo neke elemente Dablinskog

³⁵ <http://www.ontologyportal.org/>

³⁶ <http://www.cs.umd.edu/projects/plus/SHOE/onts/dublin.html>, <https://dublincore.org/>

³⁷ <https://dublincore.org/schemas/rdfs/>

jezgra: *contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type*.

Na Rudarsko-geološkom fakultetu Univerziteta u Beogradu razvijena je ontologija RudOnto, koja bi se mogla svrstati u kategoriju laganih ontologija jer sadrži koncepte, relacije i svojstva ali ne i mehanizme zaključivanja (Stanković et al. 2012; Kolonja et al. 2016). Delovi ontologije su taksonomija rudarske opreme, klasifikacija povreda, terminologija iz oblasti nafte i gasa i sl. Glavni cilj ontologije RudOnto jeste struktuiranje i efikasna upotreba znanja u oblasti rudarstva, u okviru sistema poslovne inteligencije za upravljanje zaštitom na radu u rudarstvu (Kolonja 2016). RudOnto sadrži blizu 7.000 termina na srpskom jeziku, oko 1200 prevedenih termina na engleskom jeziku, kao i neznatan broj na francuskom, ruskom i drugim jezicima. Određeni broj termina ima pridružene sinonime, a osim korišćenja za pronalaženje informacija, upotrebu je našla i za kontrolu podataka koji se koriste u okviru sistema poslovne inteligencije (Kolonja et al. 2016). Osnove prednosti upotrebe ove ontologije jesu interoperabilnost, prelistavanje i pretraga, ponovna upotreba i kontrola domena. Interoperabilnost se ogleda u mogućnosti da informacioni sistemi i zaposleni u različitim sektorima rudarskog preduzeća koriste istu ontologiju kojom su jednoznačno i dosledno označeni koncepti koji se koriste pri radu.

GeoSciML³⁸ (eng. Geoscience Markup Language) je model podataka i standard za prenos geoloških podataka, koji pokriva domen geologije – geološke jedinice, stratigrafiju, geološke strukture, geomorfologiju, geohemiju i materijale i karakteristike koje se pojavljuju pri uzorkovanju u geonaukama – podaci vezani za bušotine, geološke uzorke i laboratorijsko ispitivanje. Jezik GeoSciML koriste brojni geološki projekti, između ostalih i projekat OneGeology Europe koji ima za cilj da se geološki prostorni podaci koje poseduju evropski geološki zavodi usaglase i postanu deljivi izvan nacionalne geološke zajednice (Blagojević et al. 2014). Jezik za obeležavanje resursa o zemlji - EarthResourceML³⁹ (eng. Earth Resource Markup Language) služi za razmenu informacija o svojstvima koja se koriste u oblastima mineralnih resursa, rudnika i rudničkih aktivnosti, a oslanja se na GeoSciML.

Ontologija za upravljanje rizikom u rudarstvu je eksplicitna formalna specifikacija koncepata sa njihovim definicijama, objektima, relacijama, pravilima međusobne interakcije, koji pružaju osnovu za analizu i procenu rizika u rudarstvu (Tserng et al. 2009). Objedinjuje značenje pojmova vezanih za rizik i definiše model koji objašnjava njihovu upotrebu. Rudarske kompanije upravljaju rizikom tako što ga identifikuju, analiziraju i zatim procenjuju da li rizik treba modifikovati tretiranjem rizika, kako bi se zadovoljili postavljeni kriterijumi. Važno je što koncepti i definicije kriterijuma rizika, zahtevi, granične vrednosti, šeme ocenjivanja mogu biti svi zajedno deo ontologije rizika i potom se integralno koristiti, deleći pohranjene podatke, informacije i znanje. Vrednosti koje se vezuju za identifikovane rizike mogu imati skalu različitih uticaja, koja uz to može varirati, tako da, težinski faktor svake klase i potklase zahteva kvantitativno merenje visine, odnosno ozbiljnosti stepena rizika (Tserng et al. 2009). Metoda analitičkih hijerarhijskih procesa (AHP) je korišćena za određivanje težine svake klase i potklase kroz ekstrakciju informacija i sistematizaciju znanja o eksplicitnom, a naročito implicitnom (skrivenom) riziku (eng. tacit risk knowledge). Korišćene su i različite vrste upitnika, kako bi se sastavili profili rizika pohranjeni u bazi znanja o riziku. Profili rizika sastoje se od rizičnih događaja, uključujući opis događaja, uzrok, posledicu, meru, status rizika i kvalitativnu analizu.

³⁸ <http://geosciml.org/>

³⁹ <http://earthresourceml.org/>

3.2. Resursi za ekstrakciju informacija

3.2.1. Rečnici iz domena rudarstva

Istraživanje u okviru disertacije se u velikoj meri oslanjalo na dostupne resurse, kako u papirnom tako i u elektronskom obliku, korišćeni su tradicionalni rečnici na papiru koji se koriste u rudarstvu, baze podataka koje pokrivaju rudarsku terminologiju, korpusi tekstova iz domena rudarstva, kao i elektronski rečnici opšte namene na srpskom jeziku.

Biro za rudarstvo Ministarstva unutrašnjih poslova SAD-a je pionir u sistematizovanju rudarske terminologije, počevši od 1918. godine kada je objavljen „Glossary of the Mining and Minerals Industry“, pa sve do 1968. godine kada je objavljen „A Dictionary of Mining, Minerals, and Related Terms“. U ovom pedesetogodišnjem projektu, više od 100 stručnjaka (inženjera, naučnika i urednika) bilo je angažovano u procesu izrade rečnika i tehničkog pregleda. Rečnik sadrži 28.750 termina objašnjenih sa 37.180 definicija (Graham 1996). Rečnik se već nekoliko decenija koristi na Rudarsko-geološkom fakultetu i u ovom istraživanju je glavni resurs koji pokriva rudarsku terminologiju na engleskom jeziku. Internet verzija rečnika objavljena je na platformi Edumine⁴⁰ koja pruža obuku za stručno usavršavanje ljudi u rudarskoj industriji.

Višejezični “Rudarski rečnik: srpsko-hrvatski, engleski, francuski, nemački i ruski rečnik rudarskih termina” objavljen je 1970. godine (Nešić 1970). Sadrži 16.500 termina koji se odnose na: površinsku eksploataciju ležišta mineralnih sirovina, podzemnu eksploataciju ležišta mineralnih sirovina, pripremu mineralnih sirovina, mehaniku stena, mehaniku tla, kao i pojmove iz oblasti geologije, metalurgije, elektrotehnike, građevinarstva i matematike sa računarskim metodama, u onoj meri u kojoj su oni povezani sa rudarstvom. Svaki pojam ima srpski termin kao odrednicu, ponekad praćenu sinonimima, koja je prevedena na četiri jezika - engleski, francuski, nemački i ruski. Međusobna povezanost svih pet jezika data je dodatnim indeksima. Termini nemaju definicije, niti primere upotrebe. S obzirom na to da je rečnik star skoro 50 godina, mnogi pojmovi su zastareli, dok neki novi pojmovi nedostaju.

Prvi terminološki „Englesko-hrvatsko-srpski naftni rečnik“ za oblast naftnog inženjerstva (S. Lazić 1976), koji prethodi „Englesko-hrvatski enciklopedijski rječnik istraživanja i proizvodnje nafte i plina“ (Perić 2007), koristi se i u Hrvatskoj i u Srbiji. Sa 12.200 definicija i 7100 pojmova, sadrži sveobuhvatan rečnik kako naučnih, tako i stručnih pojmova koje koriste naučnici, stručnjaci i studenti u oblasti istraživanja i proizvodnje nafte i gasa, ali takođe i naftne geologije, geofizike, razvoja nalazišta, bušenja i opremanja bunara, ekologije i drugih disciplina.

Obraden je, takođe, i dvojezični Terminološki rečnik iz pripreme mineralnih sirovina (P. Lazić 2020) sa 2415 prevodilačkih parova, u oba smera, englesko-srpski i srpsko-engleski, ali takođe bez definicija. Na kraju, na Rudarsko-geološkom fakultetu se koristi i rečnik „Mali terminološki rečnik pripreme mineralnih sirovina: englesko-srpski i srpsko-engleski“ sa 1400 definicija na srpskom jeziku, koji nije još zvanično objavljen, ali je dozvoljeno korišćenje za potrebe istraživanja vezanog za ovu disertaciju.

Razvoj digitalnih resursa za rudarsku terminologiju kontinuirana je aktivnost na Rudarsko-geološkom fakultetu već nekoliko godina. Počelo se sa istraživanjima vezanim za razvoj ontologije rudarske opreme (Stanković et al. 2012), u skladu sa drugim istraživanjima usmerenim na razvoj dvojezičnih leksičkih izvora (Obradović et al. 2013). Zatim je fokus bio usmeren na razvoj terminološke baze za domen opšteg rudarskog inženjerstva i njenu transformaciju iz početne prilagođene šeme u TermBase eXchange (TBX) Standard (Stanković, Obradović, and Utvić 2014). Drugi terminološki resurs, uglavnom ručno izrađen, takođe je razvijen za podršku upravljanju znanjem u određenim poddomenima rudarskog inženjerstva,

⁴⁰ <http://xmlwords.infomine.com/xmlwords.htm>

kao što su rudarska oprema, bezbednost mina i geostatistika (Kolonja et al. 2016). Tezaurus rudarske terminologije dostupan je na mreži, ali se sistematski ne ažurira. Štaviše, aplikacija nema novih zapisa, interfejs nije prilagođen novim veb čitačima niti mobilnim uređajima. Na Rudarsko-geološkom fakultetu je urađen i poduhvat sa učešćem studenata u izradi rečnika rudarstva i geologije putem kartica (eng. flashcards) u okviru predmeta Engleski jezik, gde se u nastavi koristila metoda CLIL (Content and language integrated learning) (Beko, Obradović, and Stanković 2015), ali pilot rešenje nije zaokruženo kao javno dostupan resurs.

Na Rudarsko-geološkom fakultetu ranije su razvijena tri digitalna resursa koja su uključena u ovo istraživanje. To su dve terminološke baze: Termi⁴¹ i GeoliSSTerm⁴², i ranije opisana ontologija Rudonto⁴³. Termi podržava razvoj terminoloških rečnika u različitim oblastima (matematika, računarstvo, rudarstvo, bibliotekarstvo, računarska lingvistika, elektroenergetika, itd.) (Andonovski, Šandrih, and Kitanović 2019; Radojičić et al. 2018; Šandrih, Krstev, and Stanković 2020), a odabran je kao najprikladniji resurs za sveobuhvatnu višejezičnu leksičku bazu podataka rudarske terminologije, dok su preostala dva izvora ugrađena u sistem za proizvodnju rečnika.

3.2.2. Korpusi tekstova

Za sistematski razvoj rudarske terminologije bili su potrebni odgovarajući tekstualni izvori. U tu svrhu korišćeni su članci iz naučnog časopisa Podzemni radovi (eng. Underground Mining), objavljeni na srpskom i engleskom jeziku, koji su skladišteni u dvojezičnu digitalnu biblioteku Bibliša, zbirku paralelizovanih englesko-srpskih tekstova (Stanković et al. 2015; Stanković, Krstev, Vitas, et al. 2016). Pod korpusom se podrazumeva kolekcija tekstova u elektronskoj formi koja je reprezentativna po: tematici, veličini, funkcionalnom stilu i vremenskom periodu u kome su tekstovi nastali. Reprezentativnost određenog korpusa zavisi od njegove namene (Đorđević 2017), a u ovom radu korišćen je korpus različitih tekstova, od zakona i propisa, preko udžbenika, naučnih i stručnih radova, monografija, planova, izveštaja, projekata, čime se pokušao napraviti reprezentativan resurs jezika rudarske struke.

Razvoj jednojezičnog korpusa iz domena rudarstva počeo je kao deo projekta vezanog za upravljanje rudarskom projektnom dokumentacijom korišćenjem jezičkih tehnologija (Tomašević et al. 2018). Tada je jednojezični korpus obuhvatio tekstove iz rudarskog domena i srodnih istraživanja, sa ukupno 172 dokumenta (na srpskom jeziku) i 2,7 miliona reči u prvom izdanju (Tomašević et al. 2018). Tokom istraživanja vezanih za ovu tezu proširen je sa 63 dokumenta. Trenutna verzija sadrži 4,1 milion reči. Klasifikovano prema izvorima korpus obuhvata projektnu dokumentaciju (26%), zakonodavstvo (11%), doktorske disertacije (31%), udžbenike i ostalu rudarsku literaturu (32%) (Kitanović et al. 2021).

Važno je napomenuti da je obrada teksta i sa njim povezano proširivanje rečnika iz rudarskog domena iterativni proces. Naime, između ostalih zadataka, korpusi se koriste za izdvajanje rudarske terminologije, definicija i primera upotrebe primenom različitih metoda i alata.

Kao referentni korpus savremenog srpskog jezika korišćen je SrpKor (Vitas and Krstev 2012; Utvić 2014), koji je razvio D. Vitas sa saradnicima. Njegova izrada je počela 1981. godine, a njegovo dopunjavanje traje i danas. Ovaj korpus sadrži književne i umetničke tekstove pisaca 20. i 21. veka, naučne tekstove iz oblasti društvenih i prirodnih nauka, administrativne tekstove i tekstove iz dnevnih novina, časopisa i magazina počevši od 1991. godine. Manji deo korpusa čine prevodi književnih, umetničkih i novinskih tekstova. Njegova trenutna veličina je 122 miliona reči. Tekstovi u korpusu su obeleženi bibliografski i morfološki. Bibliografske oznake

⁴¹ <https://termi.rgf.bg.ac.rs/>

⁴² <http://geoliss.mre.gov.rs/recnik/>

⁴³ <http://rudonto.rgf.bg.ac.rs/>

nose informacije o autoru, registru teksta, kao i izvornom jeziku teksta, dok morfološke oznake podrazumevaju da je svakoj reči pridružena lema i njena vrsta reči. U ovom radu je korišćen za poređenje frekvencija reči sa onim u domenskom, rudarskom korpusu na osnovu čega se određuje verovatnoća pripadnosti rudarskoj terminologiji. Osim određivanja rudarske terminologije, isti metod je primenjen i na potkorpuse rizika, kvarova, profesija.

Tekstovi koji su ušli u sastav korpusa takođe se obrađuju pomoću elektronskih rečnika i lokalnih gramatika. Lokalne gramatike su alati za opisivanje značajnih lingvističkih fenomena, počevši od jednostavnih, opštih konstrukcija koje se često pojavljuju u tekstu do složenijih konstrukcija. Osnovne prednosti lokalnih gramatika su što se one mogu nezavisno konstruisati, održavati i primenjivati i što se mogu referisati na informacije iz elektronskih morfoloških rečnika (Krstev 2008).

Korpusni tekstova su podeljeni na rečenice, rečenice na reči, a potom je svaka reč lematizovana i određena joj je vrsta reči. Lematizacija je postupak kojim se flektivni (gramatički) oblici lekseme grupišu pod osnovni rečnički, odnosno kanonski oblik⁴⁴. Srpski korpus je anotiran vrstom reči i lematizovan pomoću obučenog modela za tagiranje (Stanković et al. 2020), korišćenjem alata Treetagger ⁴⁵(Schmid 1999; 2013).

3.2.3. Morfološki rečnici

Srpski jezik ima složeni sistem gramatičkih pravila koji otežava ekstrakciju terminologije, pa je stoga upotreba morfoloških rečnika opšte namene neophodna za svaki zadatak ekstrakcije i reprezentacije znanja (Stanković et al. 2018). Važan jezički resurs koji se koristi za morfološku analizu i ekstrakciju je sveobuhvatni elektronski morfološki rečnik za srpski jezik (SrpMD), koji obuhvata opštu leksiku, vlastita imena, enciklopedijsko znanje i terminologiju iz niza različitih domena (Krstev 2008), sa skoro 200.000 leksičkih zapisa. Razvoj morfoloških elektronskih rečnika je započeo 80-tih godina prošlog veka (Vitas and Krstev 1989), a njihovo dopunjavanje neprekidno traje.

Tokom proteklih godina, veliki broj termina iz domena rudarstva dodat je u SrpMD, tako da sadrži više od 4000 zapisa prostih reči i više od 2500 fraza iz domena rudarstva. Broj njihovih morfoloških oblika, zabeleženih u ovom izvoru je znatno veći. Oblici reči koji se odnose na rudarsku terminologiju i koji su obrađeni i uključeni u SrpMD (Krstev et al. 2015), omogućili su dalje izdvajanje povezanih terminoloških fraza prema metodologiji opisanoj u (Krstev et al. 2015). Naime, da bi ekstrakcija bila efikasna, veoma je važno da je resurs relativno dobro pokriven morfološkim oblicima termina specifičnih za taj domen.

SrpMD zapisi svakom obliku pridružuju lemu i flektivne oblike sa gramatičkim informacijama, semantičke markere, informacije o domenu i relacije različitih tipova: derivacione, leksičke varijacije, relacije između komponenti višečlanih izraza, između pojedinih reči i terminoloških fraza. Ovi rečnici se nalaze u takozvanom LADL⁴⁶ formatu, a koji se koristi i za niz drugih jezika. Format ovog rečnika podrazumeva postojanje dve izdvojene grupe rečnika – rečnika prostih reči DELAS i njihovih flektivnih oblika DELAF s jedne strane i rečnika polileksemskih jedinica DELAC i njihovih flektivnih oblika DELACF. Primer DELAS i DELAF zapisa daje Tabela 3-1 Rečnik prostih (jednočlanih) reči.

Tabela 3-1 Rečnik prostih (jednočlanih) reči

	DELAS	DELAF
--	-------	-------

⁴⁴ Za imenice i zamenice nominativ jednine, za glagole infinitiv

⁴⁵ <https://cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

⁴⁶ fran. Laboratoire d'Automatique Documentaire et Linguistique

Obrazac	lema.kodKlase[+sinSem]*	obikReči.lema.vrstaReči[+sinSem]*[:kategorije]*
Primer	rizikovati.V18 +Imperf+Perf+Tr+Iref	rizikovaće,rizikovati.V+Imperf+Perf+Tr+Iref:Fzp:Fzs rizikovaćemo,rizikovati.V+Imperf+Perf+Tr+Iref:Fxp rizikovaćeš,rizikovati.V+Imperf+Perf+Tr+Iref:Fys rizikovaćete,rizikovati.V+Imperf+Perf+Tr+Iref:Fyp rizikovaše,rizikovati.V+Imperf+Perf+Tr+Iref:Azp:lys:lzs rizikovati,rizikovati.V+Imperf+Perf+Tr+Iref:W rizikuje,rizikovati.V+Imperf+Perf+Tr+Iref:Pzs rizikujem,rizikovati.V+Imperf+Perf+Tr+Iref:Pxs rizikujemo,rizikovati.V+Imperf+Perf+Tr+Iref:Pxp rizikuješ,rizikovati.V+Imperf+Perf+Tr+Iref:Pys rizikujete,rizikovati.V+Imperf+Perf+Tr+Iref:Pyp ...
Objašnjenje primera	V18 kod glagolske klase	Imperf – nesvršen Perf – svršen Tr – prelazni Iref – neprelazni V: glagol P, F, W: prezent, futur, infinitiv A, I: aorist, imperfekat x,y,z: prvo, drugo, treće lice s, p: jednina,množina

Slika 3-3 prikazuje panel alata Leximirka (Stanković et al. 2018) na kome se vide gramatički podaci vezani za ovaj termin: u dodatnom panelu na desnoj strani su prikazani svi flektivni oblici leme sa pridruženim gramatičkim informacijama o vremenu, licu, broju,... Takođe se navode leme sa kojima je osnovna lema (rizikovati) u relaciji, u ovom slučaju su to rizikovanje, kao glagolska imenica i rizikovani kao trpni pridev. Određeni broj rečnika u kojima je „rizikovati“ definisano kao odrednica se može istražiti, i u konkretnom slučaju je „otvoren“ deo panela sa jednodimenzionalnim Rečnikom Matice srpske, gde se navodi rečnički članak:

rizikovati, -kujem svr. i nesvr. izložiti, izlagati riziku, opasnosti, dovesti, dovoditi u pitanje, staviti, stavljati na kocku: ~ život, ~ u novčanim transakcijama.:

U semantičkoj mreži WordNet rečnički članak bi izgledao:

```
ENG30-02544348-v, BCS=2, +0.50000, -0.12500
Staviti na kocku u nadi za povoljnim ishodom.
rizikovati 1x V18+Imperf+Perf+Tr+Iref
okušati sreću 1 nema
ogledati 2 V601+Perf+It+Ref
DOM: factotum, SUMO: Betting +
---> hypernym: ENG30-02530167-v pokušati; probati; nastojati
---> eng_derivative: ENG30-00796315-n avantura; rabota
---> eng_derivative: ENG30-00802238-n rizik
... <--- <- eng_derivative: ENG30-00802238-n rizik
... <--- <- eng_derivative: ENG30-00796315-n avantura; rabota
```

```
ENG30-02545578-v, BCS=2, +0.00000, -0.25000
Izlagati mogućnosti da se nešto izgubi ili ošteti.
rizikovati 1 V18+Imperf+Perf+Tr+Iref
staviti na kocku 1 nema
DOM: factotum, SUMO: SubjectiveAssessmentAttribute +
---> hypernym: ENG30-02530167-v pokušati; probati; nastojati
---> eng_derivative: ENG30-00802238-n rizik
```

rizikovati

Relations:

- To rizikovanje using **glagolska imenica (ti_nxe)**
- To rizikovani using **trpni pridev (Tms_Tms)**

Check in dictionaries:

- show **RMSJ**

РИЗИКОВАТИ, -кујем свр. и несвр. изложити, излагати ризику, опасност на коцку: ~ живот, ~ у новчаним трансакцијама.

- show **Stranih reči**
- show **WordNet**
- show **RSinonima**
- show **Bi-lista**

Check in external dictionaries: [Wiktionary](#) [Babelnet](#) [Termi](#) [Glosbi](#)

Frequencies:

- Top 10000 most frequent in SrbCorp122M Corpus by D.Vitas, M.Utvić (13.86 per 1000000)

Search corpora: [Concordances](#) [Form Frequencies](#)

- SrpKorpRGF

Senses (1):

1. **+Imperf+Perf+Tr+lref**

Domains:

Properties: nesvršen, svršen, prelazni, nije povratni

rizikova, rizikovati. V: Ays: Azs
 rizikovacxe, rizikovati. V: Fzp: Fzs
 rizikovacxemo, rizikovati. V: Fxp
 rizikovacxesx, rizikovati. V: Fys
 rizikovacxete, rizikovati. V: Fyp
 rizikovacxu, rizikovati. V: Fxs
 rizikovah, rizikovati. V: Axs: lxs
 rizikovahu, rizikovati. V: lzp
 rizikovala, rizikovati. V: Gpn: Gsf: Gwm: Gwn
 rizikovale, rizikovati. V: Gpf: Gwf
 rizikovali, rizikovati. V: Gpm
 rizikovalo, rizikovati. V: Gsn
 rizikovani, rizikovati. V: Tms
 rizikovana, rizikovati. V: Tfs: Tmw: Tnp: Tnw
 rizikovane, rizikovati. V: Tfp: Tfw
 rizikovani, rizikovati. V: Tmp

Slika 3-3 Panel alate Leximirka za reč "rizikovati"

Markeri su nosioci informacija od izuzetnog značaja za pronalaženje i ekstrakciju informacija iz korpusa, a označavaju domen upotrebe date reči ili ih bliže određuju po svom značenju. Postoji više grupa markera: domenski, semantički ili sintaktički. Tako, na primer, marker +Hum označava nešto živo, marker +Org označava vrstu organizacije itd. Značaj markera se ogleda u tome što oni u kombinaciji sa gramatičkim oznakama bliže definišu reči. Njima se pravi finija distinkcija među rečima neophodna za precizniju ekstrakciju informacija.

Markere je moguće koristiti kroz regularne izraze, za neke jednostavnije upite nad korpusom ili kroz postavljanje upita konstrukcijom automata sa složenijim zahtevima. Ekstrakcija domenske terminologije, mernih jedinica, za ekstrakciju imenovanih entiteta (naziva ustanova, ličnih imena, ulica...), za uspostavljanje relacija među rečima je uspešno rađena u brojnim istraživanjima (Krstev et al. 2013; Obradović et al. 2017; Krstev et al. 2015; Stanković, Krstev, Obradović, et al. 2016; Pajić et al. 2018; Andonovski, Šandrih, and Kitanović 2019; Šandrih, Krstev, and Stanković 2020).

Rudarstvo je veoma složena industrijska grana, sa aktivnostima koje se međusobno prožimaju i nadopunjuju, ali se istovremeno i veoma razlikuju. U (Krstev 2008) definisan je opšti domen rudarstva DOM=+Mining, a kasnije u (Obradović et al. 2017) je uvedeno detaljnije obeležavanje rudarskih termina, koje je uvelo obeležavanje grana rudarstva, kao i specifične semantičke markere i time omogućiti precizniju ekstrakciju informacija iz tekstova rudarskog domena. U ovom istraživanju su korišćeni uvedeni markeri i dodati još neki potrebni za prepoznavanje koncepata rizika. Karakteristične markere sa primerima iz rečnika navodimo u nastavku:

- **+Mining** – Rudarstvo: *rudnik, rudar, ruda, eksploatacija, ugalj, mineralna sirovina, ruda bakra, etaža, jalovina, okno, bušenje, cevovod, rudarska merenja, ventilacija, nafta,...*
- **+Mach**- Mašinstvo: *rotorni bager, bager vedričar, kamion, dozer, ripper, transportna traka, cevopolagač, železnica, odlagač, transportni most, grejder, deponijska mašina,...*
- **+Safety** - Zaštita na radu: *ventilacija, zaštitna oprema, buka, vibracije, povreda na radu, profesionalno oboljenje, zaštita od požara, eksplozija, služba spasavanja,...*
- **+Transport** – Transport: *transportna traka, kamion, damper, hidraulički transport, izvoz, izvozno postrojenje, samohodno transportno sredstvo,...*

- **+RockMech** - Mehanika stena: *pritisak, smicanje, naponsko stanje, nosivost stene, fizičko-mehaničke osobine, deformabilnost stena, stabilnost kosina, nosivost tla,...*
- **+Surveying** – Geodezija: *rudnička mreža, jamski poligon, nivelmanski vlak, rudarska viseća busola, teodolit, žiroteodolit,...*
- **+EnvProt** - Zaštita životne sredine: *aerouzagađenje, zagađenje, monitoring vode, monitoring vazduha, remedijacija, rekultivacija, odlaganje otpada, ...*
- **+Surface**- Površinska eksploatacija: *etaža, završna kosina, stabilnost kosina, otkrivka, koeficijent otkrivke, deponija, miniranje, granica kopa, jalovina, rekultivacija,...*
- **+Underground** - Podzemna eksploatacija ležišta mineralnih sirovina - *podzemni istražni radovi, okno, hodnik, uskop, niskop, zaštitni stub, široko čelo, ventilacija,...*
- **+MinProcess** - Priprema mineralnih sirovina kretanje mase, uzorak, usitnjavanje, drobljenje, sito, mlevenje, mlin, klasiranje, koncentracija, flotacija, flotacijska pulpa,...
- **+Ore** - Mineralne sirovine: *ugalj, lignit, mrki ugalj, kameni ugalj, ruda gvožđa, hematit, magnetit, halkopirit, ruda olova, nafta, gas, geotermalna voda, šljunak, pesak, kamen,...*
- **+Activity** - Rudarska aktivnost: *projektovanje rudnika, miniranje, bušenje, eksploatacija mineralne sirovina, ventilacija rudnika, drobljenje rude, mlevenje rude,...*
- **+Object** – objekti: *okno, hodnik, uskop, niskop, bunker, usek, potkop, zasek, jama, drobilana, bušotina, horizont, sipka, jamska prostorija,...*
- **+Prof+Hum** – Profesije: *rudarski inženjer, rudar, geolog, merač, rudarski tehničar, bagerista, kopač, miner, površinac, podzemljaš, pripremaš, pomoćni radnik,...*
- **+Org** – Organizacije: *Kolubara, Kostolac, Rudnik, RTB Bor, Drmno, EPS,...*
- **+Instrum** - Rudarska oprema: *teodolit, bušaći čekić, drobilica, odlagač, dodavač,...*

Tabela 3-2 donosi primer DELAC i DELACF zapisa kojima se predstavljaju složene reči, odnosno reči koje se sastoje od više komponenti, u literaturi imenovane i kao: višečlane reči, polileksemske jedinica, fraze.

Tabela 3-2 Rečnik složenih reči

	DELAC
Obrazac	lemaPJ.KodKlasePJ[+SinSem]* lemaPJ = oblik1(lema1.kodKlase[:kategorije]) oblik2(lema2.kodKlase[:kategorije]) oblik3(lema3.kodKlase[:kategorije])...
Primer	odnos(odnos.N1:ms1q) uzrok-posledica.NC_N4X1 +Ek+Comp +DOM=BIinf +DOM=BIspec+DOM=BI
Oznake	N1: od imeničke klase NC_N4X1: kod klase polileksemske jedinice m: muški gramatički rod s: jednina (singular) 1: nominativ q: neanimatan Ek: ekavski izgovor Comp: polileksemska jedinica (složena reč) +DOM=BIinf +DOM=BIspec+DOM=BI: domen bibliotekarstvo i informatika, specifičan domen informatika
	DELACF
Obrazac	oblikPJ.lemaPJ.VrstaReci[+kategorije]+
Primer	odnos uzrok-posledica, odnos uzrok-posledica.N:ms1q:ms4q odnosa uzrok-posledica, odnos uzrok-posledica.N:ms2q odnosu uzrok-posledica, odnos uzrok-posledica.N:ms3q:ms7q odnose uzrok-posledica, odnos uzrok-posledica.N:ms5q odnosom uzrok-posledica, odnos uzrok-posledica.N:ms6q
Oznake	1-7 padeži: nominativ-lokativ

Slika 3-4 prikazuje rečnički članak za višičlanu lemu „*odnos uzrok-posledica*“, vide se svi flektivni (kosi) oblici i gramatičke informacije, a uz to u donjem delu panela informacije o komponentama složene reči.

Lexical Entry #227357



odnos uzrok-posledica

Check in dictionaries:

- show Terminološki

kauzalna relacija (f), odnos uzrok-{posledica}{posljedica} (m) causal relation . inf //+Blinf . Izvor: Reč. Bib. Inf.

Check in external dictionaries: Wiktionary Babelnet Termi Glosbi

Senses (1):

1. +DOM=Blinf+Comp+DOM=Blispec+DOM=Bl+Ek

Domains: informatika (BI), bibliotekarstvo i informatika
 Properties: polileksemska jedinica, ekavski oblik
 Note: BIR,jan12
 Is composed of:

Form	Lemma	FST Code	Gram Cat	Separator
odnos	odnos	N1	ms1q	
uzrok				-
posledica				

https://leximirka.jerteh.rs/LexicalEntry/Delaf?eid=227357

odnos uzrok-posledica,odnos(odnos.N1:ms1q) uzrok-posledica.NC:ms1q:ms4q

odnosa uzrok-posledica,odnos(odnos.N1:ms1q) uzrok-posledica.NC:ms2q

odnose uzrok-posledica,odnos(odnos.N1:ms1q) uzrok-posledica.NC:ms5q

odnosom uzrok-posledica,odnos(odnos.N1:ms1q) uzrok-posledica.NC:ms6q

odnosu uzrok-posledica,odnos(odnos.N1:ms1q) uzrok-posledica.NC:ms3q:ms7q

Slika 3-4 Panel alate Leximirka za reč “*odnos uzrok-posledica*”

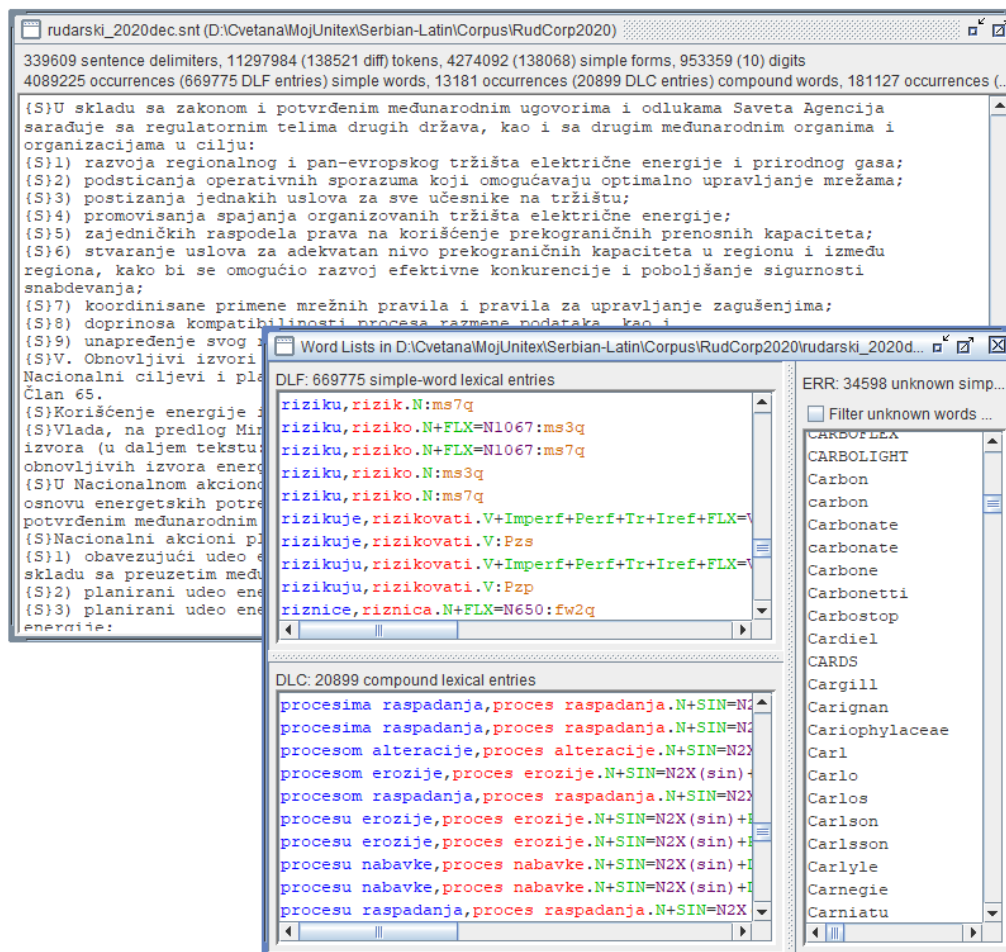
3.2.4. Obrada teksta na prirodnom jeziku - Unitex

Obrada prirodnih jezika se implementira različitim pristupima: simboličkim pravilima, statističkim metodama, mašinskim učenjem, i hibridnim pristupom. Tehnike mašinskog učenja koriste korpus za obučavanje matematičkih modela. U ovom radu je korišćen hibridni pristup, ali najvećim delom se obrada zasnivala na simboličkim pravilima Unitex⁴⁷ softvera otvorenog koda za analizu tekstualnih podataka. Obrada teksta u Unitex-u se zasniva na tehnologiji konačnih automata posebno pogodnoj za:

- Umetanje, pomeranje ili zamenu karaktera obrađenog teksta;
- Kreiranje pravila i rečnika kao konačnih automata;
- Korišćenje promenljivih generisanih na osnovu dela teksta ili bilo kojim znakovima;
- Korišćenje regularnih izraza;
- Kreiranje kaskada pravila.

Grafičko integrisano razvojno okruženje Unitexa omogućava korisnicima da kreiraju projekat, kreiraju specifične rečnike, testiraju pravila zasnovana na grafovima i, primenjuju sve ove resurse na tekstualne datoteke. Slika 3-5 prikazuje obrađen tekst korpusa, gde se u zaglavlju vidi da se korpus sastoji od 339.609 rečenica, 11,3 miliona tokena, 4,2 miliona reči, od toga 138 hiljada različitih. Donji deo slike prikazuje rečnik obrađenog teksta sa prozorima za: DLF proste reči (eng. simple words), DLC složene reči (eng. compound words) i ERR nepoznate reči, odnosno reči kojih nema u elektronskom morfološkom rečniku.

⁴⁷ <https://unitexgramlab.org/>



Slika 3-5 Panel alata Unitex sa obrađenim rudarskim korpusom

Na ovako obrađen korpus se mogu primeniti različiti obrasci za pretragu, recimo: „<A><rizik>“ koji će da vrati sve konkordance odnosno rečenice u kojima iza prideva sledi reč „rizik“. U upitima je moguće kombinovati i semantičke markere, recimo „<N+Hum><rizik>“ će ekstrahovati rečenice u kojima se može naći: menadžer ~, analitičar ~, teoretičar ~, izazivač ~, preuzimač ~,... Slika 3-6 prikazuje primere konkordanci prethodnog upita. Više o obrascima i načinima na koje se mogu istražiti spojevi reči će biti u odeljcima 4.1.1 i 4.1.

požarne zaštite, menadžmentu sistema i [analitičaru rizika](#). {S} Model pruža široke mogućnosti za
 tuju pre svega vrednosti i interesi. {S} [Izazivač rizika](#) mora da računa ne samo sa posledicama na
 n, ugalj, rudarstvo, menadžment rizika, [menadžer rizika](#). {S}1. UVOD Optimalna realizacija planir
 broj informacija o prethodnim štetama, [menadžer rizika](#) može da ove informacije iskoristi za pr
 }3.1.2 Funkcija upravljanja finansijama [Menadžer rizika](#) može direktno da bude odgovoran finansi
 {S} Male i srednje firma najčešće imaju [menadžera rizika](#) koji ne radi sa punim radnim vremenom.
 nuli", što znači da se prema mišljenju [menadžera rizika](#) ovaj događaj neće ni desiti, - "neznat
 a, jer da bi neki rizik privukao pažnju [menadžera rizika](#), on mora da pređe neki minimalni prag
 vatnu ponudu pokrića71. {S}Prioritet za [menadžere rizika](#) današnjice postali su rizici koji izla
 i upravljanja rizikom U mnogim firmama, [menadžeri rizika](#) imaju odgovornost da paze da firma bud
 sikolozima, ekonomistima, inženjerima i [menadžerima rizika](#)170 {S}5. {S}1 Rizik zagađivanja život
 ili projektovanog budućeg dohodka. {S}- [Menadžeru rizika](#) su potrebni podaci o nabavnoj vrednost
 sursa i potrebnog kapitala na različite [nosioce rizika](#) je od suštinske važnosti za takvo upravl
 će biti odgovorni za rizik nazivaju se [nosiocima rizika](#). {S}Planovi odgovora se izrađuju za on
 }0 inkorporaciji firma prenosi na svoje [poverioce rizik](#) od nemogućnosti isplate šteta i drugih
 ovaj način špekulator preuzima rizik, a [prenosilac rizika](#) je obično firma koja želi da rizik ce
 g podizvođača. {S}- Kod ugovora kada se [preuzimač rizika](#) obaveže da će isplatiti sve štete za k
 su zajednički odgovorni i prenosilac i [preuzimač rizika](#). {S} Na primer, pre nego što započne sa
 tima, mora posvetiti posebna pažnja. {S} [Preuzimači rizika](#) moraju veoma pažljivo sagledati proje
 slučaju mogu poslužiti kao orijentacija [preuzimačima rizika](#), da u zavisnosti od stanja faktora
 pavnosti i štetnosti, koje prema nalazu [procenjivača rizika](#) mogu da izazovu povredu ili ugroze
 entivnih službi {S}7. {S}7. {S} Proizvod «[procenjivača rizika](#)» je prepoznavanje opasnosti koje tr
 procenjivanje rizika (u daljem tekstu: [procenjivači rizika](#)); {S}4) metode za vršenje procene r
 drugi gubitak. {S}Ekonomisti, naučnici, [teoretičari rizika](#), statističari i aktuari, svi oni ima
 pod kojima se vrednuje značaj rizika. • [Vlasnik rizika](#) je osoba ili entitet koji su krajnje odg

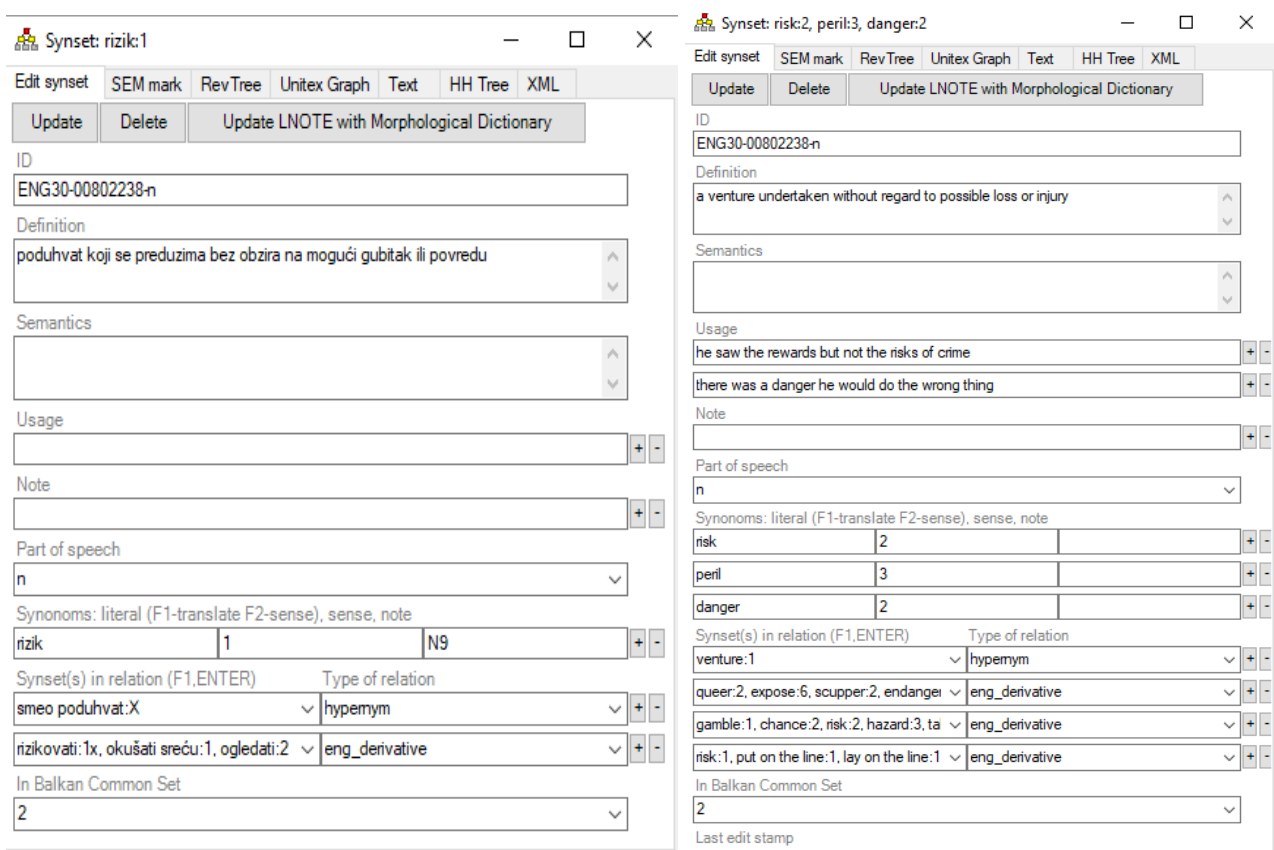
Slika 3-6 Panel alata Unitex sa konkordancama za upit „<N+Hum><rizik>“

3.2.5. Leximir i Wordnet

Leximir je alat za rukovanje leksičkim resursima i za kreiranje proširenih upita, koji sadrži nekoliko komponenti za izvršavanje različitih funkcija. Imajući u vidu da su resursi korišćeni u tezi bili dostupni u različitim formatima i kodnim šemama, korišćen je modul za konverziju koji omogućava transformisanje resursa iz jednog formata u drugi (uglavnom je reč o konvertovanju tekstova, rečnika, grafova i regularnih izraza iz jedne kodne šeme u drugu) (Stanković et al. 2011). Generisanje teksta na osnovu pdf dokumenata je podrazumevalo razrešavanje prekinutih rečenica znakom paragrafa, prelomljenih reči crticom, kao i druge probleme nastale optičkim prepoznavanjem teksta (Krstev, Stanković, and Vitas 2018).

Podsistem za održavanje morfoloških rečnika je omogućio pretraživanje i izdvajanje lema na osnovu poređenja sa zadatom niskom karaktera, vrstom reči, kodom flektivne klase, sintaksnim i semantičkim oznakama. Uz to, Leximir obezbeđuje vezu sa morfološkim grafovima koji opisuju flektivna svojstva pojedinačne reči (leme).

Wordnet je semantička mreža reči pohranjena u leksičkoj bazi podataka koja je organizovana kao graf čvorova i relacija između tih čvorova. Ovi čvorovi se nazivaju sinsetovi (od eng. synset ili synonymous set) i predstavljaju skupove reči koje u nekom kontekstu imaju isto (sinonimno) značenje. Srpski wordnet kao semantička baza razvijen je u okviru projekta BalkaNet koji se bavi razvojem wordnet-a kao višejezičke baze podataka za balkanske jezike (Koeva, Krstev, and Vitas 2008). Slika 3-7 prikazuje na levoj strani srpski, a na desnoj engleski sinset čiji je ID: ENG30-00802238-n. Slika 3-8 prikazuje hijerarhijsku strukturu: nadređene i podređene sinsetove selektovanog sinseta.

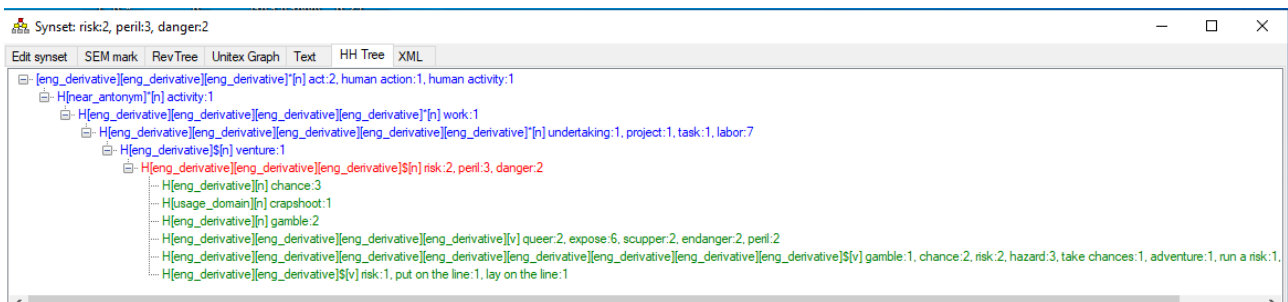


Slika 3-7 Panel alata Leximir sa sinsetima srpskog i engleskog WordNeta

Podsistem za interakciju sistema elektronskih rečnika i semantičke mreže Wordnet omogućava razmenu informacija između wordnet-a i morfoloških rečnika – morfosintaksičke informacije iz rečnika se pridružuju literalima u sinsetovima, a semantičke informacije iz sinsetova lemmama u rečnicima, što je iskorišćeno za leksikalizaciju ontologije: pronalaženje sinonima kojima se opisuju koncepti sa svim odgovarajućim flektivnim oblicima, ortografskim i leksičkim varijantama.

Interna reprezentacija WordNeta je XML struktura koju dajemo u nastavku:

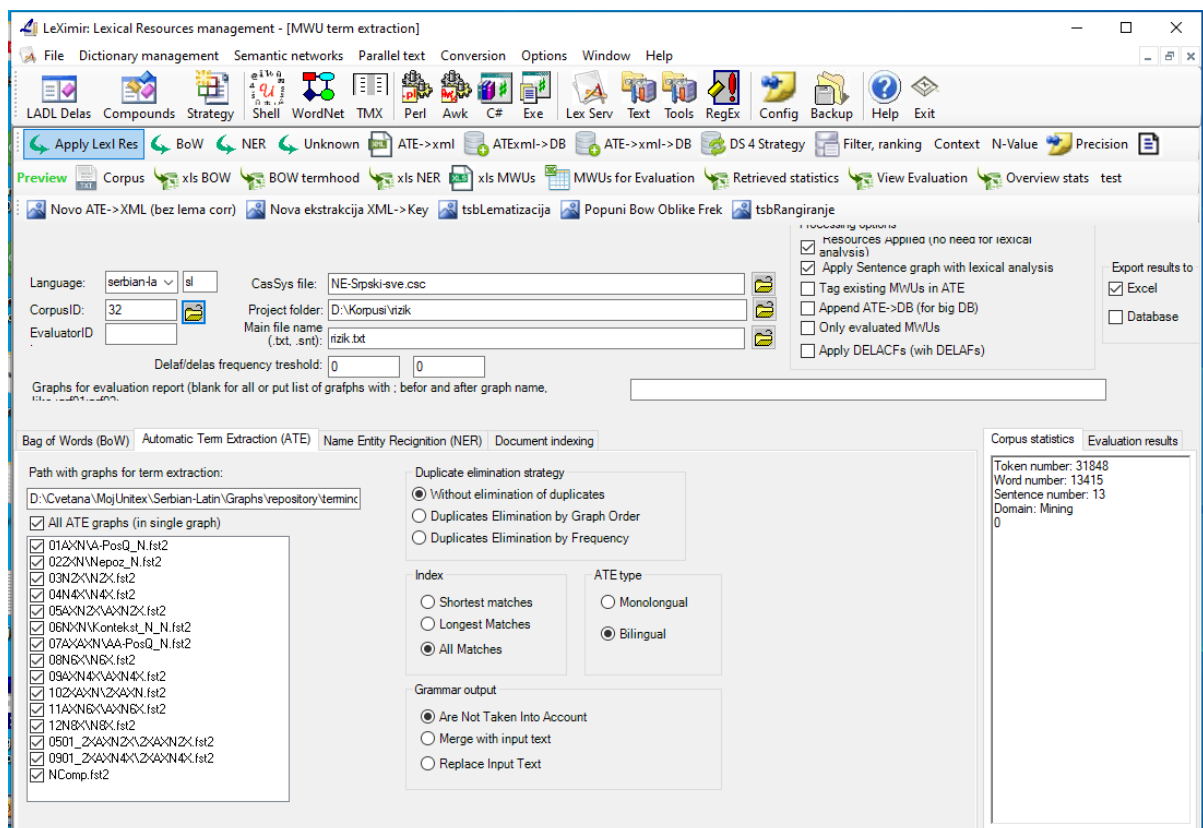
```
<SYNSET>
  <ID>ENG30-00802238-n</ID>
  <POS>n</POS>
  <SYNONYM>
    <LITERAL>rizik<SENSE>1</SENSE><LNOTE>N9</LNOTE></LITERAL>
  </SYNONYM>
  <DEF>poduhvat koji se preduzima bez obzira na mogući gubitak ili
  povredu</DEF>
  <BCS>2</BCS>
  <ILR>ENG30-00797878-n<TYPE>hypernym</TYPE></ILR>
  <ILR>ENG30-02544348-v<TYPE>eng_derivative</TYPE></ILR>
  <NL>yes</NL>
  <STAMP>Cvetana</STAMP>
  <SUMO>SubjectiveAssessmentAttribute<TYPE>+</TYPE></SUMO>
  <SENTIMENT><POSITIVE>0</POSITIVE><NEGATIVE>0.625</NEGATIVE></SENTIMENT>
  <DOMAIN>factotum</DOMAIN>
  <RILR>ENG30-00797878-n<TYPE>hyponym</TYPE></RILR>
  <RILR>ENG30-02544348-v<TYPE>eng_derivative</TYPE></RILR>
  <RILR>ENG30-02545578-v<TYPE>eng_derivative</TYPE></RILR>
</SYNSET>
```



Slika 3-8 Panel alata Leximir sa stablom nadređenih i podređenih sinseta „rizik“

Podsistem za predikciju flektivne klase (prostih) reči zasnovan je na pravilima koja koriste metapodatke o flektivnim klasama i statistiku sufiksa postojećih rečničkih odrednica. Podsistem za izdvajanje termina iz teksta i generisanje ispravne leme i cele rečničke odrednice na osnovu morfološke i sintaksičke analize, (Stanković, Krstev, Obradović, et al. 2016) je korišćen za ekstrakciju kandidata – instanci ontologija, što će biti prikazano u odeljcima 4.1.1 i 4.1.

Podsistem za predikciju klase višestanih reči i generisanje rečničke odrednice cele leme za zadatu listu termina zasnovan je na sistemu tvorbenih pravila i opisa morfoloških osobina srpskog jezika uz korišćenje grafova. Panel za ekstrakciju višestanih reči prikazuje Slika 3-9, a u donjem levom uglu se vide obrasci sintaksičkih grafova razvijenih u alatu Unitex koji su integrisani u sistem. Kako se veći deo instanci ontologije leksikalizuje polileksemskim jedinicama ovaj modul je korišćen za određivanje flektivnih oblika višestanih termina, uglavnom za opisivanje kvarova, opasnosti, posledica itd., kako bi bile prepoznate u tekstu i obeležene na odgovarajući način (Stanković et al. 2011).



Slika 3-9 Panel alata Leximir sa panelom za ekstrakciju višestanih reči iz korpusa

3.3. Semantička mreža okvira FrameNet

Teorija semantike okvira (eng. frame semantics) je kognitivno-semantički pristup Čarlsa Filmora koja eksplicitno povezuje značenja neke reči sa sintaksičkim kontekstima u kojima se ta reč javlja (Fillmore 1982). Brojni autori su saglasni da je teorija semantike okvira prigodan način da se analizira i opiše način korišćenja neke reči, tako da se značenja većine reči mogu najbolje razumeti na osnovu semantičkog okvira, opisa vrste događaja, relacije ili entiteta i učesnika u njemu. Analiza njenog semantičkog sadržaja i otkrivanje gramatičkih konstrukcija u kojima se javlja se bazira na ispitivanju njenog ponašanja u korpusu. Smatra se da je neophodno opisati sve dopune i odredbe u tim konstrukcijama u onoj meri u kojoj značenje date reči zavisi od njih. Posebna pažnja poklanja se rečima koje se ne mogu pravilno upotrebiti ukoliko se ne znaju konstrukcije u kojima se javljaju. Takve reči nazivaju se rečima koje prizivaju okvire (eng. frame-evoking words), a to su imenice, pridevi i glagoli (Расулић and Кликовац 2014).

Analiza podrazumeva nekoliko koraka: 1) utvrđivanje elemenata okvira (elementi odgovaraju semantičkim ulogama); 2) utvrđivanje valentne grupe, koju čine jedan od elemenata okvira zajedno sa svojom gramatičkom realizacijom (tip imeničke jedinice i njena funkcija u rečenici); 3) valentni⁴⁸ model (skup valentnih grupa realizovanih u jednoj rečenici); 4) opis valence (skup svih valentnih modela koje realizuje jedna leksička jedinica).

Termin „okvir“ Filmor koristi kao opšti termin za skup pojmova koji su u literaturi o razumevanju prirodnih jezika poznati pod nazivima shema, scenario, kognitivni model, narodna teorija i neki drugi. FrameNet⁴⁹ je organizovan oko konceptualnih struktura poznatih kao okviri. Semantički okvir predstavlja scenu, vrstu događaja, stanja ili drugog scenarija, može biti univerzalan ili kulturološki specifičan, uz to može biti opšti za domen ili specifičan za domen. Okvir definiše uloge učesnika ili elemente okvira (FE), čiji odnosi čine konceptualnu podlogu potrebnu za razumevanje (određenih čula) stavki (eng. items) rečnika.

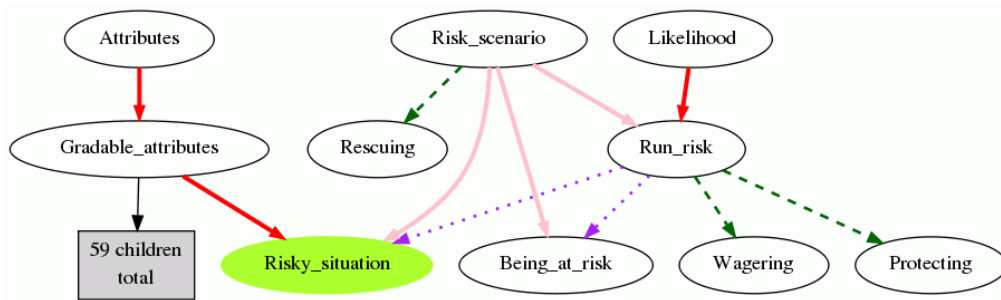
U odnosu na scenario komercijalne transakcije (okvir) su definisani: glagoli kao što su kupiti, prodati, platiti (buy, sell, pay), i imenice kao kupac, prodavac, cena, kupovina (as buyer, seller, price, and purchase). Elementi okvira (FE, frame elements) koji su centralni za ovaj okvir—mogu se ili ne moraju pominjati izričito u tekstu sa jednom od pomenutih leksičkih stavki: kupac, prodavac, roba koju prodavac prodaje i novac dat kao plaćanje u zamenu od strane kupca.

Elementi okvira su formalno navedeni uz opis njihove funkcije u svom okviru. Okviri su organizovani u mrežu, uključujući hijerarhiju nasleđivanja i drugih vrsta odnosa između okvira. Formalno, leksička jedinica povezuje lemu sa okvirom, a kada se analizira u tekstu kaže se da priziva (eng. evoke) okvir. Rečenice su anotirane uzimajući u obzir reči koji prizivaju okvir i ostale elemente okvira. Alat za vizuelizaciju relacija između okvira i elemenata okvira⁵⁰ pruža mogućnost da se izabere početni okvir i sistematski istražuju veze između okvira. Slika 3-10 prikazuje semantički okvir Rizična_situacija (eng. Risky_situation):

⁴⁸ Svojsvo glagola da za sebe „vezuje“ rečenične članove naziva se valenca (vrednost, važnost), a označava broj veza koje sintaksički elementi rečenice uspostavljaju između sebe.

⁴⁹ <https://framenet.icsi.berkeley.edu/>

⁵⁰ <https://framenet.icsi.berkeley.edu/fndrupal/FrameGrapher>



Slika 3-10 Semantički okvir Rizična_situacija (eng. Risky_situation)

U nastavku je prikazan okvir *Rizična_situacija* gde se nakon definicije daju primeri, a potom ključni elementi okvira i sporedni elementi. Svakom elementu okvira se pridružuje različita boja, tako da se u tekstu definicije i njihova pojavljivanja boje na odgovarajući način. Leksičke jedinice koje karakterišu ovaj okvir su: *opasnost.N*, *opasan.A*, *rizik.N*, *rizično.ADV*, *bezbedan.A*, *bezbedno.ADV*, *pretnja.N*, *nebezbedno.A*, *pretnja.N*, *pretiti.V*, gde N označava imenicu, A pridev, ADV prilog, V glagol i oni se markiraju crnom pozadinom u anotiranom tekstu. Za svaki element okvira se navodi definicija i jedan obeležen primer upotrebe.

Rizična_situacija

Definicija: Određena **Situacija** može da odvede (ali ne mora) do štetnog događaja koji je zadesio **Dragocenost**. **Situacija** može biti stanje, aktivnost ili određeni fokalni entitet za koji mora važi da je deo neke šire **Situacije** koja uključuje i entitet i **Dragocenost**. Za ovaj okvir važno je razumevanje ideje o štetnom događaju ali ne predstavlja argument leksičkih jedinica ovog okvira.

Da li su **klimatske promene** **OPASNE** **po čovečanstvo**?

Kupci se mogu žaliti na **NEBEZBEDNE** **proizvode**.

Najveća **OPASNOST PRETI** **našoj infrastrukturi**.

Ključni elementi:

Asset [ass]

Nešto što se smatra vrednim i poželjnim a postoji mogućnost da će mu šteta biti naneta.

Med nije **BEZBEDAN** **za bebe**.

Dangerous entity [dan ent]

Konkretan ili apstraktan entitet koji može naneti štetu ili dovesti do gubitka **Dragocenosti**

Isključuje: Situation

Taj čovek je **RIZIČAN**!

Situation [sit]

Situacija koja može dovesti do štetnog događaja.

Većina ljudi se slaže da je **NEBEZBEDNO** **voziti brže od 120 km/h**.

Sporedni elementi:

Circumstances []

Uslovi pod kojim je **Dragocenost** u opasnosti.

Degree [deg]

Modifikator koji ukazuje na odstupanje nivoa bezbednosti od očekivane vrednosti uzimajući u obzir **Situaciju**.

Odmetnute nacije su naša **najveća** **PRETNJA**.

Domain [dom]

Domen u kome **Dragocenost** ostvaruje određeni nivo bezbednosti.

Sve naše sajtove je **BEZBEDNO** koristiti u **obrazovanju dece**.

Frequency []

Koliko se često **Dragocenost** susreće sa Rizičnom_situacijom.

Place [pla]

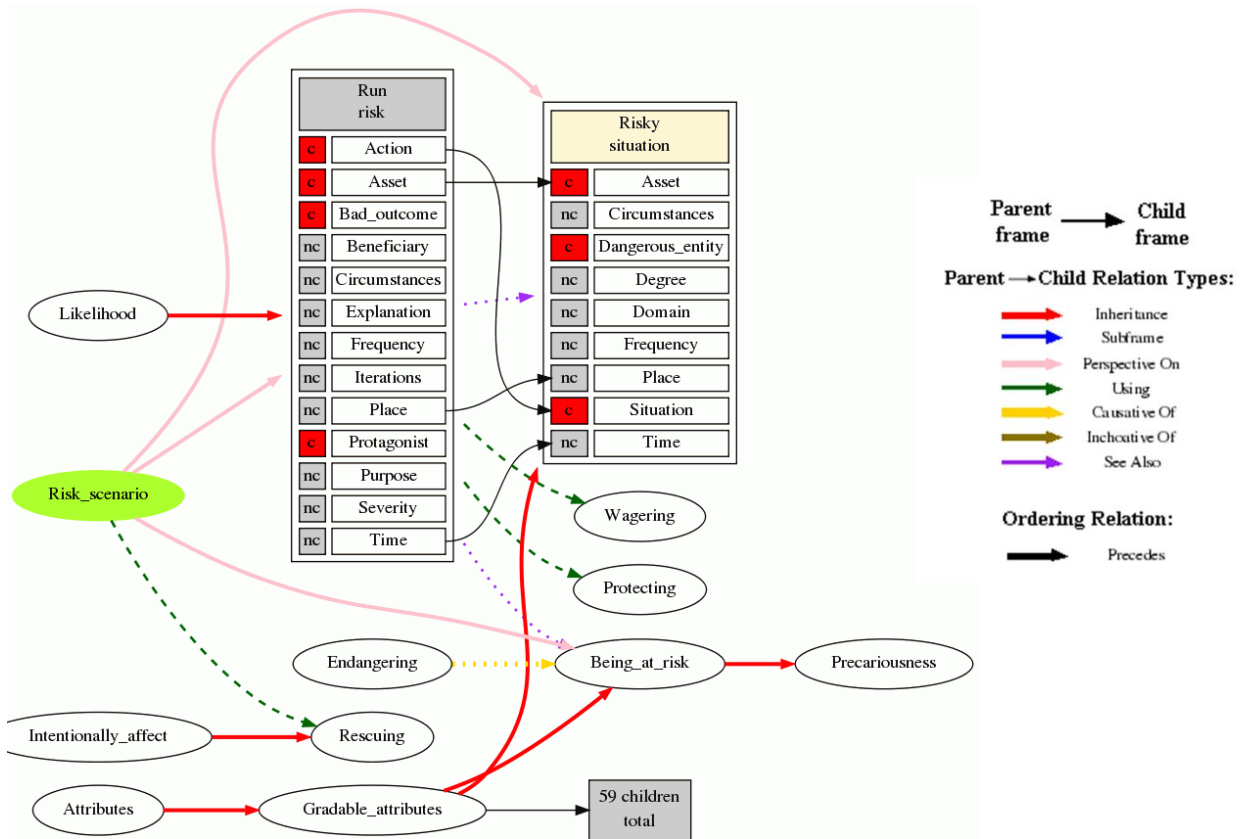
Određena lokacija na kojoj je **Situacija** bezbedna.

Često se može zaključiti da svojstva koja lokacija poseduje čine određene **Situacije** bezbednim ili nebezbednim.

Time [tim]

Vremenski period u kom **Situacija** ima iskazani nivo zaštite.

Slika 3-11 prikazuje okvir *Scenario_rizika* (eng. Risk_scenario) sa detaljnim prikazom okvira *Run_risk* i *Risky_situation*, gde su onda obeleženi ključni (eng. core, c) i sporedni (eng. noncore, nc) elementi. Na desnoj strani slike je data legenda sa vrstama relacija koje se uspostavljaju između okvira, na primer: nasleđivanje, podokvir, iz_perspektive, koristi, uzrokovan_od, i sl.



Slika 3-11 Semantički okvir Scenario_rizika (eng. Risk_scenario)

Okvir *Biti_u_opasnosti* (eng. Being_at_risk) prikazuje Slika 3-12, a realizuje se sledećim leksičkim jedinicama: *opasnost.N*, *nebezbedno.A*, *opasnost.N*, *bezbedno.A*, *sigurno.A*, *bezbednost.N*, *podložnost.N*, *biti podlozan.N*, *ugrožen.A*. Boje elemenata su kao u prethodnom primeru, uz dodatak elementa *Harmful_event*.

Definicija: **Dragocenost** je izložena **Štetnom događaju** dejstvom **ugrožavajućeg entiteta**. Reči koje iskazuju relativnu bezbednost (odsustvo rizika) takođe su deo ovog okvira.

Nijedno dete nije **VAN OPASNOSTI** da padne u iskušenje da uradi ono što njegovi vršnjaci rade.

Naša država užasno greši pokušavajući da se **ZAŠTITI od ljudi**- ona mora da štiti ljude.

Ukoliko ste bagerista, **vi** ste pod **RIZIKOM** da vam se sluh ošteti zbog izloženosti buci tokom rada.

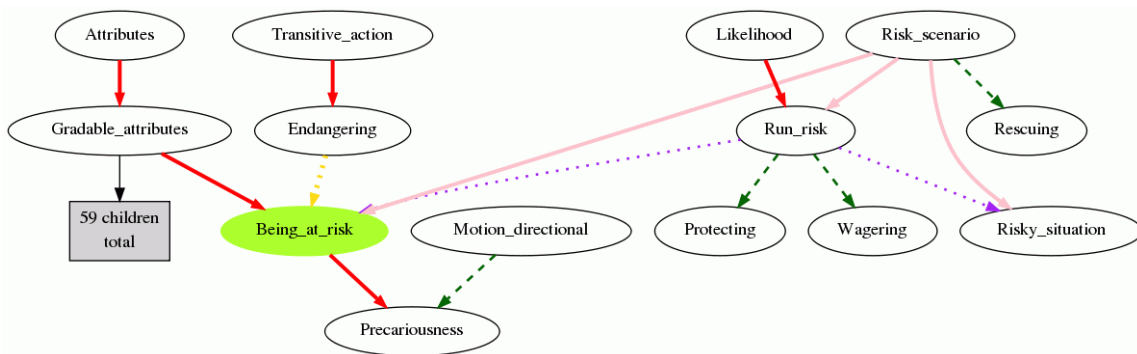
Vi nemate **SIGURNOST** zaštite privatnosti se mako imate sigurnost protiv prisluškivanja.

Harmful event [har]

Isključuje:

Dangerous_entity

Događaj koji se može desiti ili stanje koje se može promeniti na štetu po **dragocenost** ili da rezultira gubitkom **dragocenosti**. Obezbeđuje da su informacije koje se čuvaju na hardveru **ZAŠTIĆENE** od napada hakera kao i **pokušaja fizičke krađe**.



Slika 3-12 Semantički okvir Biti_u_opasnosti (eng. Being_at_risk)

Slika 3-13 prikazuje okvir Run_risk, za koji se vezuju sledeće leksičke jedinice: ugrožen.A, opasnost.N, rizik.N, rizikovati.V, ugroziti.V. U nastavku dajemo definiciju, primere i karakteristične elemente ovog okvira:

Definicija: **Protagonista** je u potencijalno opasnoj situaciji koja može da se okonča **Lošim ishodom** po njega ili nju. **Protagonisti** može da preti **Loš ishod** ili gubitak svoje **Dragocenost**. Ne postoje naznake da se **Protagonista** namerno izlaže rizičnoj situaciji. Moguće je da **Protagonista** pokušava da ostvari neki **Cilj** što ga stavlja u opasnu situaciju. **Ozbiljnost** rizika kojem se izlaže takođe može biti izražena. Postojao je **RIZIK** da se sve zapali.

Asset

Nešto što **Protagonista** poseduje ili nešto što ga karakteriše i smatra ga poželjnim a postoji opasnost da ga može izgubiti.

Iključuje:

Bad_outcome

RIZIKOVAO je **svoju reputaciju** zbog nje.

Protagonist

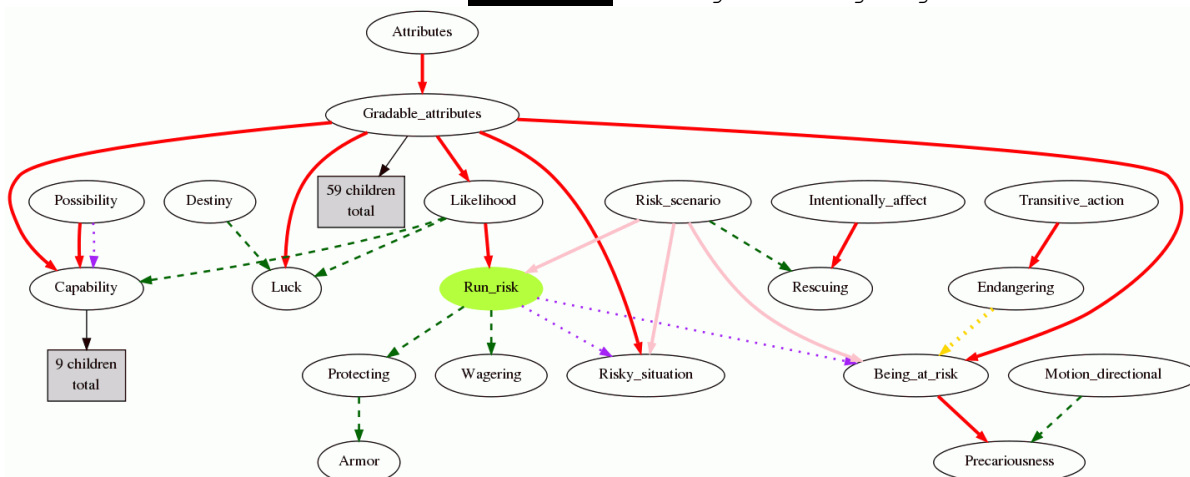
Osoba kojoj preti **Loš ishod**.

Bad outcome

Situacija koju bi **Protagonista** hteo da izbegne. **RIZIKOVAO** je da izgubi **svoje zdravlje**.

Severity

Verovatnoća da će se nešto loše desiti **Protagonisti**. Kada više ne raspoznaješ dobro od zla u **ozbiljnoj** si **OPASNOSTI** da izgubiš svoju ljudskost.



Slika 3-13 Semantički okvir Rizikovati (eng. Run_risk)

3.4. Izgradnja ontologije

U informacionim i komunikacionim tehnologijama razmena znanja postala je kritičnija i izazovnija, naročito za organizacije kod kojih su digitalne tehnologije i znanje važne komponente poslovanja. Različite aktivnosti se mogu realizovati u domenu upravljanja znanjem i deljenjem informacija (Shu-hsien 2003). Pristup razvoju ontologija se kretao od ručnog unosa instanci klasa, da bi se kasnije razvile tehnike za automatsko popunjavanje. Iskustva za druge jezike pokazuju da je moguće ubrzati razvoj ontologija alatima za ontološko učenje (eng. ontology learning environment) oslanjajući se na obradu narativnog, nestruktuiranog teksta tehnikama obrade prirodnih jezika (eng. NLP, natural language processing), uglavnom ekstrakcijom informacija (konceptata, primeraka i relacija) koje se koriste i u ovom istraživanju.

3.4.1. Principi razvoja ontologije

Izgradnja ontologija se odvija kroz više razvojnih koraka-(Noy, McGuinness, and others 2001; Yadav et al. 2016):

- identifikacija opsega: - ontologija je model baze znanja koji se sastoji od klasa, svojstava i njihovih instanci vezanih za određeni domen, tako da je neophodno prvo odrediti obim datog problema,
- opisivanje resursa: - navođenje alata i editora koji se koriste,
- definisanje taksonomije: - nakon identifikacije obima i resursa, koncepti se organizuju u hijerarhijski model, tj. taksonomiju,
- definisanje svojstava: - svojstva se navode u vidu izraza.
- definisanje aspekta: - podrazumeva upotrebu RDF šeme (rdfs) za predstavljanje,
- definisanje instanci: - za klase se definišu njihovi primerci (recimo za klasu *Rang_tretmana* primerci su: *prihvata_se*, *prati_se*, *eliminiše_se_u_potpunosti* i *smanjuje_se_na_nivo_prihvatljivog*),
- provera višeznačnosti: - ontologiju treba verifikovati tako da pokriva opseg datog domena i da ujedno nema nepreciznosti i nedostataka.

Prilikom razvoja ontologije treba pratiti pet osnovnih principa razvoja (Tomašević 2018):

- proširivost, koja treba da omogući izgradnju ontologije specijalizovanog domena nad opštim, prethodno razvijenim ontologijama,
- nezavisnost od aplikacija, što znači da struktura i sadržaj ontologije treba da se oslanja na znanje konkretne discipline, a ne na način kako se koristi domensko znanje,
- nezavisnost od prirodnog jezika, što znači da se predstavljaju koncepti, a ne termini. Koncepti ostaju isti bez obzira na dijalekat, tehnički žargon, strane jezike, i slično. Sinonimi, kao reči sa istim ili sličnim značenjem se mogu mapirati u okviru ontologije korišćenjem relacije „sameAs“,
- ortogonalnost, koji govori da složeni koncepti treba da se raščlane na svoje sastavne delove, i tako da olakšaju ponovno udruživanje konceptata na nov način.
- grupna raspodela rada, odnosno učešće zajednice dobrovoljaca koja treba da razvija i koristi ontologiju, jer ona ima smisla samo ako je koristi više aplikacija (i timova).

Formalno kodirani koncepti modeliraju se u ontologiji korišćenjem jezika OWL (eng. Web Ontology language). Ukoliko postoji mogućnost da se tokom izgradnje ontologije, traže i mapiraju koncepti sa dostupnim međunarodno prihvaćenim ontologijama (rečnicima, standardima i klasifikacijama). Odnosi između klasa i pravila se kodiraju. Pravila se izvode iz domenskog znanja i analize, a predstavljaju se jezikom semantičkih veb pravila SWRL (eng. Semantic Web Rule Language).

Razvoj sistema zasnovanog na znanju čini pet faza: identifikacija (definicija problema, učesnici i ciljevi), konceptualizacija (definicija ključnih koncepata), formalizacija, primena i testiranje, nakon čega sledi revizija u kojoj se sistem redizajnira i ponovo primenjuje u odnosu na dobijene rezultate iz postupaka ispitivanja. Procesi takozvanog inženjerstva ontologije (eng. ontology engineering) uključuju specifikaciju i konceptualizaciju, formalizaciju, implementaciju i održavanje ontologija. OWL kao standard koji preporučuje W3C je najčešće korišćeni jezik koji pruža interoperabilnost i moćan mehanizam zaključivanja. Osim predstavljanja koncepata klasama i primeraka instancama, mogu se definisati tipska svojstva, svojstva objekata i skup aksioma za dodatnu specifikaciju ograničenja i mogućnosti klasa (O'Connor et al. 2008). OWL se zasniva na deskriptivnoj logici, gde domen predstavljaju objekti i relacije sa njihovim mogućim stanjima, tako da omogućava pronalaženje instanci (Zhao and Liu 2008). Dekriptivna logika podrazumeva semantičku mrežu i konceptualni graf, kao i jezike za predstavljanje znanja sa formalnom sintaksom i pravilima zaključivanja za izvođenje semantički validnih rezultata. Jezik semantičkih veb pravila (SWRL) je dodatak jeziku OWL, koji se koristi za definisanje pravila i rezonovanje nad instancama ontologije (Zhao and Liu 2008). Korišćenjem SWRL moguće je definisati Hornova logička pravila⁵¹, koja imaju dva dela, naime premisu kao telo pravila i posledicu kao glavu. Korišćenje alata otvorenog koda Protégé za predstavljanje taksonomije domenskog znanja i odnosa između njih, podržan aksiomama i pravilima koriste i dalje unapređuju brojni istraživači (Shu-hsien 2003; Knublauch et al. 2004).

Poslednjih godina agilni pristup razvoju softvera je inspirisao istraživače da koriste agilnu metodologiju za razvoj ontologije, odnosno da usvoje agilne principe i praksu u razvoju ontologije. Razvoj platforme podrazumeva da se različite ontološke aktivnosti pretoče u faze Scrum agilne metodologije, koja se odvija kroz faze: planiranje, analiza zahteva, projektovanje, implementacija, testiranje, isporuka. Scrum je jedan od najčešće korišćenih načina za implementaciju agilne metodologije, koji se sastoji od definisanja niza uloga, odgovornosti i sastanaka (sprintevi) (Abdelghany, Darwish, and Hefni 2019) .

Razvoj sistema zasnovanog na znanju na Rudarskom fakultetu Univerziteta u Beogradu je započeo pre dve decenije. Jedan od prvih sistema bio je hibridni sistem INVENTS sa bazom znanja koja predstavlja formalizaciju stručnog znanja u polju rudarske ventilacije (Lilic, Obradovic, and Stankovic 2002). Ontologija rudarske opreme uspostavila je taksonomiju tipova opreme, njihove konstrukcione karakteristike i tehnološke parametre (Stanković et al. 2012). Razvoj terminoloških resursa sa bazom znanja je demonstriran na studiji slučaja u rudarstvu kojom je pokazana mogućnost obrade polustrukturiranih podataka (Kolonja et al. 2016). Novija istraživanja usredsređena su na korišćenje tehnologija prirodnog jezika: u poboljšanju pretraživanja digitalnog rečnika pomoću ontologije izvedene iz semantičke mreže reči WordNet (Mladenović, Stanković, and Krstev 2017), potom za upravljanje projektnom dokumentacijom rudarstva (Tomašević et al. 2018) i u medicinskom domenu za klasifikaciju dokumenata izvlačenjem taksonomskih koncepata iz MeSH Ontologije (Škorić and Dragoni 2019).

3.4.2. Automatizacija popunjavanja ontologije

Tradicionalan pristup razvoja ontologije podrazumeva da se pojmovi (koncepti), relacije između njih i instance ručno popunjavaju u ontologiji, što zahteva mnogo vremena, skup je i složen zadatak (Makki, Alquier, and Prince 2008b). Postojali su različiti pristupi za (polu) automatsko učenje i dopunjavanje ontologija, koje se sastoji u ekstrakciji pojmova i njihovih odnosa, unoseći ih obično kao primerke (instance) ontologije. Problemi koji se postavljaju su: razumevanje specifične semantike jezika, ekstrakcija terminologije konkretnog domena,

⁵¹ <https://mathworld.wolfram.com/HornClause.html>

razumevanje implicitnog znanje u tekstovima i nedostatak domenskih eksperata. Imajući u vidu da je rizik u rudarstvu osetljiv domen, popunjavanje ontologije se ne može osloniti na potpuno automatske metode i zahteva validaciju od strane eksperata. Makki sa kolegama (2008) predlaže poluautomatsko popunjavanje ontologije za upravljanje rizikom za domen rudarstva, zasnovan na tehnikama obrade prirodnog jezika kojim se dobijaju novi primerci koncepata iz različitih rudarskih tekstova (Makki, Alquier, and Prince 2008a).

Tseng je sa kolegama (2009) razvio platformu za upravljanje rizicima koja je podržana projektnom ontologijom rizika da ciljem da poboljšaju proces upravljanja rizikom u građevinskoj industriji oslanjajući se na studije slučaja. Krajnji cilj je smanjenje uticaja rizika na građevinski projekat u celosti. Algoritam za pronalaženje informacija je osnova razvijenog alata za dinamičku ekstrakciju termina koji se koriste za dopunu i ažuriranje ontologije (Tseng et al. 2009). Razvijeni modeli za pronalaženje informacija su računali relevantnost dokumenata, potom su sažimali dokumenta tako što su izdvajali važne izraze kao skup semantički relevantnih ključnih reči. Dokumenti su bili predstavljeni kao skupovi semantički relevantnih ključnih reči koje se koriste za proračun relevantnosti dokumenta. Relevantnost dokumenata je određivana vektorskim modelom, gde se dva dokumenta koja se porede, predstavljaju sa dva pojedinačna vektora u virtuelnom prostoru. Rezultati njihovog istraživanja ukazuju da se relacija između rizika može se utvrditi razvijem vektorskim modelom. Praćenjem i nadgledanjem izvršavanja algoritma za ekstrakciju informacija, kao dodatak su razvili alat za ekstrakciju ontologije koje je podržao dinamičko ažuriranje i dopunu ontologije. U ovom vektorskom modelu, sličnost dve ključne reči, koje predstavljaju dva koncepta rizika, se određivalo rastojanjem vektora ključnih reči.

Sličnost dve reči ili dve fraze se može izračunati kosinusom vrednosti dva vektora kojima su predstavljanje ključne reči: kw^i i kw^j , kao u jednačinama koje slede (Salton 1989), kojima se određuje težinski koeficijent, odnosno ponder svake ključne reči i . Štaviše, koeficijent 0,5 je korišćen za normalizaciju frekvencije termina, odnosno ključnih reči koju su predložili Salton i Buckley (Salton and Buckley 1988).

$$W_{dj,kw} = (0.5 + 0.5 \times tf_{dj,kw}) \times idf_{kw} = (0.5 + 0.5 \times \frac{fre_{kw}}{\max fre}) \times \log \frac{N}{n_{kw}}$$

$$tf_{dj,kw} = \frac{fre_{kw}}{\max fre}, \quad idf_{kw} = \log \frac{N}{n_{kw}},$$

$$similarity (\overrightarrow{kw^j} \times \overrightarrow{kw^k}) = cosine (\overrightarrow{kw^j} \times \overrightarrow{kw^k}) = \frac{\sum_{i=1}^n W_{di,kwj} \times W_{di,kwk}}{\sqrt{\sum_{i=1}^n W_{di,kwj}^2} \times \sqrt{\sum_{i=1}^n W_{di,kwk}^2}},$$

gde je $w_{di,kw}$ težinski koeficijent ključne reči kw u dokumentu di ; fre_{kw} je frekvencija svake ključne reči; $\max fre$ je maksimalna frekvencija ključne reči kw u dokumentu di ; $\sum_{i=1}^S fre_{si}$ je ukupna frekvencija svih ključnih reči u svakom dokumentu; N je ukupan broj dokumenata u bazi znanja; n_{kw} je ukupan broj dokumenata u kojima se ključna reč pojavljuje.

Alat za ekstrakciju ontologije dinamički izvlači važne koncepte rizika i na taj način dopunjava ontologiju projektnog rizika. Obrada teksta je podrazumevala: čišćenje dokumenata (uklanjanje stop reči, prefiksa i sufiksa); ekstrakciju ključnih reči; i izbor ključnih reči koje jesu koncepti domena rizika. Da bi se obezbedila semantička normalizacija teksta, sinonimi se u analiziranom tekstu (odnosno indeksu teksta) zamenjuju pojmom sa istim ili sličnim značenjem na osnovu odgovarajućeg tezaurusa. Evaluacija je pokazala da je procenat validnih ekstrahovanih termina bio uglavnom veći od 80% za niske (0,7) i visoke (0,9) granične vrednosti mere sličnosti, tako da je testirani prototip ocenjen kao uspešan alat za ekstrakciju ontologije.

Todirascu i saradnici (2002) kombinuju sintaksu i različite ontologije za identifikovanje novih elemenata ontologije koji su specifični za domen, koristeći parcijalnu sintaksičku analizu i domensko znanje predstavljeno deskriptivnom logikom. Metoda koju su predložili koristi

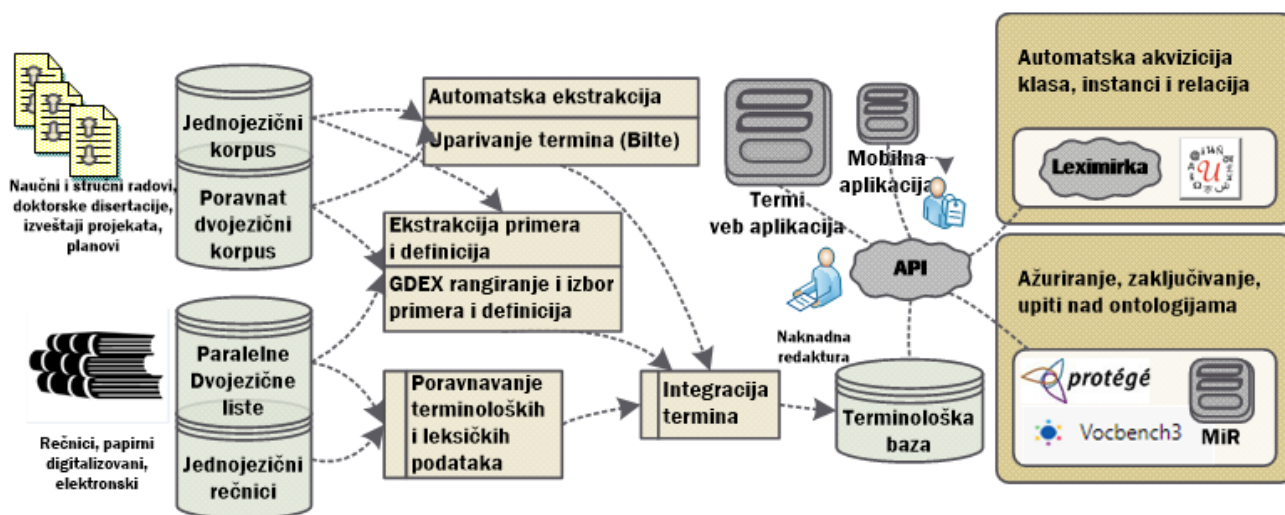
delimične sintaksičke strukture koje obezbeđuje LTAG parser (Lexical Tree Adjoining Grammars) za identifikovanje koncepata kandidata, kao i mehanizam zaključivanja (pravila) deskriptivne logike za predstavljanje znanja. Predložena metodologija je uključila nekoliko koraka kako bi se identifikovali ključni elementi ontologije: primerci (instance) koncepta, odnosi između koncepata i odnosi između primeraka i koncepata. Primerci pojmova su identifikovani delimičnim parsiranjem, dok relacije između koncepata i između instanci i koncepata zahtevaju resurse specifične za konkretan domen. Znanje domena je ključni element u sistemu ekstrakcije informacija, tako da domenske ontologije često sadrže specifična znanja koja se ne bi mogla naći u ontologiji opšte namene (Todorascu, Romary, and Bekhouche 2002).

Metode zasnovane na statistici i mehanizmu zaključivanja su u širokoj upotrebi za automatsku ekstrakciju ontologija: za identifikovanje kandidata pojmova u tekstovima, kreiranje klasa sličnih pojmova i utvrđivanje relacija između ovih klasa. Statističke metode grupišu pojmove sa sličnim kontekstom u klase povezane sa jedinstvenim konceptom (Daille 1994; Assadi and Bourigault 2000). Odnosi između različitih klasa identifikuju se tumačenjem sintaksičkih struktura (Capponi and Toussaint 2000) ili korišćenjem statističkih podataka (Bouaud et al. 2000). Statističke metode zahtevaju velike, stabilne korpuse za obuku da bi izdvojili klase, kao i eksperte za validaciju i tumačenje rezultata. Metode zasnovane na zaključivanju koriste poluautomatske metode za proveru valjanosti postojećeg znanja. Novi koncepti izvedeni pravilima zaključivanja dodaju se hijerarhiji domena ako su u skladu sa postojećim znanjem. Odnosi se identifikuju pomoću sintaksičkog znanja, kao potkategorizacija (Capponi and Toussaint 2000). Prekomerna ekstrakcija koncepta i troškovi provere nekoherentnosti i nedoslednosti ekstrahovanog znanja su glavni nedostaci pristupa.

4. MODEL SISTEMA ZA UPRAVLJANJE RIZIKOM

Zadatak sistema koji se razvija je da se napravi sistem za podršku upravljanju rizikom, koji će modelirati (prepoznavati) uzročno-posledične veze, događaje, bazirati se na pronalaženju pokretača ili uzroka. Previđeni sistem je fokusiran na izgradnju baze znanja koja treba da pomogne analitičarima i ostalim osobama koje se bave rizicima, naročito u fazi identifikacije rizika i u pronalaženju odgovarajućih rešenja.

Slika 3-2 prikazuje arhitekturu celokupnog sistema sa svim neophodnim komponentama za njegovu realizaciju. Korišćeni resursi, kao i komponente razvijene u okviru disertacije detaljno su predstavljene u okviru trećeg i četvrtog poglavlja, a ovde se daje samo kratki osvrt na ceo sistem. Proces pripreme potrebnih resursa započinje prikupljanjem relevantnih dokumenata, njihovom obradom i pripremom za jednojezični korpus i poravnavanjem englesko-srpskih tekstova za dvojezični paralelni korpus. Nad pripremljenim resursima vrši se ekstrakcija termina i deponovanje u terminološku bazu podataka. Ontološka baza se dodatno popunjava od strane pojedinačnih korisnika kroz veb aplikaciju MIR. Nad popunjenim podacima se primenjuju unapred definisani mehanizmi zaključivanja ili se kreiraju novi prema potrebama.



Slika 4-1 Arhitektura Sistema za ekstrakciju i publikovanje terminologije

4.1. Terminološki resursi i korpusi

U sklopu doktorske disertacije unapređeno je ranije predstavljeno softversko rešenje Termi. Implementirane su dodatne funkcionalnosti koje su bile potrebne za ovo istraživanje, od kojih su najznačajnija: povezivanje sa rudarskim korpusom i razvijenom ontologijom, kao i izvoz podataka u rdf formatu. Uz pomenuti razvoj terminološke platforme, sama terminološka baza značajno je dopunjena terminima za oblast rudarstva i rizika, što se izlaže u ovom odeljku.

4.1.1. Arhitektura sistema za ekstrakciju ontoterminologije

Metodologija ekstrakcija informacija obuhvata nekoliko celina: kompilaciju specijalizovanih korpusa, identifikaciju termina (polu-automatski), ekstrakciju pojmova, definisanje strukture argumenata, anotaciju konteksta, definisanje semantičnih okvira, kodiranje frejmova, odnosno okvira i definisanje odnosa između okvira. Deo arhitekture sistema za ekstrakciju i publikovanje terminologije prikazuje Slika 4-1.

Ontoterminologija⁵² je terminologija čiji je konceptualni sistem formalna ontologija. Ontoterminologija eksplicitno čini dvostruku dimenziju terminologije (lingvističku i konceptualnu). Ova nova paradigma čuva jezičku raznolikost uzimajući u obzir višejezičnost i omogućava standardizaciju domenskog znanja (Roche 2012).

Polazni resurs je korpus iz rudarskog domena, podkorpus iz domena rizika i korpus savremenog srpskog jezika, koji su opisani u odeljku 3.2.2. Na osnovu razlika frekvencija između korpusa savremenog srpskog i korpusa RudKorp i posebno podkorpusa Rizik kao njegove particije, sastavljen je izbor pojmova, što je uključilo i objedinjavanje njihovih morfoloških i leksičkih varijanti koristeći aplikaciju Leximir (Stanković, Krstev, Obradović, et al. 2016). Nepotrebno je isticati da su ekstrahovane ključne reči i takozvane proste (jednočlane) reči (eng. simple words⁵³), ali i fraze u vidu polileksičkih jedinica.

Naročito izazovan deo istraživanja je svakako identifikacija termina sa identičnom strukturom argumenata:

- isti tip argumenata (oprema za transport, otkopavanje, utovar,...) ,
- isti broj argumenata,
- iste semantičke uloge (kvar, neispravnost, zastoj, nedostatak kvaliteta ili proizvodnje...),
- slični dodaci odnosno dopune.

Leksički sadržaj okvira nastaje na osnovu korpusa i koristeći distribuciju reči, tako što se analizira kontekst oko reči i pronalaze leksičke jedinice koje imaju sličan kontekst. Pretpostavka je da ako ove jedinice dele sličan kontekst, verovatno prenose slična značenja ili dele važne semantičke komponente. U kontekstu semantike okvira, pretpostavka je da ako ove jedinice dele veliki broj semantičkih komponenti prizivaju isti okvir ili da prizivaju okvire koji su povezani sa tezaurusom ili ontologijom. Relacije se ručno kodiraju koristeći leksičke funkcije (Mel'cuk 1995; Polguère 2014), uzimajući u obzir do tri različita jezička svojstva: sintaksičku strukturu kolokacije, opšte i apstraktno značenje relacije i, konačno, vezu između srodnog pojma i strukture argumenata ključne reči. Tehnički, leksičke funkcije se primenjuju na ključnu reč i daju jednu ili uži spisak vrednosti, kao što je dato u primeru u sledećoj tabeli.

Razvoj alata za dinamičku ekstrakciju ontologije razvijan u okviru ove disertacije, kroz koji će se ontologija povremeno dopunjavati i ažurirati se može razložiti na tri zadatka:

- određivanje osnovnih koncepata, na primer: rizik, događaj rizika, faktor rizika, uzrok rizika, uticaj (posledice), mere za otklanjanje rizika, hazard (opasnost),
- razvoj taksonomija za svaki od osnovnih koncepata,
- prepoznavanje relacija među konceptima.

Poluautomatski pristup podrazumeva da osim ekspertskeg znanja koristimo pristup zasnovan na podacima (eng. data-driven approach) koji će vršiti ekstrakciju informacija iz korpusa tekstova, iz tabelarnih podataka, izveštaja, intervjuja, iz ISO procedura i sličnih dokumenata. Za vršne (eng. top) koncepte su važne i druge ontologije, konkretno povezivanje sa drugim ontologijama i bazama znanja. U navedenim pasusima ukratko su opisani resursi i alati neophodni za ekstrakciju elemenata za izgradnju ontologije rizika.

⁵² <http://ontoterminology.com/>

⁵³ <https://unitexgramlab.org/releases/3.1/man/Unitex-GramLab-3.1-usermanual-en.pdf>

Tabela 4-1 Primeri leksičkih funkcija

Relacija	Primer	Leksička funkcija
Isto značenje		
Tačna sinonimija, varijante, simboli, skraćenice	povlata → krovina; ugljen dioksid → CO ₂	Syn Symb
Srodna značenja		
Skoro sinonimi	hazard → opasnost	QSyn
Šire značenje	rotorni bager → bager	Gener
Antonimi	opasan → bezbedan održiv → neodrživ	Anti1, Anti2, Rev1, Rev2
Kontrastivnost	ruda → jalovina	Contr
Konverzivnost	pogon → proizvodnja	Convij
Porodice reči		
Isto značenje, različita vrsta reči	napušten → napuštanje otkop → otkopavanje	A0, S0, V0, Adv0
Pridev sa dodatim značenjem	zagađen → zagaditi; zagađen → zagađenje	A1, A2, Able1, Able2, i sl.
Jezička realizacija argumenata		
Oznaka uloge (npr. Agent, pacijent)	kvar → (Patient) oprema, postrojenje, bager	Oznake uloga se navode u posebnoj listi uloga
Tip		
Intenzifikacija	rizik → visok ~; udar → jak ~	Magn
Prema lokaciji	eksploatacija → podzemna ~	Hypo - Lieu
Kombinacije		
Tipična upotreba	deponija → odlagati na ~	Reali, Facti, Labreal12
Postojanje	nesreća → preživeti ~	Funci
Stvaranje	zemljište → otkup ~	Causi Func0
Drugi		
Meronimi	etaža → kop	[Part], [Tot]. Mult, Sing
Količina	gas → koncentracija gasa	Quant

Glavni cilj pristupa je objedinjavanje i povezivanje svih raspoloživih pojmova u domenu rudarstva i rizika u povezanu bazu znanja, koju uz terminološku bazu podataka Termini čine i povezani podaci dostupni preko SPARQL pristupne tačke, sistem pravila i konačnih automata razvijen u alatu Unitex, ontologija u alatu Protégé. Pored cilja objedinjavanja pojmova iz različitih izvora, jedan od razloga za harmonizaciju pojmova iz više rečnika bio je i procena

upotrebe konkretnih termina u rudarskoj praksi i leksikalizacija rudarskih pojmova na odgovarajući način. S druge strane, poravnanje pojmova sa SrpMD bilo je neophodno, jer su ovi rečnici osnovni resurs za lematizaciju i ekstrakciju informacija. Budući da su SrpMD već u leksičkoj bazi podataka Leximirka (Stanković et al. 2018), koju je razvio i kojom upravlja isti istraživački tim, ova vrsta poravnanja je bila moguća. Jedna od dodatnih aktivnosti je bilo uspostavljanje dvojezične terminologije, za čije je poravnanje korišćena aplikacija Bilte⁵⁴ (Šandrih, Krstev, and Stanković 2020).

Ekstrahovani pojmovi su obeleženi markerima uvedenim za specifikaciju domena termina (rudarstvo i druge nauke), poddomena kao pojedinih grana rudarstva, potom semantičkih markera za obeležavanje materijala, procesa, aktivnosti, mehanizacije, alata i sl. Od 16.491 unosa dobijenih Rudarskog rečnika (Nešić 1970), 12.018 (73%) je ručno klasifikovano i dodeljeni su im markeri za domen i poddomen, kao i semantičke oznake. Preostalih 4473 (27%) neklasifikovanih unosa uključivalo je reči iz opšte leksike i neke retko korišćene izraze. Klasifikovani zapisi uglavnom su iz domene rudarstva, tačnije postoji 4793 (40%) unosa karakterističnih za različite grane rudarstva. Uključen je i osnovni rečnik iz srodnih domena, na primer, 2398 (20%) zapisa koji se odnose na geologiju, hidrogeologiju i geografiju, 860 (7%) zapisa koji se odnose na transport, mehaniku stena, snimanje, zaštitu životne sredine, bezbednost, izgradnju, transport i elektrotehniku, dok 3082 (26%) zapisa pripadaju opštoj tehničkoj terminologiji. Postoje i zapisi iz osnovne nauke, na primer, 885 (7%) pojmova koji se odnose na biologiju, hemiju, matematiku, informatiku i fiziku.

Među terminima iz rudarskog domena, one koji se odnose na određenu granu rudarstva konsultovani su eksperti iz domena rudarstva tako da je dobijeno, na primer termini koji se odnose na preradu mineralnih sirovina (251), transport (243) ili podzemna eksploatacija (469). Dodeljene su i dodatne semantičke oznake, na primer, materijal (699), uređaj (536), mašina (384), mineral (313), objekat (288), instrument (279) itd. Vrsta reči je dodeljena poluautomatski, gde je samo 40 termina označeno kao pridevi, 250 kao glagoli, a svi ostali kao imenice.

Važan deo svake baze znanja predstavljaju naučne i leksikografski ispravne definicije. Na srpskom jeziku je vrlo malo terminoloških rečnika koji ih sadrže, već se većina objavljenih srpskih terminoloških rečnika ograničava na preciziranja prevodilačkih ekvivalenata. U okviru istraživanja na ovoj tezi je otpočeta dugoročna aktivnost na pronalaženju i adaptaciji rečenica kandidata za definicije. Prvi izvor je svakako rudarski korpus na srpskom jeziku, a potom rečnici na engleskom koji sadrže definicije termina. Prilagođavanje engleskih definicija srpskom jeziku i domaćoj rudarskoj praksi se radi u fazi naknadne obrade ili redakture. Definicije su najpotpunije date u američkom rudarskom rečniku (Graham 1996), a u predloženom pristupu se daje prioritet u obradi prvo najčešće korišćenim terminima, kako u korpusima, tako i u rečnicima.

Konačno, ekstrahovani termini su usklađeni i usaglašeni sa mikrostrukturom baze podataka Termini, koja osim samo naziva termina sadrži i sinonime, skraćenice, definicije, za svaki jezik, potom bibliografski izvor, a postoji mogućnost uključivanja slika i ilustracija i drugog spoljnog sadržaja.

Termini su organizovani u hijerarhijsku strukturu, kojom se određuje jed vid rudarske taksonomije, a predviđen je unos i dodatnih relacija između pojmova. Aktuelna taksonomija je urađena automatski koristeći informacije o domenu, poddomenu i semantičkim markerima, što je podložno reorganiziranju u fazi naknadne obrade. Slika 4-2 prikazuje panel sa hijerarhijom unetih rudarskih termina na levoj strani, a na desnoj je deo za unos, gde se može videti kako je definisan pojam „podzemna eksploatacija“. Panel pruža mogućnost uvida korišćenja termina u

⁵⁴ <http://bilte.jerteh.rs/>

kontekstu kroz pretraživanje digitalne biblioteke Bibliša i Rudarskog korpusa RudKorp i to uz mogućnost povlačenja konkordanci ili frekvencija oblika. Za detaljan prikaz termina, takozvanu ličnu kartu termina na raspolaganju je dugme „Detaljno“. Editor može da doda bibliografsku referencu iz koje je preuzeta ili prilagođena definicija pojma. Vidljivost svakog termina zavisi od statusa termina: u radu, verifikovan, publikovan, poništen.

The screenshot shows the 'Ažuriranje' (Update) interface for the term 'podzemna eksploatacija' (underground mining) in the RudKorp database. The interface is divided into several sections:

- Left Panel (Ažuriranje):** A hierarchical tree structure showing the term's location within the 'rudarstvo' (mining) domain. The selected term is 'podzemna eksploatacija'.
- Main Editing Area:**
 - Naziv:** podzemna eksploatacija
 - Sinonimi:** (empty field)
 - Skracenica:** (empty field)
 - Definicija:** Metode izvođenja rudarskih radova na otvaranju, razradi, pripremi, otkopavanju, bušenju i miniranju, izradi podzemnih prostorija, transportu, izvozu, provetranju u podzemnim prostorijama i na otkopima, odvodnjavanju uz pridržavanje mera bezbednosti i zdravlja na radu i mera zaštite radne i životne sredine.
 - Embed:** Bibliša, NaRA, Detaljno
 - Pretraži korpus:** RudKorp, Plain lemma, Konkordance
 - Frekvencije oblika:** (empty field)
 - Komentari:** (empty field)
 - Prevod:** English, German, Français
 - ID:** 115900
 - Language:** EN
 - State:** Verifikovan
 - Title:** underground mining
 - Synonyms:** (empty field)
 - Definition:** Method of performing mining works on opening, developing, preparing, excavating, drilling and blasting, construction of underground premises, transport, haulage, ventilation in underground premises and on excavations, and drainage, while adhering to safety and health measures and environmental protection measures.
 - Bibliografija:** (empty field)
- Bibliography Table:**

Naslov	Autor	Izdavač	Godina izdanja
Zakon o rudarstvu i geološkim istraživanjima	Republika Srbija - Ministarstvo rudarstva i energetike	Službeni glasnik RS	2015

Slika 4-2 Terminologija iz oblasti rudarstva

Osim rudarskog domena, terminološki je pokriven i domen rizika, osnovnim konceptima na srpskom i engleskom, gde su definicije usaglašene sa domaćom praksom i ISO procedurama. Slika 4-3 prikazuje panel sa hijerarhijom termina iz oblasti rizika na levoj strani, a na desnoj je panel za unos.

The screenshot shows a web application for a terminology database. The browser address bar displays <https://termi.rgf.bg.ac.rs/Azuriranje/Azuriranje/Azuriranje?korenID=115997>. The navigation menu includes 'Prelistavanje', 'Pretraga', 'Ažuriranje', 'Bibliografija', 'Import iz excela', 'Profili', 'Pomoć', 'Korisnik', and 'Odjava'. The main content area is titled 'Srpski' and shows details for the term 'deljenje rizika' (ID: 116036). The status is 'Verifikovan'. The definition is: 'Način postupanja sa rizikom koji obuhvata dogovorenu podelu rizika sa drugim stranama. Pravni ili regulatorni zahtevi mogu ograničiti, zabraniti ili dodeliti ovlašćenje za deljenje rizika. Deljenje rizika može se vršiti putem osiguranja ili drugih vidova ugovora. Do koje mere će rizik biti raspodeljen može zavistiti od pouzdanosti i jasnoće sporazuma za deljenje rizika. Prenošenje rizika je oblik podele rizika.' Below the definition, there are sections for 'Komentari', 'Prevod', and 'Bibliografija'. The search bar contains 'matematika' and 'Plain lemma'.

Slika 4-3 Terminologija iz oblasti rizika

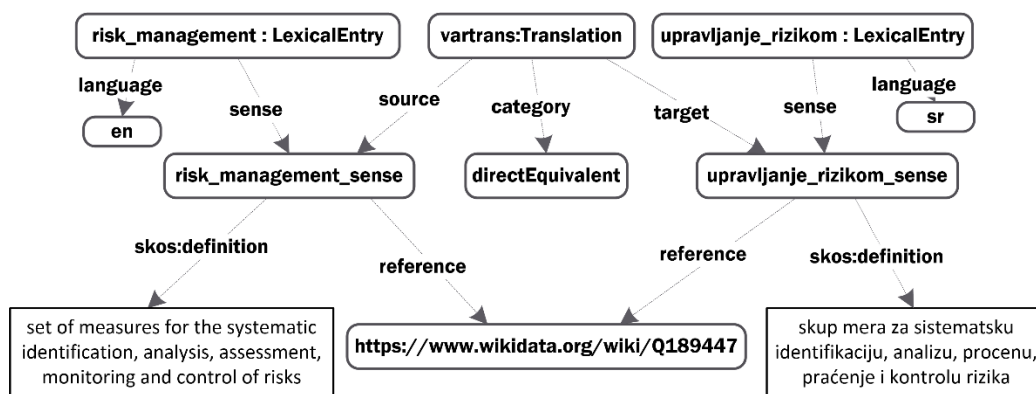
Integracija informacija koja prevazilazi nivo pojedinačnih rečnika i individualnih korisnika jezičkih resursa postao je važan cilj, nezavisno od grane privrede, nauke ili obrazovanja. Tehnologija koja, za sada na najbolji način omogućava postizanje ovog cilja je usvajanje paradigme povezanih podataka, Linked Data⁵⁵, tehnologije koja omogućava publikovanje semantički obogaćenog i međusobno povezanog sadržaja na webu u standardnom, mašinski čitljivom i razumljivom formatu, poznatijem kao RDF. W3C -standardi, u prvom redu RDF daju preporuke kako se koriste URI⁵⁶-je za jednoznačno identifikovanje resursa na webu, njihovih komponenti i njihovih relacija u mreži podataka. Deo neba otvorenih podataka koji se naziva Linked Open Data (LOD), predstavlja sadržaje koji su dostupni za korišćenje i distribuciju bez

⁵⁵ <https://www.w3.org/standards/semanticweb/data>

⁵⁶ Uniform Resource Identifier (URI) je jedinstveni niz karaktera koji identifikuje logički ili fizički resurs koji se koriste u veb tehnologijama.

pravnih, tehnoloških ili društvenih prepreka koristeći otvorene standarde. Linked Open Data je danas sve popularniji termin koji predstavlja otvorene podatke u RDF formatu, bez obzira na domen iz kog potiču isti (državna uprava ili poslovni sektor), tako da se i Srbija uključila u ovaj vid povezivanja resursa.⁵⁷

U istraživanju vezanom za ovu disertaciju takođe smo težili da publikujemo resurse saglasno tehnologiji povezanih podataka, koristeći njegove standarde modeliranja i principe deljenja mašinski čitljivih međusobno povezanih podataka na webu. Ova vizija globalno dostupnih i povezanih podataka na webu, zasnovana na RDF standardima semantičke mreže, je inicirala implementaciju modula za RDF sterilizaciju podataka o rudarskim terminima i terminima rizika. U tom cilju, aplikacija Termini je obezbedila izvoz podataka u RDF koji je u skladu sa OntoLex Lemon Lexicography Module⁵⁸ (Bosque-Gil et al. 2016; Bosque-Gil, Gracia, and Montiel-Ponsoda 2017) kao proširenje standarda Lexicon Model for Ontologies (lemon)⁵⁹ (J. McCrae, Spohr, and Cimiano 2011; J. McCrae et al. 2012). Slika 4-4 daje primer RDF izvoza za kojim sledi Turtle RDF Syntax⁶⁰ koja ilustruje upotrebu modela.



Slika 4-4 Primer RDF serijalizacije termina iz oblasti rizika

Napomenimo samo važnost povezivanja koncepata u različitim resursima, ako je ovde, osim veze ka vikipodacima o kojima je ranije bilo reči, ovde uspostavljena veza i sa oficijelnim rečnicima Evropske Unije⁶¹ i RDF grfom koji predstavlja pojam „upravljanje rizikom“ na adresi http://eurovoc.europa.eu/c_406ad4cc.

```
:risk_management a ontolex:LexicalEntry;
dct:language <http://lexvo.org/id/iso639-1/en> ;
lexinfo:partOfSpeech lexinfo:noun;
ontolex:lexicalForm :risk_management-form;
ontolex:sense :risk_management_sense.
:risk_management-form a ontolex:Form;
ontolex:writtenRep "risk management"@en.
:risk_management_sense skos:definition "set of measures for the systematic
identification, analysis, assessment, monitoring and control of risks"@en;
ontolex:reference <https://dbpedia.org/page/Risk_management>;
ontolex:reference <https://www.wikidata.org/wiki/Q189447>;
```

⁵⁷ <https://linkeddata.rs/>

⁵⁸ <https://www.w3.org/2019/09/lexicog>

⁵⁹ <https://www.w3.org/2016/05/ontolex/>

⁶⁰ <https://www.w3.org/TR/turtle/>

⁶¹ <https://op.europa.eu/en/web/eu-vocabularies>

```

ontolex:reference <http://eurovoc.europa.eu/c_406ad4cc>.

:upravljanje_rizikom a ontolex:LexicalEntry;
dct:language <http://id.loc.gov/vocabulary/iso639-1/sr> ;
lexinfo:partOfSpeech lexinfo:noun;
ontolex:lexicalForm :upravljanje_rizikom-form;
ontolex:sense :upravljanje_rizikom_sense.
:upravljanje_rizikom-form a ontolex:Form;
ontolex:writtenRep " upravljanje rizikom "@sr.
:upravljanje_rizikom_sense skos:definition "skup mera za sistematsku
identifikaciju, analizu, procenu, praćenje i kontrolu rizika"@sr;
ontolex:reference <https://www.wikidata.org/wiki/Q189447>.

:trans_risk_management_sense-upravljanje_rizikom_sense a
vartrans:Translation;
vartrans:source :risk_management_sense;
vartrans:target :upravljanje_rizikom_sense;
vartrans:category
<http://purl.org/net/translation-categories#directEquivalent>.

```

4.1.2. Dopuna korpusa RudKorp

Tokom istraživanja vezanih za ovu tezu proširen je sa 63 dokumenta. Trenutna verzija sadrži 4,1 milion reči, što obuhvata projektnu dokumentaciju (26%), zakonodavstvo (11%), doktorske disertacije (31%), udžbenike i ostalu rudarsku literaturu (32%) (Kitanović et al. 2021).

Korpus je publikovan na tri platforme: Unitex⁶², SketchEngine⁶³ i NoSketch⁶⁴, jer imaju različite mogućnosti, tako je cilj bio da se prikupljeni tekstualni podaci iskoriste na najbolji mogući način. O korišćenju korpusa u sistemu Unitex se više informacija može naći odeljcima 3.2.4 i 3.2.5. Korpus na NoSketch platformi je povezan sa terminološkom bazom Termini i leksičkom platformom Leximirka, što je demonstrirano u odeljku 3.2.3., a u poglavlju 4 će se videti primena sva tri sistema u funkciji ekstrakcije podataka, informacija i znanja iz teksta.

Ključno pretprocesiranje proširenog korpusa je anotacija reči vrstom reči⁶⁵ i kanonskim oblikom reči, poznatom kao lematizacija. Označavanje vrste reči je obavljeno automatski, tako da je svaka korpusna reč dobila informaciju da li je imenica, glagol, pridev, veznik, ili neka druga vrsta reči saglasno gramatici srpskog jezika. Za anotaciju je korišćen servis za obradu teksta Društva za jezičke resurse i tehnologije Jerteh⁶⁶ (Stanković et al. 2020), koji omogućava različite vrste obrade teksta. Slika 4-5 prikazuje alat za anotaciju teksta gde se na ulazu otpremi tekst u vidu fajla ili direktno unese u tekstualno polje, a kao rezultat se dobije vertikalizovan fajl gde je svaka reč data u zasebnom redu i svakoj reči je pridružena vrsta reči i kanonski oblik reči, odnosno lema.

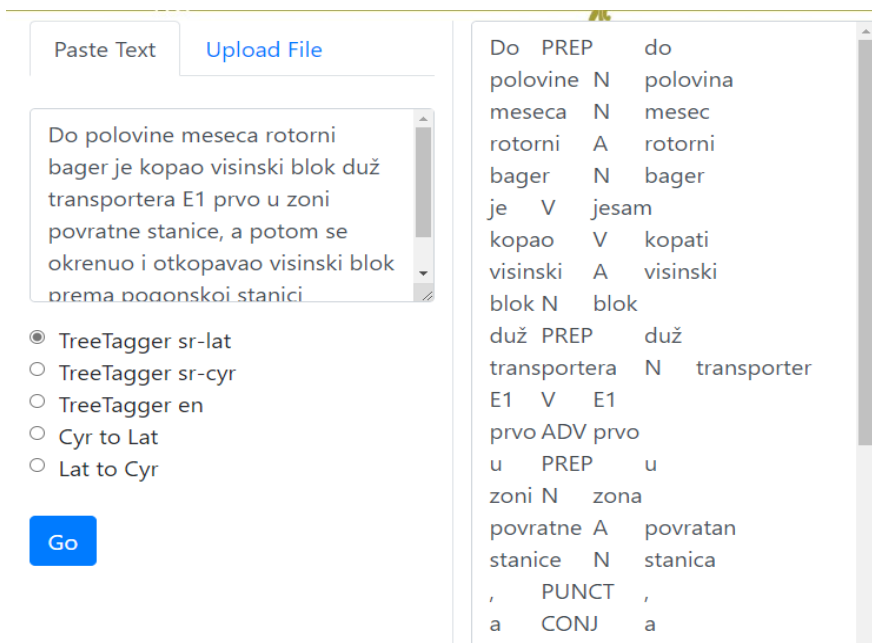
⁶² <https://unitexgramlab.org/>

⁶³ <https://www.sketchengine.eu/>

⁶⁴ <https://noske.jerteh.rs/>, <https://www.sketchengine.eu/nosketch-engine/>

⁶⁵ Deset vrsta reči se obeležavaju na sledeći način: N imenice, PRO zamenice, A pridevi, ADV prolozi, V glagoli, NUM brojevi, CONJ veznici, PREP predlozi, PAR rečice, INT uzvici

⁶⁶ <http://obrada.jerteh.rs/>



Slika 4-5 Veb alat za anotaciju teksta vrstom reči i lemom

Slika 4-6 prikazuje deo panela sa konkordancama⁶⁷ korpusa za upit „pridev iz koga sledi bilo koji gramatički oblik reči *rudnik*“. Ispod svake korpusne reči prikazane su anotacije: vrsta reči i gramatičke kategorije, potom i lema.

ičaja	španskog	rudarstva	ističući	da	6.000	aktivnih	rudnika	zapošljava	80.000	radnika	2,2	% o					
jznačaj	Agpnsy/španski	Nompr/rudarstvo	Zi, Rr/isticati	Cs/da	Mdo/[number]	Agmpgy/aktivan	Nompr/rudnik	Zi, Vmr3s/zapošljavati	Mdo/[number]	Nompr/radnik	Zi, Mdo/[number]	Z/% Sg					
istva	sertifikacijom	i	stepena	primene	TQM	u	anketiranim	rudnicima	[144]	.	</s><s>	Mišljenje	autora	je	da		
vojtstvo	Ncfs/serifikacija	CoI	Nomsg/stepen	Ncfs/primena	Npmsn/tqm	Si/lu	Agpmsy/anketirani	Nompr/rudnik	Zi[Npmsn/144]	Zi.		Nonsn/mišljenje	Nompr/autor	Var3s/bitl	Cs/da		
Nemačkoj	.	</s><s>	Kao	primer	navodi	se	najveći	rudnik	kamenog	uglja	u	Nemačkoj	.	</s><s>	U	1991	
ipfsl/nemačka	Zi.		Cs/kao	Nomsn/primer	Vmr3s/navoditi	Px--sa/sebe	Agmsayn/velik	Nomsn/rudnik	Agpnsy/kameni	Nomsg/ugalj	Si/lu	Npfsi/nemačka	Zi.		Sa/lu	Mdo/[nun	
imenog	uglja	u	Nemačkoj	.	</s><s>	U	1991.	nemački	rudnik	uglja	RAG	Aktiengesellschaft	(RAG)	odk	
gy/kameni	Nomsg/ugalj	Si/lu	Npfsi/nemačka	Zi.		Sa/lu	Mdo/[number]	Agpmsayn/nemački	Nomsn/rudnik	Nomsg/ugalj	Npmsn/rag	Npmsn/aktiengesellschaft	Zi(Npmsn/rag	Zi)	Vmp-sr	
</s><s>	Razvoj	sistema	upravljanja	kvalitetom	u	nemačkim	rudnicima	kamenog	uglja	započet	je	1989	.	kad			
/.	Nomsn/razvoj	Nomsg/sistem	Nomsg/upravljanje	Ncfsi/kvaliteta	Si/lu	Agpfdy/nemački	Nompr/rudnik	Agpnsy/kameni	Nomsg/ugalj	Vmn/započeti	Var3s/bitl	Mdo/[number]	Zi,	Cs/ka			
ani	su	postojeći	i	započinje	otvaranje	novih	rudnika	.	</s><s>	Slika	2.1	Proizvodnja	uglja				
znajati	Var3p/bitl	Agpmsy/postojeći	CoI	Vmr3s/započinjati	Nonsa/otvaranje	Agpmpgy/nov	Nompr/rudnik	Zi.		Ncfsn/slika	Mdo/[number]	Zi,	Agpmsy/proizvodnja	Nomsg/ugalj	!		
ugljenokopa	u	Srbiji	.	</s><s>	Prvi	vlasnik	Kostolačkog	rudnika	prvih	desetak	godina	bio	je	Franjo			
r	Nomsg/ugljenokop	Si/lu	Npfsi/srbija	Zi.		Mlomsn/prvi	Nomsn/vlasnik	Agpmsy/kostolački	Nomsg/rudnik	Zi,	Mlofpg/prvi	Mls/desetak	Nofpg/godina	Zi,	Vap-sm/bitl	Var3s/bitl	Npmsn/ra
io	21,68	miliona	tona	uglja	.	</s><s>	Od	nekadašnjih	rudnika	sa	jamskom	eksploatacijom	nastali	su	površ		
opati	Nonsn/z1,68	Nompr/milion	Nofpg/tona	Nomsg/ugalj	Zi.		Sg/od	Agpmpgy/nekadašnji	Nompr/rudnik	Su/sa	Agpmsy/jamski	Ncfsi/eksploatacija	Vmp-pm/nastao	Var3p/bitl	Agpmsnyj		
oko	Jasenovac	Štavalj	Vrška	Čuka				Ibarski	rudnici	Lubnica	i	Rudarsko-građevinsko	preduzeće	Alek			
n/soko	Zi,	Npmsn/jasenovac	Zi,	Npfpv/štavalj	Zi,	Npfsi/vrška	Npfsi/čuka	Zi,	Agpmpny/ibarski	Nompr/rudnik	Zi,	Npfsn/lubnica	CoI	Agpnsy/rudarsko-građevinski	Nomsn/preduzeće	Npmsr	
u	Boru	počelo	je	Francusko	društvo			borskih	rudnika	koncesija	Sveti	Đorđe	sa	sedištem	u	P	
Sa/lu	Npmsn/bor	Vmp-sr/početi	Var3s/bitl	Agpmsy/francuski	Nomsn/društvo			Agpmpgy/borski	Nompr/rudnik	Zi,	Ncfsn/koncesija	Agpmsy/sveti	Npmsn/dorđe	Su/sa	Nonsa/sedište	Si/lu	Npr
doo	(RB	Bor)	-	zavisno	preduzeće	.	Rudnik	bakra	Majdanpek	doo	(RB	Majdanpek		
msn/doo	Zi(Npmsn/rb	Npmsn/bor	Zi)	Zi-	Rgp/zavisno	Nomsn/preduzeće	Zi,	Agpmsny/	Nomsn/rudnik	Nomsg/bakar	Npmsn/majdanpek	Npmsn/doo	Zi(Npmsn/rb	Npmsn/majdanpek	
st	oblika	eksploatacije	mineralnih	sirovina	u	anketiranim	rudnicima	sa	ISM	Površinska	eksploatacija	je	zastupije				
enost	Nompr/oblik	Ncfs/eksploatacija	Agpfdy/mineralan	Nofpg/sirovina	Si/lu	Agpmsy/anketirani	Nompr/rudnik	Su/sa	Nomsn/ism	Agpmsy/površinski	Ncfsn/eksploatacija	Var3s/bitl	Agpmsny/zas				

Slika 4-6 Deo panela sa konkordancama sa prikazanim anotacijama vrstom reči i lemom

4.1.3. Terminološki resursi

Svi ovi rečnici, kao i niz drugih, ukupno njih 22, digitalizovani su u svrhu ovog istraživanja. Pod digitalizacijom se ne misli samo na puko skeniranje i optičko prepoznavanje karaktera (OCR), već na raščitavanje rečničkih članaka gde je prvo ceo tekst rečnika podeljen na rečničke članke, a onda svaki članak na elemente članka: odrednicu, definiciju, prevod, dodatne markere ukoliko postoje. Nakon raščitavanja, članci su pohranjeni u terminološku bazu, da bi se proširili i obogatili dostupni, pretraživi, digitalni resursi. Ove aktivnosti ne treba shvatiti kao jednokratne

⁶⁷ Konkordanca je format u kom se najčešće prezentuju automatski pronađeni primeri upotrebe date reči u kontekstu: (levi kontekst) → (ključna reč) → (desni kontekst)

aktivnosti, jer se svaka od njih može periodično ponavljati kada se pojave nove mogućnosti za dopunu i obogaćivanje resursa.

Dakle, u pripremljenoj fazi nakon skeniranja i optičkog prepoznavanja karaktera (OCR) rečnici su transformisani u MS Word, uz očuvanje formata (podebljano, kurziv) kako bi se moglo izvršiti detaljnije raščitanje rečničkih članaka. Ručna korekcija je neophodan korak i vrlo verovatno, postoje propusti jer je nekada potrebno više provera kako bi se dobio tekst bez grešaka sa dobro formatiranim člancima. Word dokumenti su zatim raščitani, postupkom razdvajanja rečničkog članka koji je morao da se prilagodi za svaki rečnik i da se uskladi sa njegovom strukturom. Raščlanjeni podaci konačno su transformisani u strukturirane formate: Excel i XML, pre nego što su uvezeni u relaciju bazu podataka. Na primer, od višejezičnog Rudarskog rečnika (Nešić 1970) je proizvedeno 15.016 rečničkih zapisa koji su sadržali samo jedan srpski pojam, ali bilo je 1355 zapisa sa dva termina, a 120 sa 3-5 termina, što je rezultiralo sa ukupno 18.092 srpskih termina, od kojih 16.916 različitih. Što se tiče engleskog dela rečnika, bilo je 13.163 zapisa sa jednim pojmom, 2553 sa dva pojma i 775 sa 3–8 pojmova, što je rezultiralo sa ukupno 20.878 pojmova na engleskom, od kojih 17.774 različitih.

Rudarska terminologija, slično opštoj tehničkoj terminologiji, sadrži veliki broj termina sastavljenih od više reči (višečlanih termina ili polileksemskih jedinica). U skupu podataka dobijenom iz rečnika što se tiče srpskih zapisa, 22% su jednočlani termini, 47% ima dve komponente, 17% ima tri, a preostalih 14% ima četiri ili više komponenata. Ove jezičke konstrukcije najčešće su sastavljene od dve ili više imenica, na primer, 'odlagalište jalovine', 'pritisak gasa',... međutim, mogu sadržati i tri, četiri ili više imenica, na primer 'lična zaštitna sredstva od gasova'.

Pomenuti rečnici iz rudarstva i srodnih domena koji su raščitani i pohranjeni u bazu podataka, su dodali još 63.571 novih zapisa, uključujući i one na engleskom. Pet jednojezičnih engleskih rečnika iz rudarskog domena proizvelo je 5933 zapisa, tri dvojezična rudarska englesko-srpska rečnika proizvela su 24.049 zapisa, tri jednojezična engleska rečnika koja pokrivaju terminologiju iz domena zaštite od mina doprinela su sa 655 zapisa, a englesko-srpski rečnik terminologije u oblast upravljanja otpadom dao je 1968 zapisa. Takođe su obuhvaćeni rečnici iz srodnih domena, naime četiri engleska rečnika koji su proizvela 21.448 zapisa i tri dvojezična rečnika koja su proizvela 9518 zapisa.

Jedno od zapažanja, čak i pre nego što je ovo istraživanje započelo, bilo je da neki pojmovi u papirnim rečnicima više nisu u upotrebi. To zapažanje nas je navelo da se izračuna učestalost termina u rudarskom korpusu na srpskom jeziku. Učestalost u korpusu i broj rečnika koji potvrđuju pojam bili su glavni kriterijumi za prioritet obrade termina u fazi post-procesiranja.

Zapisi iz svih digitalizovanih rečnika bili su uskladišteni u istoj bazi podataka, ali u različitim strukturama, koje odgovaraju njihovoj izvornoj šemi podataka, i uz referencu na originalni izvor. Sve strukture se, generalno, mogu preslikati na objedinjavanje struktura tako da se terminološki unos u zajedničku bazu podataka može sastojati od zaglavlja (liste), retko vrste reči, ekvivalenata u drugim jezicima, obično jednog, ali ponekad i više, markiranih značenja koja uključuju definicije, ponekada su dati i sinonimi i skraćenice, veze do drugih zapisa, bibliografija, i retko određeni domen.

4.2. Ekstrakcija važnih pojmova

Zadatak automatske ekstrakcije važnih pojmova ima za cilj identifikaciju skupa reči koje sažeto i jezgrovito opisuju tematiku teksta. Na osnovu prikupljenih i kreiranih resursa predstavljenih u Poglavlju 3, započeta je ekstrakcija važnih pojmova za razvoj ontologije. Slično istraživanjima prikazanim u poglavlju sa pregledom literature, i predmetno istraživanje je usvojilo niz metoda

za ekstrakciju važnih pojmova ciljnog domena, pri čemu su korišćene i postojeće taksonomije, strana iskustva, standardi, analiza uzorka dokumenta i slično.

Metode za ekstrakciju važnih pojmova se grubo dele na: nadgledane (eng. supervised), nenadgledane (eng. unsupervised) i učenje uz podsticaje (eng. reinforced learning). Osim takve grube podele, savremeni pristupi često metode dele i na statističke metode, lingvističke metode, metode bazirane na mašinskom učenju, neuronskim mrežama, grafovima i sl. Nadgledane metode zahtevaju unapred definisane ključne reči na osnovu kojih uče svoj model. Za nenadgledane metode to nije potrebno, one ekstrahuju iz teksta ključne reči bez potrebe da uče na prethodno definisanim ključnim rečima za pojedini tekst. Oslanjajući se na prethodna iskustva za ekstrakciju ključnih reči i termina za srpski jezik korišćena je kombinacija statističkih, lingvističkih i metoda zasnovanih na grafovima.

Metoda ekstrakcije ključnih reči zasnovana na selektivnosti (SBKE, Selectivity Based Keyword Extraction) zasnovana je isključivo na statističkim i strukturnim podacima izvedenim iz grafa reči teksta. SBKE metod zasniva se na izdvajanju ključnih reči iz mreže, u kojoj se mera selektivnosti čvora definiše se kao prosečna težina na vezama pojedinačnog čvora i koristi se u postupku izdvajanja kandidata za ključne reči (Beliga et al. 2017). U usmerenom grafu, stepen čvora i u oznaci k_i je broj ulaznih i izlaznih veza čvora, a težina s_i ulaza / izlaza čvora i definiše se kao broj njegovih dolaznih i odlaznih veza, a mera selektivnosti e_i je zapravo prosečna težina čvora koja se izračunava kao udeo težine čvora i stepena čvora:

$$s_i^{in/out} = \sum_j w_{ji/ij}, \quad e_i^{in/out} = \frac{s_i^{in/out}}{k_i^{in/out}}$$

Jedna od opcija je da se tekst preprocesira i radi sa lematizovanim tekstom (Kitanović 2016), što onda ipak znači da zahteva određene lingvističke podatke, tako da je SBKE uključena kao metoda da bi se postigao što bolji odziv ekstrakcije, imajući u vidu da metode zasnovane na preciznim sintaksičkim obrascima daju precizne podatke, ali odziv može biti nešto niži. Dakle, prihvaćena strategija je da se koriste svi raspoloživi resursi i sve raspoložive metode.

Analizom postojećih ontologija se uočavaju ključni termini, koji su kandidati za klase, tako da ih pišemo velikim početnim slovom: „*Risk, Cause, Impact, Likelihood, Consequence, Treatment, Gaps, Control, Exposure, Event, Proccess, Action, Activity...*“, što odgovara srpskim terminima „*Rizik, Uzrok, Uticaj, Verovatnoća, Posledica, Tretman, Propust, Kontrola, Izloženost, Događaj, Proces, Akcija, Aktivnost*“. Neke od fraza koje se vezuju za rizik, u zavisnosti od sintaksičke strukture upućuju na aktivnosti ili akcije, kao na primer: „*Risk (~cause, ~exposure, ~responsibility, ~management, ~treatment, ~identification, ...)*“, odnosno na srpskom „*Rizik (~uzrok rizika, izlaganje riziku, odgovornost za rizik, upravljanje rizikom, tretman rizika, identifikacija rizika...*“. Iz ovog primera se može uočiti razlika u kompleksnosti gramatičkih varijacija engleskog i srpskog jezika i jasno je da se modeli razvijeni za engleski ne mogu primeniti u srpskom već je potrebno razviti rešenje koje je specifično za srpski jezik i konkretan domen.

Jedan od načina za ekstrakciju ključnih reči je poređenje relativnih frekvencija domenskog korpusa sa korpusom standardnog jezika. Za potrebe ovog istraživanja je izdvojen deo korpusa RudKorp (Kitanović et al. 2021) sa tekstovima iz domena rizika, tako da je u ovom slučaju poređenje relativnih frekvencija podkorpusa RiskKorp sa relativnim frekvencijama referentnog korpusa (SrpKorp) ukazalo na reči koje su karakteristične za domen rizika.

ključnost = $\frac{fpm_{domen} + n}{fpm_{ref} + n}$ ⁶⁸, gde fpm_{domen} je normalizovana vrednost (per milion) frekvencije reči u korpusu koji je u domenskom korpusu, u konkretnom korpusu RiskKorp, dok fpm_{ref} je

⁶⁸ <https://www.sketchengine.eu/wp-content/uploads/ske-statistics.pdf>

normalizovana vrednost (per milion) frekvencija reči u referentnom korpusu, u ovom slučaju SrpKorp. Parametar zaglađivanja (eng. smoothing) kao podrazumevanu vrednost uzima 1. Rezultat prve faze ekstrakcije je grupisao ključne proste reči po vrstama reči, uz analizu njihove uloge u rečenici.

Tako su se izdvojile kao ključne imenice: *rizik, verovatnoća, hazard, izloženost, podsistem, pouzdanost, varijansa, rekultivacija, štetnost, varijabla, zagađivanje, otkaz, modeliranje, rangiranje, monitoring, polutant, radilište, ležište, sulfat...* U opisima rizika, imenica može da vrši sve službe reči, tj. ona može biti: subjekat, imenski deo predikata, objekat, apozicija, atribut i priloška odredba, a njima se označavaju predmeti (oprema, alati), materija (ruda, voda), bića (radnici, menadžeri), kao i neke nevidljive i neopipljive pojmove (rizik, izloženost). Kao karakteristične dvočlane termine već prisutne u elektronskom rečniku, sistem je prepoznao: donošenje odluka, tehnološki proces, vremenski period, podzemne vode, radni proces, sistem upravljanja,...

Među ključnim skraćenicama se izdvaja više grupa: metode i alati: *FTA, ALARP, FMEA, HAZOP, PROMETHEE, HSE, ANOVA, GIS...*, delovi naziva mašina: *SRS, ERS...*, organizacija: *BZR, LZS, LZO*, hemijskih elementa: *Zn, Pb, Fe, Cu, Ca...*, institucija: *TENT, JP, MUP*.

Među glagolima koji se izdvajaju kao karakteristični su: *otkazati, kvantifikovati, identifikovati, generisati, štetiti, izračunavati, umeriti, preraditi, utovarati, detektovati, osiguravati, proceniti, klasifikovati, kategorisati, modelirati, proračunati...* Glagoli su nesamostalne reči koje označavaju radnju koju neko vrši (*utovarati, detektovati, preraditi*), stanje u kojem se neko ili nešto nalazi (*razmišljati, pocrveneti, voleti, želeći*) i zbivanje (*lomiti, rositi, sevati, goretiti*), što će naročito biti važno za prepoznavanje uzročno-posledičnih relacija u tekstu kao izvora za pronalaženje instanci ontologije.

Karakteristični pridevi su: *zagađujući, rudnički, težinski, stohastički, rezidualni, vodni, stenski, osiguran, elektro-energetski, rudarski, požarni, eksploatacioni, zavistan, kvantitativan, rudni, regulatorni...* Važni su jer bliže određuju imenicu uz koju stoje, odnosno označavaju neku od osobina te imenice.

Prilozi obično određuju glagolsku radnju po mestu, vremenu, načinu, uzroku ili količini: *tehnološki, nizvodno, ekstremno, konstantno, potencijalno, drastično, tehnički, prostorno, propisno...*, dok pored i osim funkciju imaju ulogu predloga.

Glagolske imenice su posebno interesantne jer označavaju vršenje radnje, rezultat svršene radnje, stanja ili zbivanja, kao na primer: *reosiguranje, drobljenje, zagađivanje, modeliranje, prosejavanje, rangiranje, upravljanje, identifikovanje, osiguranje, bušenje, minimiziranje, grupisanje, otkopavanje, merenje, zavarivanje, habanje*

Određena lista reči je posebno označena, kao na primer reč *zbog* kao naročito važna za prepoznavanje fraza od interesa.

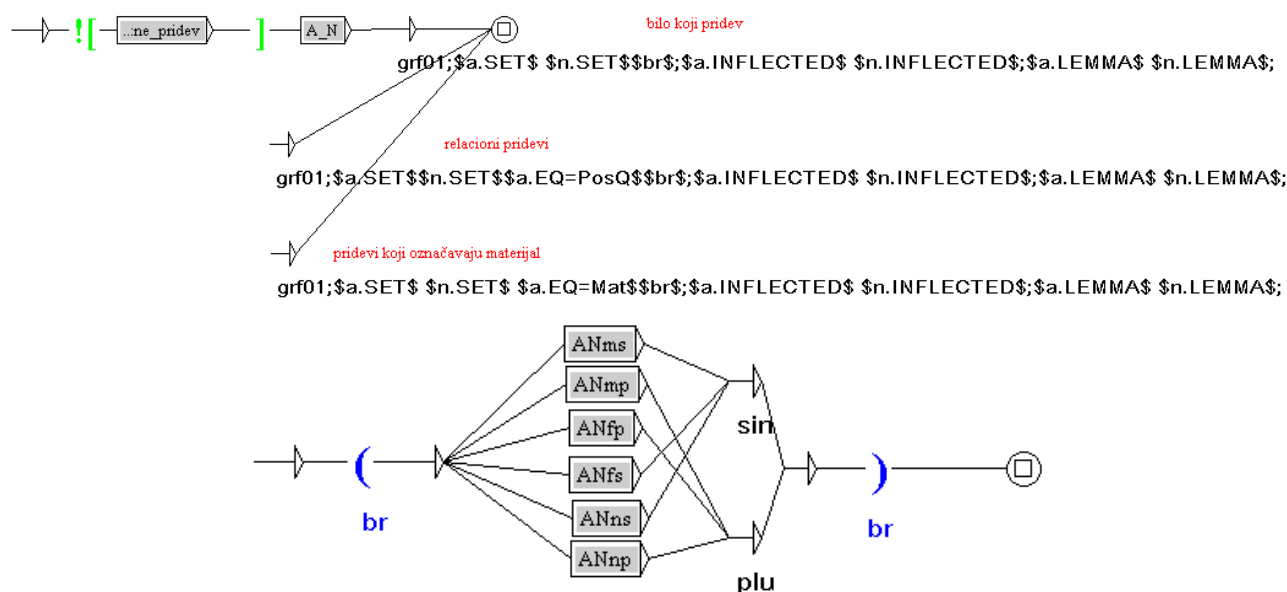
Zahvaljujući semantičkim markerima (Obradović et.al, 2017) je moguće izdvojiti koncepte karakteristične za pojedine semantičke grupe reči. U nastavku teksta navodimo reči koje su izdvojene iz korpusa rizika i rudarstva.

Plitkim parsiranjem teksta i ekstrakcijom terminologije zasnovanom na kolekciji grafova (Krstev et al. 2015; Stanković, Krstev, Obradović, et al. 2016) su dobijeni karakteristični višočlani termini za izabrane ključne reči. U nastavku teksta prikazujemo primere analize višočlanih termina na primeru reči. „rizik“.

Odgovor na pitanje „Kakav je rizik?“ daje obrazac pridev-imenica (*AXN, grf01*), koji izdvaja karakteristične opise rizika. Slika 4-7 prikazuje transduktor, odnosno konačni automat sa izlazom koji prepoznaje i ekstrahuje dvočlane izraze pridev imenica koje se slažu u rodu, broju, padežu, što ilustruje donji deo slike. Transduktori definišu relacije između dva skupa niski

karaktera, odnosno transformišu zadatu nisku u drugu nisku nad istom ili nekom drugom azbukom; prepoznaju jednu nisku karaktera a generišu drugu. Graf koristi rečničke promenljive u izlazu transduktora, tako da uz prepoznati oblik \$a.INFLECTED\$ daje i lemu, odnosno kanonski oblik reči: \$a.LEMMA\$ \$a\$, izvršavajući na taj način jednostavnu lematizaciju reči (Krstev et al. 2015). Promenljiva \$br\$, koja je u donjem delu grafa označena plavim zagradama ukazuje na gramatičku množinu. Dobijene informacije su iskorišćene za klasifikaciju rizika, a evo nekoliko primera:

- {strategijski, sistematski, ekološki, tržišni, finansijski, kreditni, požarni, politički, sistemski, proizvodni, kamatni, regulatorni, ekonomski, dinamički, dopunski, pravni, valutni, imovinski, statički, industrijski, tehnološki, makroekonomski, hemijski, tehnički, zdravstveni, nuklearni, strateški, netehnički, transportni, tranzitni, ergonometrijski, biološki, računovodstveni, klimatski, neekonomski, procesni, investicioni, devizni} **rizik**.



Slika 4-7 Konačni automat za ekstrakciju strukture pridev-imenica

Odgovor o pripadnosti i poreklu rizika, odnosno delu rizika koji se dobijaju na pitanja: *Koga? Čega? Od koga? Od čega?* Takođe pitanja *S kime? Čime?* će dati konstrukcija gde se iza imenice nalazi reč koja ne podleže fleksiji, najčešće genitiv ili instrumental ($N2X$, $grf03$), pri čemu se rizik može javiti kao prvi ili drug član termina:

- {upravljanje, monitoring, menadžment, sistem upravljanja, sistem za upravljanje} **rizikom**, izloženost **riziku**,
- {menadžer, analitičar, procenjivač, preuzimač, nosilac} rizika
- {nivo, uticaj, stepen, kategorija, pokazatelj, indeks, efekat, efekt, vrsta, definicija, izvor, intenzitet, prag, tip, klasa, matrica, problem, vrednost, pojam, scenario, karakteristika, priroda, veličina, strategija, komponenta, posledica, verovatnoća, aspekt, parametar, model, ekspertiza, prenos, cena, stanje, mapa, prihvatljivost, obim, sekvenca, mera, učestalost, rang, faktor, činilac, neizvesnost, opasnost, registar, kontekst, uzrok} **rizika**
- {procena, analiza, smanjenje, merenje, preuzimanje, redukcija, ocena, kontrola, identifikacija, sprečavanje, prenošenje, rangiranje, prihvatanje, teorija, tretman, otklanjanje, klasifikacija, procenjivanje, ispoljavanje, definisanje, transfer, menadžment, karakterizacija, raspodela, finansiranje, adresiranje, praćenje, modeliranje, smanjivanje, zadržavanje, nošenje, ublažavanje, monitoring, osiguranje, održavanje, identifikovanje, eliminisanje, tretiranje, sagledavanje, posmatranje, redukovanje, umanjivanje, određivanje, vrednovanje, pozicioniranje, izučavanje, ostvarivanje, realizacija, prevencija, upravljanje,

pokrivanje, povećanje, izbegavanje, mapiranje, iskorišćavanje, tumačenje, označavanje, uključivanje, planiranje, kategorizacija, savladavanje, snižavanje, ocenjivanje, umanjivanje } **rizika**

- **rizik** {*zemlje, preduzeća, sistema, elemenata, procena, loma, procesa, ulaganja, bezbednosti, likvidnosti, primenom, požara, terorizma, analiza, poslovanja, organizacije, gubitka, društava, delovanja, bankrotstva, implementacije, zagađivanja, osiguranja, opasnosti, korišćenjem, verovatnoća, transporta, života, matrice, posledica, nemogućnosti, fluktuacije, strategije, finansiranja, zastarevanja, podsistema, portfelja, društva, osiguranika, havarije, sklopova, akcija, preduzetnika, cena, instituta, banke, pogoršanja, neuspeha, tehnike, krađe, stopa, prevara, nesolventnosti, fluktuiranja, neizvršenja, partnera, proizvođača, prekoračenja, suvereniteta, reinvestiranja, investitora, obavljanja, efikasnosti, pojava, mašina, obaveze, samokontrole, kreditiranja, fondova, firme, reputacije, nesposobnosti, kamata, postrojenja, regiona, učešća, smrti, nezgode, invalidnosti, odgovornosti, greške, projektanta, mraza, razvoja*}

Na pitanja: *Od čega je rizik? Na šta se odnosi rizik?* Odgovor daje rezultat grafa N4X koji traži fraze u kojima se iza imenice nalaze dve reči koje ne podležu fleksiji, a mogu biti sintaksičke strukture gde iza reči rizik slede predlog i imenica.

- **rizik od** {*zagađivanja, požara, otkaza, povreda, loma, pada, eksplozije, odgovornosti, terorizma, povrede, nastanka, udara, udesa, poplava, prignječenja, opekotina, posekotine, povređivanja, nezaposlenosti, poplave, opekotine, opasnosti, kontakta, prekida, zagađenja, štete, orkana, prevrtanja, okliznuća, uboda, degradacije, ugrožavanja, pojave, azbesta, prosipanja, izlaganja, oluje, uvlačenja, upada, akcidenta, otpada, važnosti, curenja, pregrevanja, infekcije, zemljotresa, izostanka, proizvoda, oštećenja, krađe, prevara, nesolventnosti, uništenja, nezgode, oboljenja, greške, smrti, smrkavanja, odbijanja, udarca, spoticanja, odrona, uganuća, aeroxagađenja*};
- **rizik za** {*elemente, proces, sklop, preduzeće, aspekt, aktivnosti, društvo, rezultate, podsisteme, populaciju, osiguranje, procese, pod, cilj, potrebe, sliv, slivove, odlagališta, poslovanje, ocenu, članove, organizaciju, vlasnike, konstrukciju, grad, imovinu, vlasnika, postrojenja, poslove, kičmu, stanja, implementaciju, sisteme, preradu, povređivanje*}
- **rizik prema** {*izvoru, načinu, preporukama, izvorima, dobavljaču, podacima, šemi, značaju, izgledima, posledicama, mogućnosti, uticaju*}
- **rizik u** {*preduzeću, rudarstvu, procesu, preradi, organizaciji, kontekstu, poslovanju, modelu, preduzećima, sistemu, analizi, osiguranju, svetu, zemlji, zakonu, poljoprivredi, organizacijama, okruženju, firmi, praksi, konstruisanju, igrama, izolaciji, sektoru, bilansima, postupku, procesima, sistemima, slučajevima, firmama, bankama, prihodu, javnosti, rudnicima, saobraćaju, industriji, pogonu, području, algoritmu, privredi*}
- **rizik usled** {*opasnosti, diversifikacije, otkaza, pojava, udara, dejstva*}
- **rizik unutar** {*preduzeća, kompanije*}
- **rizik po** {*zdravlje, bezbednost, zagađivanje, većinu, obrascu, tabeli, segmentima, vozače, potrebi, delovima, elementima*}
- **rizik sa** {*aspekta, predlogom, strategijom, indeksom, uticajem, namerom, kategorizacijom, tendencijom, delegiranjem, definicijama, stabilnošću, vlasnicima, misijom, analizom, modelom, vremenom, metodom, dokazima, izradom, ciljem*}

Sintaksičke strukture gde iza imenice i predloga sledi reč rizik daje sledeće rezultate:

- {*osiguranje, zaštita, uzdržavanje, obezbeđenje, osiguranik, zavisnost*} **od rizika**
- {*premija, odgovornost, kapacitet, kompenzacija, servis, osiguranje, rešenje*} **za rizik**
- {*podatak, izveštavanje, svest, parametar, izveštaj, analiza, zaključak, utisak, slika, saznanje, informisanje, dokument*} **o riziku**
- {*usklađivanje, veza, postupanje, vrednost, suočavanje, nagrada, paket, odnos, korelacija*} **sa rizikom**

Odgovor na pitanje *Čega je rizik?* Daje sintaksička struktura u kojoj iza imenice se nalaze dve reči koje ne podležu fleksiji unutar višočlane reči; druga reč može biti imenica ili pridev. Tako se reč rizik može naći na bilo kojoj od tri pozicije u frazi.

- **rizik** {*tehničkog, proizvodnih, celog, elemenata, tehničkih, industrijskih, tehnoloških, informacionih, kompleksnih*} **sistema**
- **rizik** {*sistematske, regulatorne, jednostrane, različite*} **prirode**
- **rizik velikog** {*intenziteta, opsega, obima*}
- **rizik elemenata** {*podсистема, sklopa, sistema*}
- **rizik prirodnih** {*nepogoda, katastrofa*}
- **rizik** {*loma mašina, radnog mesta, finansijske institucije, prirodnih katastrofa, katastrofalnog događaja, radne sredine, implementacije strategije, celokupnog preduzeća, finansiranja strategije, regulatorne prirode, zastarevanja kapaciteta, uticajnih faktora, radne delatnosti, upravljačke delatnosti, ugroženih lokacija*}
- {*proces, kontekst, prikaz, koncept, aspekt, potreba, rezultat, pristup, podatak, diskusija, postavka, objekt, unapređenje, faza, primer, cilj, većina slučajeva, model, disciplina, gledište, kriterijum, metod, metoda, problem*} **analize rizika**
- {*merenje, delovanje, osobina, prirod, priroda, efekat, efekt, definisanje, kategorija, faktor, oblik, aspekt, vrsta, shvatanje, definicija, dekompozicija, komponenta, izvor, preuzimanje, potcenjivanje, verovatnoća, rešavanje problema*} **operativnog rizika**
- {*definisanje, dispozicija, identifikacija, identifikovanje, klasifikacija, mnoštvo, model, modelovanje, otkrivanje, postojanje, proizvod, realizacija, scenarij, scenario, učestalost, uticaj, varijacija, značaj*} faktora rizika
- {*definisanje, izračunavanje, kategorija, ocena, određivanje, oscilacija, oznaka, podatak, pokazatelj, praćenje, procena, promena, smanjenje, utvrđivanje, vrednost*} **nivoa rizika**
- **upravljanje** {*operativnim, ekološkim, profesionalnim, strategijskim, celokupnim, sistematskim, finansijskim, regulatornim, komercijalnim, ličnim, unutrašnjim, povećanim*} **rizikom**
- **procena** {*ekološkog, jedinstvenog, katastrofalnih, konstruktivnih, kreditnog, mogućih, nivoa, potencijalnih, požarnog, profesionalnog, sekundarnih, stanja, stepena, stvarnog, ukupnog, uticaja, zona*} **rizika**
- **analiza** {*efekata, ekološkog, fenomena, globalnih, pojedinačnog, pokazatelja, relativnog, strategijskog, uticaja*} **rizika**
- **nivo** {*indeksa, kreditnog, preuzetog, prihvatanja, profesionalnog, sistematskog, smanjenja, tolerisanog, ukupnog*} **rizika**
- **uticaj** {*individualnih, ključnih, makroekonomskog, nepovoljnih, odgovarajućeg, određenog, pojedinačnih, problema, relevantnih, sistematskog, strategijskog, tržišnog, valutnog*} **rizika**
- **posledica** {*čistog, definisanog, dejstva, delovanja, karakteristika, mnogih, nastupanja, ostvarenih, prisutnih, sledećih, špekulativnih, zadržanog*} **rizika**

Tročlane izraze pridev-imenica-imenica prepoznaje grf05 AXN2X:

- **ukupan** {*faktor, indeks, nivo, portfelj, stanje, stepen, uticaj*} **rizika**
- **različit** {*definicija, kategorija, nivo, nosilac, scenarij, scenario, stepen, tip, vrsta*} **rizika**
- {*preliminarna procena, izdvojena klasa, sistemska analiza, određena klasa, parcijalan pokazatelj, pojedinačan faktor, visok stepen, ključni indikator, ključni faktor, nizak stepen, negativan efekat, prostorna analiza*} **rizika**

Graf grf08a NC_N6X, prepoznaje strukturu koju obeležavamo sa NNgiPrepNp, što znači da iza imenice se nalaze imenica u genitivu ili instrumentalu i predložka fraza, kao na primer:

- **upravljanje rizikom u** {*preduzeću, rudarstvu, preradi*}

- {procena, stepen, indeks, analiza, kategorija, karta } **rizika od zagađivanja**
- **procena rizika** {na radu procena, od požara, od povreda}
- *rizik bezbednosti na radu, analiza rizika od požara, nivo rizika od otkaza, nivo rizika u procesu*

Graf grf08b NC_N6X, prepoznaje strukturu koju obeležavamo sa NNgiNgiNgi, što znači da iza imenice se nalaze tri imenice/prideva u genitivu ili instrumentalu;

- **zona praga** {niskog, ozbiljnog, umerenog, visokog} **rizika**
- {merenje efekata, model merenja, nivo preuzimanja, objašnjenje prirode, predviđanje efekata, rešavanje problema, uticaj faktora} **operativnog rizika**
- **upravljanje rizikom** { bankrotstva klijenta, celokupnog preduzeća, finansijske odluke, finansijskih institucija, fizičkog okruženja, katastrofalnog događaja, kompleksnih sistema, radne sredine, radnih mesta, tehničkih sistema}
- *procena rizika tehničkih sistema, rizik zagađivanja životne sredine, adaptacija koncepta mapiranja rizika, kombinacija scenarija individualnih rizika, postupak analize efekata rizika*

Graf grf08c NC_N6X, prepoznaje strukturu koju obeležavamo sa NprepNpNgi, što znači da iza imenice se nalaze tri reči koje ne podležu fleksiji unutar višočlane reči i iza imenice se nalazi predložka fraza;

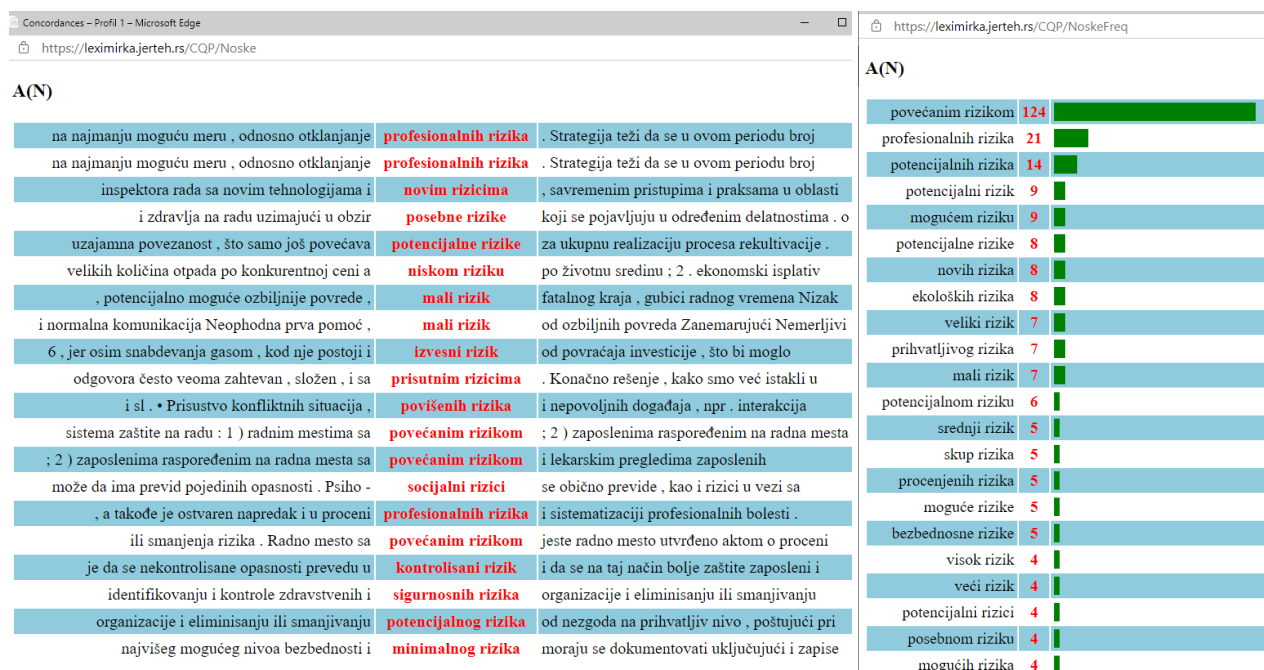
- **rizik od** {curenja rezervoara, degradacije zemljišta, eksplozije {gasa, materijala}, fluktuacija cena, karcinoma pluća, loma {držača, mašina, osovine, rotora, stakla}, nastanka povreda, oštećenja automobila, otkaza {dela, delova, drobilice, elemenata, hidromotora, podsistema, senzora}, pada {predmeta, sirovine, tereta}, pojave vlage, povreda {ruke, glave, radnika, ruku, šake}, pregrevanja ulja, prekida poslovanja, prignječenja {noge, šake}, prosipanja materijala, udarca ruke, ugrožavanja kvaliteta, ugrožavanja zdravlja, upada {čestica, piljevine}, uvlačenja prstiju}
- **{rizik u** {analizi aspekta, bilansima banaka, fazi dizajna, kontekstu pouzdanosti, modelu vrednovanja, oblasti {bezbednosti, osiguranja}, poslovanju {društava, društva}, postupku analize, preradi uglja, procesu {analize, proizvodnje, projektovanja}}
- **rizik sa** {aspekta {bezbednosti, verovatnoće}, ciljem dostizanja, definicijama pojmova, delegiranjem autoriteta, izradom akta, kategorizacijom nivoa, misijom organizacije, nivoa preduzeća, određivanjem nivoa, predlogom {metoda, modela}, tendencijom približavanja, tržišta osiguranja, uticajem faktora, vlasnicima kuća}
- **metoda za** {analizu, identifikaciju, kontrolu, kvantifikaciju, merenje, obradu, ocenu, prevazilaženje, procenu, redukciju, smanjenje, smanjivanje, tretman} **rizika**
- **mera za** {kontrolu, otklanjanje, redukciju, reosiguranje, smanjenje, ublažavanje} **rizika**

Graf grf09b AXN4X prepoznaje izraze kod kojih se ispred imenice nalazi pridev sa kojim se imenica slaže u rodu, broju, padežu i animatnosti, a iza imenice se nalaze dve reči koje ne podležu fleksiji unutar višočlane reči; moguće su dve sintaksičke strukture: ANNgiNgi – ispred imenice se nalazi pridev, a iza imenice dva prideva/imenice u genitivu ili instrumentalu:

- {povećan stepen profesionalnog, parcijalan uticaj faktora, grafički prikaz redukcije, velik broj faktora, moguć rezultat delovanja, pojedinačan scenario ključnih, originalan model faktora, adekvatan sistem merenja, ključan indikator sistematskog, makroekonomski indikator sistematskog, potencijalan scenario posmatranih, različit scenario faktora} **rizika**
- *operativan rizik {finansijske institucije, niske frekvencije, regulatorne prirode}, značajan rizik finansiranja strategije*

Alat Unitex predstavlja veoma moćan alat za analizu i ekstrakciju informacija, ali je nedostatak što je desktop aplikacija koja se izvršava lokalno na računaru koristeći lokalne resurse, a uz to

i zahteva određeno znanje i veštine. Kako bi se korisnicima, koji nemaju znanja o korišćenju CQP jezika niti Unitex-a obezbedila mogućnost pretrage korpusa vođene obrascima, u aplikacije Termi i Leximirka su ugrađeni neki od obrazaca kojima je moguća raznovrsna ekstrakcija. Rezultat pretrage u CQP-u se analizira na različite načine: lista frekvencija, kolokacije, konkordance sa užim i širim kontekstom. Slika 4-8 (levo) prikazuje konkordance dobijene iz Leximirke za obrazac pridev-imenica za reč rizik, dok se na desnoj strani slike može videti histogram sa frekvencijama oblika za isti obrazac.



Slika 4-8 Konkordance obrasca pridev imenica za "rizik"

Rudarski korpus je publikovan i na Sketch engine platformi, gde se mogu postaviti različiti upiti. Recimo možemo izvući kolokacije izabrane leme ili izraza, potom kolokacije za lemu, takođe i tezaurus srodnih reči, takozvanu skicu reči [lemma="rizik"] ili videti razliku upotrebe neke dve, obično srodne reči.

Recimo na pitanje *Čega je rizik?* Odgovor bi mogli naći sledećim CQL upitom [lemma="rizik"][tag="N"]. Odgovor na pitanje *Kakav je rizik?* Bi dao upit [tag="A"][lemma="rizik"], ili ako želimo da dozvolimo razmak između reči (ne više od 5) onda bi napisali upit na sledeći način: [pos="V"][word!="\."]{0,5}[lemma="rizik"].

Skica reči (eng. Word sketch) daje brzi pregled ponašanja izabrane reči, tako što prikuplja informacije iz hiljada i miliona primera upotrebe i pruža sažetak kategorizovanih kolokacija na jednoj stranici sa vezama do pojedinačnih primera. Slika 4-9 prikazuje primer skice reči za reč „rizik“ gde jedan pogled na stranicu ilustruje kako se ta reč koristi.

The screenshot shows the WORD SKETCH interface for the word "rizik". The search bar contains "RudKor". The main area is divided into five columns, each representing a different grammatical category:

- verb prepositional phrases:** "rizik" od ..., ... od "rizik", "rizik" sa ..., ... sa "rizik", "rizik" u ..., ... u "rizik", "rizik" po ..., ... po "rizik", "rizik" na ..., ... na "rizik", "rizik" za ..., ... za "rizik".
- modifiers of "rizik":** povećati (radnom mestu sa povećanim rizikom), potencijalan (potencijalni rizik), politički (politički rizik), identifikovati (identifikovani i rangirani rizici), proceniti (procenjeni rizici), neprihvatljiv (neprihvatljiv Uopšteno • visoki rizik), izložiti (su zaposleni posebno izloženi rizicima u slučaju nestanka), ohs (neophodna zbog upravljanja OHS rizicima. To uključuje), specifičan (kojima se pojavljuje specifičan rizik od nastanka povreda).
- verbs with "rizik" as subject:** smanjiti (bi se smanjio rizik), nastajati (preventivnih mera, rizik koji nastaje usled izloženosti zaposlenih), postojati (rizik ukoliko postoji), definisati (identifikuju opasnosti i rizici, te se definišu potrebne kontrole vezeno), moći (rizik može), trebati (rizik treba), morati (identifikuju opasnosti i rizici vezane uz promene, organizacija mora osigurati da se), predstavljati (Rizik predstavlja), imati (rizik vezan za profitabilnost ima).
- verbs with "rizik" as accusative object:** postojati (postoji rizik od), smanjiti (smanji rizik od), proceniti (da proceni rizik), smanjivati (u značajnoj meri smanjuje rizik i olakšava proces), eliminisati (način koji bi eliminisao rizik? Korak), nositi (nose dosta veliki rizik), povećavati (povećava rizik), predstavljati (predstavlja rizik), utvrditi (utvrditi i mogući rizik), sadržati (sadrži minimalni rizik).
- conjunctions after "rizik":** od (sa visokim rizikom od izbijanja požara), kako (rizik kako), kao (mestu sa povećanim rizikom kao i pre premeštaja), i (sa povećanim rizikom i), ili (mestu sa povećanim rizikom ili za upotrebu), da (postoji rizik da).

Slika 4-9 Skica reči "rizik" u alatu Sketch engine

Istraživanje kolokacija je od izuzetnog značaja tako da sa skice reči za reč rizik i kolokaciju „rizik od“ mogu se prikazati konkordance, što prikazuje Slika 4-10.

The screenshot shows the CONCORDANCE interface for the word "rizik". The search bar contains "RudKor". The main area displays a list of 20 concordance entries, each with a document ID (doc#), a snippet of text, and the word "rizik" highlighted in red. The text snippets show various contexts where "rizik" is used, often in relation to safety, risk management, and environmental factors.

Slika 4-10 Konkordance u alatu Sketch engine

Statistički prikaz frekvencija može se prikazati i grafički. Slika 4-10 prikazuje frekvencije pojave termina rizik sa pridruženim pridevima.

	Lemma	Frequency ↓	Relative ?
1	povećati rizik	114	43.26
2	potencijalan rizik	18	6.83
3	velik rizik	11	4.17
4	visok rizik	8	3.04
5	specifičan rizik	7	2.66
6	politički rizik	6	2.28
7	mali rizik	6	2.28
8	moguć rizik	6	2.28
9	sav rizik	4	1.52
10	nov rizik	4	1.52
11	značajan rizik	4	1.52
12	postojeći rizik	4	1.52
13	izložiti rizik	4	1.52
14	glavni rizik	4	1.52
15	zaposliti rizik	3	1.14
16	geološki rizik	3	1.14
17	ekonomski rizik	3	1.14
18	identifikovati rizik	3	1.14
19	ekološki rizik	3	1.14
20	izvestan rizik	3	1.14

Rows per page: 20 1–20 of 93

Slika 4-11 Frekvencije kolokacija za reč rizik u alatu Sketch engine

Slika 4-12 prikazuje panel razlika u skicama reči (eng. Word Sketch Difference) koji daje poređenje upotrebe dve reči, a predstavlja proširenje modula skice reči. Generiše skice reči za dve reči i upoređuje ih, dozvoljavajući da se uoče razlike u upotrebi. Funkcija je posebno korisna za bliske sinonime, antonime i reči iz istog semantičkog polja. Sa slike se, na primer, može videti da se uz „rizik“ koriste „povećati, politički, proceniti, prihvatljiv“ dok uz „opasnost“ ide „neposredan, ozbiljan, iznenadan“.

WORD SKETCH DIFFERENCE

rizik 1,188x | opasnost 1,177x

modifiers of "rizik/opasnost"

povećati	114	0	...
politički	7	0	...
proceniti	4	0	...
neprihvatljiv	3	0	...
potencijalan	21	21	...
identifikovati	4	8	...
izložiti	5	20	...
nov	4	18	...
prisutan	0	5	...
iznenadan	0	5	...
ozbiljan	0	12	...
neposredan	0	45	...

prepositions before "rizik/opasnost"

nad	3	0	...
prema	5	0	...
za	7	0	...
sa	20	6	...
na	12	17	...
o	12	18	...
od	0	12	...
iz	0	7	...
usled	0	3	...
zbog	0	5	...
bez	0	10	...
protiv	0	3	...

"rizik/opasnost" u ...

organizacija	6	0	...
proces	9	0	...
oblast	5	0	...
veza	3	0	...
slučaj	4	0	...
rudarstvo	6	4	...
biti	0	8	...
dovesti	0	3	...

"rizik/opasnost" na ...

odgovor	7	0	...
projekt	7	0	...
nivo	5	0	...
rad	4	9	...
odnos	3	10	...
mesto	0	5	...

... na "rizik/opasnost"

odgovor	7	0	...
projekt	7	0	...
nivo	5	0	...
rad	4	9	...
odnos	3	10	...
mesto	0	5	...

"rizik/opasnost" za ...

zdravlje	4	0	...
bezbednost	7	0	...
upozorenje	0	3	...
život	0	4	...

"rizik/opasnost" od ...

otkaz	19	0	...
poplava	17	0	...
izloženost	4	0	...
nastanak	38	3	...
lom	6	3	...
povreda	5	5	...
eksplozija	18	31	...
izbijanje	4	15	...
klizanje	0	4	...
požar	0	14	...
dodir	0	6	...
povrednjanje	0	7	...

verbs with "rizik/opasnost" as accusative object

smanjiti	14	0	...
smanjivati	7	0	...
eliminirati	3	0	...
nositi	3	0	...
povećavati	3	0	...
proceniti	10	5	...
postojati	43	68	...
predstavljati	24	42	...
prepoznati	0	4	...
prouzrokovati	0	6	...
otkloniti	0	5	...
izazvati	0	12	...

verbs with "rizik/opasnost" as subject

smanjiti	4	0	...
nastajati	6	0	...
trebati	4	0	...
morati	5	0	...
predstavljati	5	0	...
imati	4	0	...
moći	40	18	...
definisati	6	3	...
postojati	8	8	...
identifikovati	0	3	...
.	0	4	...

noun objects of "rizik/opasnost" as infinitive

upravljati	7	0	...
smanjiti	5	0	...
predstavljati	3	0	...
proceniti	7	4	...
biti	3	5	...
identifikovati	0	4	...
prouzrokovati	0	4	...
otkloniti	0	3	...
prepoznati	0	4	...

Slika 4-12 Skica razlika upotrebe reči rizik i opasnost

4.3. Ekstrakcija rudarskih entiteta iz teksta

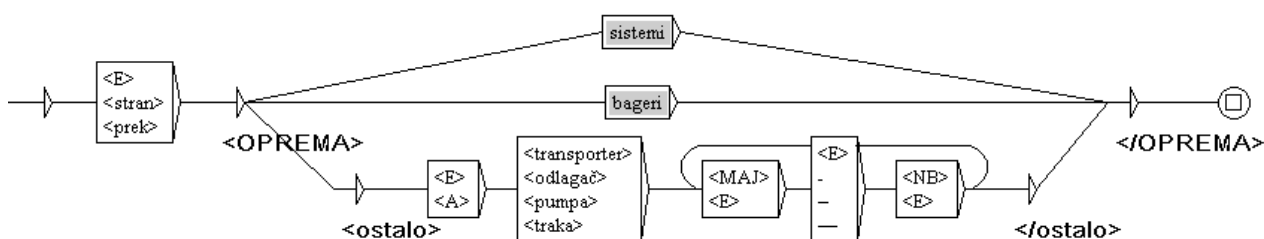
Ekstrakcija svih entiteta vrši se kroz okruženje Unitex uz pomoć prethodno pripremljenih rečnika srpskog jezika (Krstev 2008; Obradović et al. 2017; Tomašević et al. 2018). Za različite vrste i podvrste entiteta pripremljeni su različiti konačni automati u vidu grafova prepoznavanja. Ovi grafovi vrše i funkciju transduktora, istovremeno obeležavajući prepoznate entitete u tekstu. U grafovima se koriste termini iz rečnika oivičeni uglastim zagradama, što znači da će biti prepoznato njihovo pojavljivanje u bilo kom flektivnom (kosom) obliku.

4.3.1. Entiteti rudarske opreme

Za prepoznavanje (odabrane) rudarske opreme koristi se graf koji prikazuje Slika 3-2. Ovaj graf ekstrahuje entitete u tekstu koji predstavljaju bagere, bagerske sisteme, transportnu opremu i određene klase pomoćne mehanizacije i obeležava ih etiketom `<OPREMA>`. Bageri, otkopno-utovarni i proizvodni sistemi se prepoznaju pomoću specijalnih podgrafova koje prikazuje Slika 4-14 i Slika 4-15, dok se ostala oprema prepoznaje kao sekvenca elemenata u samom grafu opreme i odmah obeležava etiketom `<ostalo>`. Sistem je fleksibilan i može uključiti prepoznavanje drugih tipova opreme. Ostala oprema se prepoznaje kao pojavljivanje transportera, odlagača, pumpe ili trake praćeno opcionim pojavljivanjem mešavine velikih slova i brojeva, kako se obično obeležavaju ovi primerci opreme.

U tekstu se često pre tipa opreme može naći i opcionim pridev za približnije određenje tipa što se modelira čvorom `[<E>/<A>]`, gde `<E>` označava praznu reč a `<A>` označava pridev. Kako bi se razrešila višeznačnost prideva koji mogu dati pogrešne rezultate, pridevi *stran* i *prek* se isključuju iz prepoznavanja sekvence. Ovi pridevi imaju iste oblike kao imenica *strana* i predlog *preko*, tako da njihovo višeznačnost može dati neželjene izlaze. Graf za ekstrakciju entiteta rudarske opreme, na ovaj način prepoznaje primerke opreme poput:

- **Pumpa KV-1**
- **veznim transporterom**
- **Odlagač 3**
- **transportnih traka**

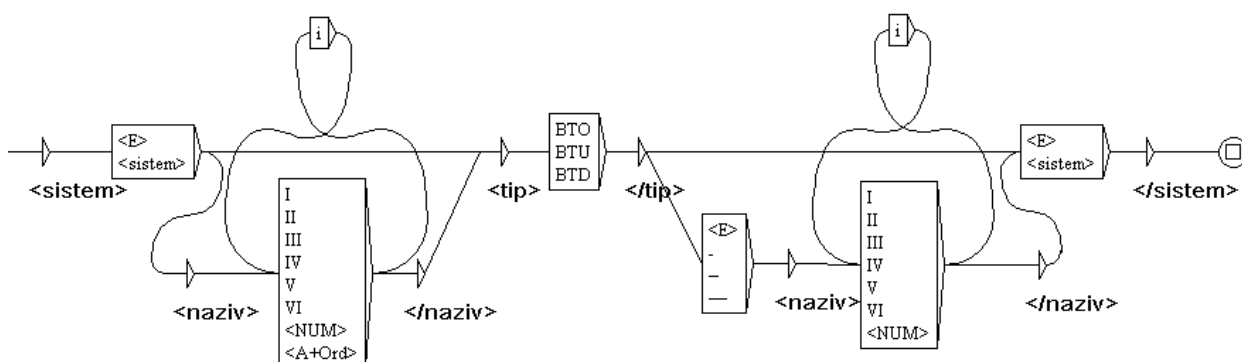


Slika 4-13 Graf za ekstrakciju entiteta rudarske opreme

Proizvodni, otkopno-utovarni sistemi se prepoznaju na osnovu pojavljivanja oznaka *BTO*, *BTU* i *BTD*, što se obeležava etiketom `<tip>`. Pre i posle oznake tipa sistema su prepoznati i opcionim brojevi (pretežno rimski), što jesu njihovi uobičajeni nazivi, pa su transduktorima i obeleženi etiketom `<naziv>`. Takođe, imenica sistem je dopisana na bilo koji kraj prepoznate niske reči (Slika 4-14). Na ovaj način su prepoznata pojavljivanja poput:

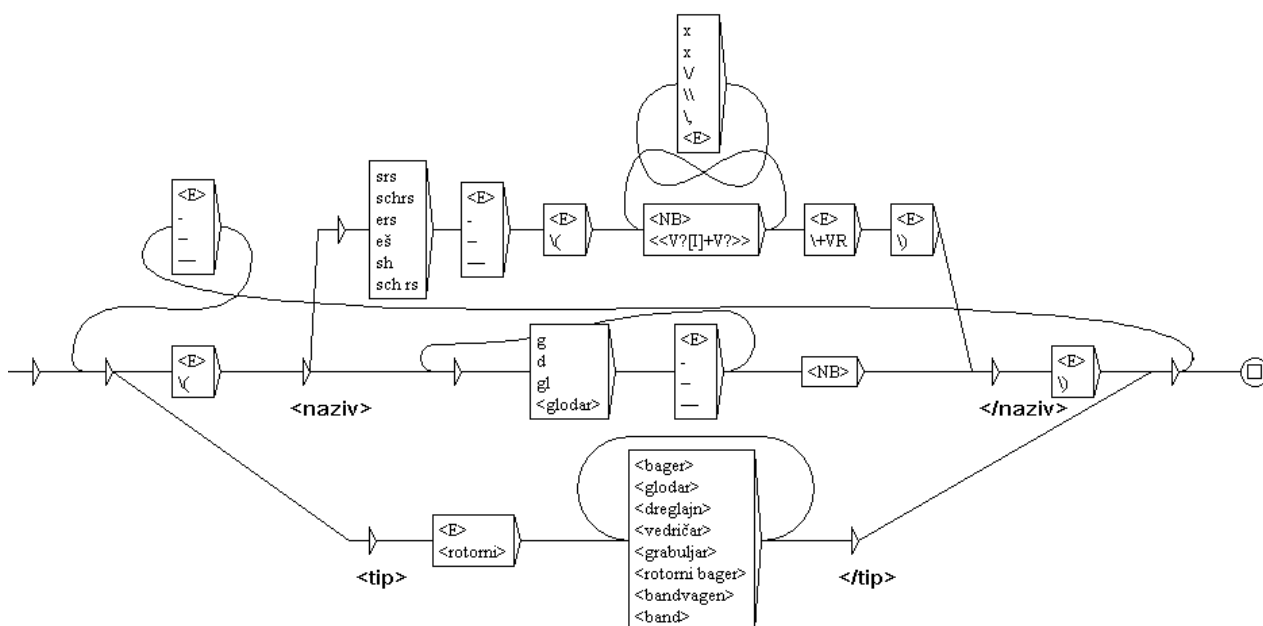
- **BTD** Sistema
- **BTO-1** sistema
- **II BTO** Sistema

- treći **BTO** sistem



Slika 4-14 Graf za ekstrakciju bagerskih Sistema

Graf za prepoznavanje samih bagera i njihovih imena je najkompleksniji. Može se grubo podeliti na dva dela: prepoznavanje tipa otkopne opreme odnosno bagera (donji deo grafa) i prepoznavanje naziva (srednji ili gornji deo grafa). Ova dva dela se mogu javljati naizmenično jedan pokraj drugog, odvojeni razmakom ili crticom.



Slika 4-15 Graf za ekstrahovanje pojavljivanja bagera i njihovih naziva

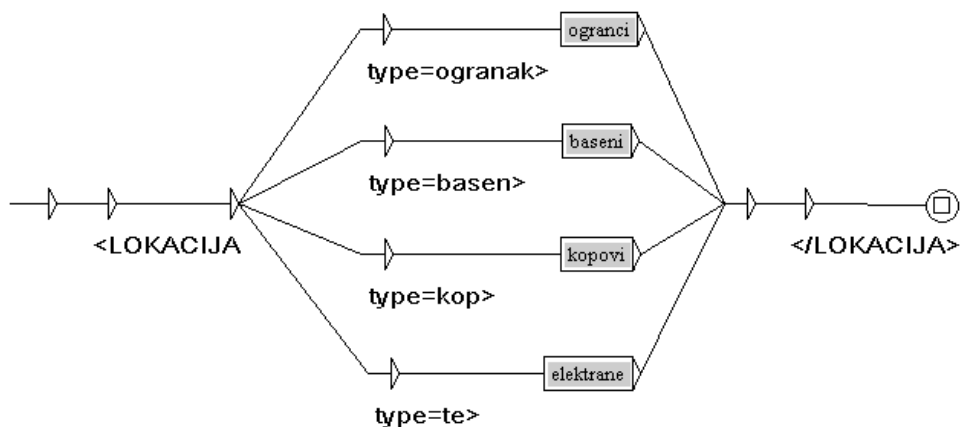
Promenljiva <tip> prepoznaje i obeležava deo teksta koji upućuje na tip bagera, a lista tipova je preuzeta iz ontologije RudOnto (Stanković et al. 2012; Kolonja et al. 2016), dok se sam naziv, ponovo, može prepoznati na dva načina. Prvi način je prepoznavanje po oznakama vrste (naziv vrste, skraćenica naziva vrste ili početno slovo) praćeno nekom brojnom oznakom (npr: G1 ili Glodar-2). Drugi način je na osnovu modela bagera praćeno brojnim ili dimenzionalnim oznakama (EŠ 6/45/I, SchRs 1600x25/3, ERs 1000...) Ovim grafom se uspešno prepoznaju različite kombinacije pominjanja poput:

- dreglajna EŠ-6/45-II
- Rotorni bager SchRs 630 (Gl-1)

- bager Vedričar eRs 1000
- rotorni bager G2

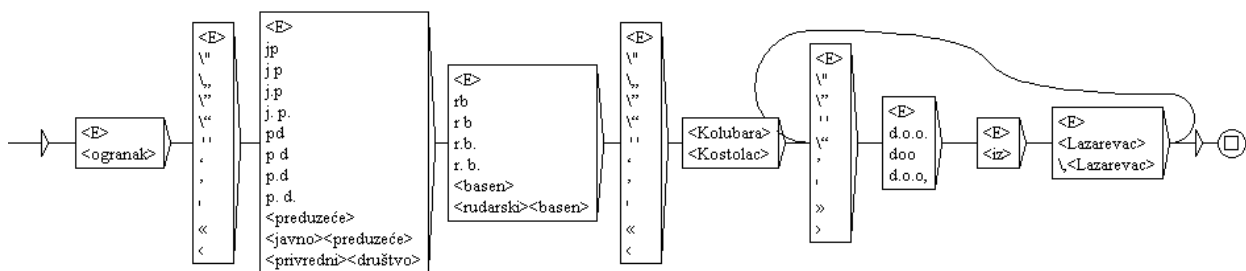
4.3.2. Entiteti rudarskih lokacija

Entiteti rudarskih lokacija su podeljeni grubo u četiri kategorije: ogranci, baseni, kopovi i elektrane. Za svaku kategoriju pripremljen je zaseban podgraf za prepoznavanje, a sva pojavljivanja su obeležena pomoću transduktora za lokacije (Slika 4-16). Transduktor obeležava sve prepoznate entitete etiketom <LOKACIJA>, i za svaku kategoriju dodeljuje i zaseban atribut tip.



Slika 4-16 Graf transduktor za prepoznavanje i obeležavanje lokacija

Podgraf koji obeležava preduzeća odnosno institucije (Slika 4-17) je vrlo linearan. Entitet počinje opcionom imenicom ogranak, zatim može ili ne mora imati oznake javnog preduzeća, privrednog društva i/ili rudarskog basena. Jedini obavezan deo je naziv ogranka (Kolubara ili Kostolac). Nakon toga mogu uslediti i oznake za društvo sa ograničenom odgovornošću – d.o.o. i mesto Lazarevac u kome je sedište JPRB Kolubara. Na određenim mestima, oko ličnih imena, mogu se, ali ne moraju, naći i navodnici. Pravni oblik današnjih ogranaka JP EPS se menjao i u tekstovima koji datiraju iz različitih vremenskih perioda se fizički isti entiteti mogu naći pod različitim nazivima, što se mora podržati razvijenim rešenjem.

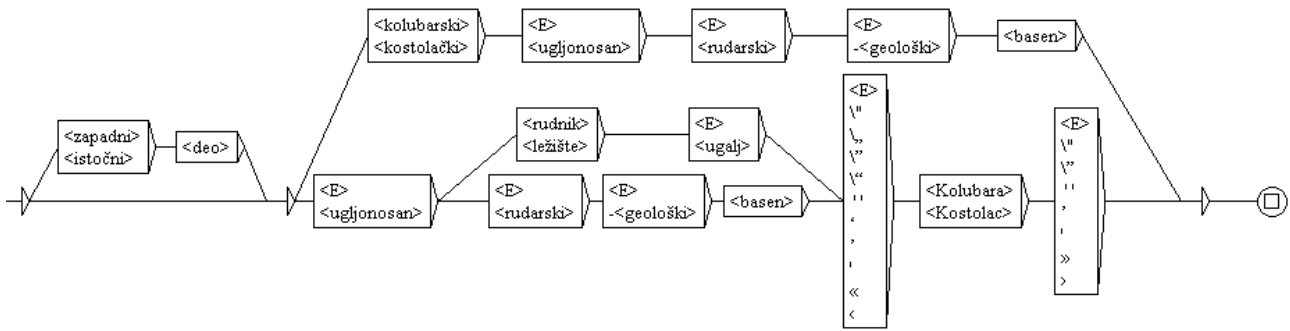


Slika 4-17 Podgraf za obeležavanje rudarskih preduzeća (institucija)

Baseni su prepoznati slično kao institucije, sa tim da se ovde radi o geografskim lokacijama a ne pravnim licima (Slika 4-18). Umesto oznaka za preduzeća uz Kolubaru i Kostolac traže se opcione oznake rudnik/ležište uglja, ili ugljonosni rudarsko geološki basen, sa svim atributima kao opcionim. Neki od entiteta prepoznatih ovim grafom su:

- Kolubarski ugljonosni basen
- Rudarsko-geološki basen „Kolubara“

- ležištu uglja “Kostolac”
- Zapadni deo Kolubarskog ugljonosnog basena

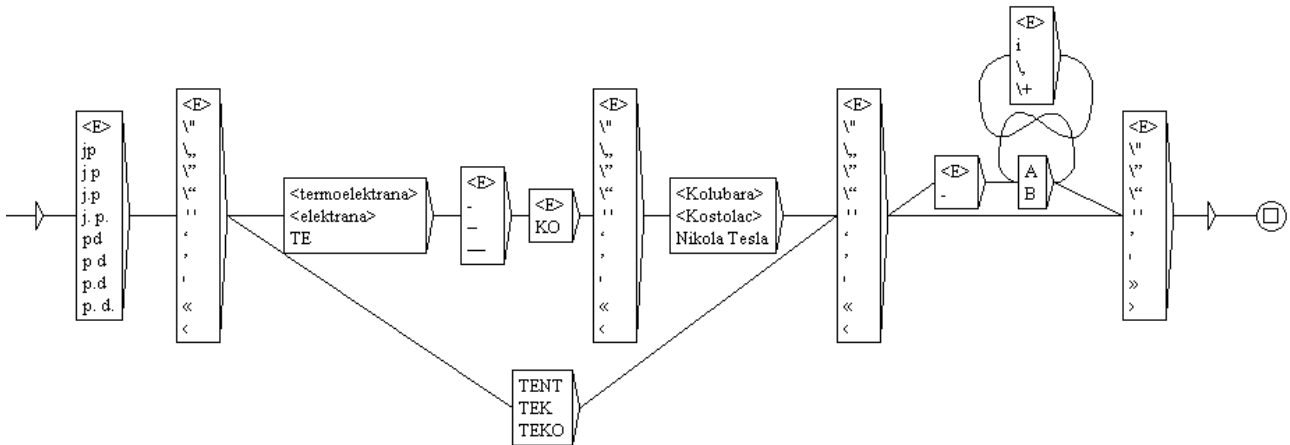


Slika 4-18 Podgraf za obeležavanje rudarskih basena kao lokacija

Treći podgraf prepoznaje entitete termoelektrana (Slika 4-19), uključujući i njihovo detektovanje kao pravnih lica, a sadrži i elemente grafa koji prepoznaje basene kao institucije (Slika 4-17). Razlika je što su obavezne reči sada pored ličnog imena i elektrana, termoelektrana ili njihove nedvosmislene oznake poput TENT (termoelektrana Nikola Tesla) ili TEK (termoelektrana Kostolac).

Uz obeležja preduzeća, elektrane i naziva, mogu su u entitetu naći i približnije oznake termoelektrana, gde postoje ogranci. Prepoznati nazivi podgrafa su:

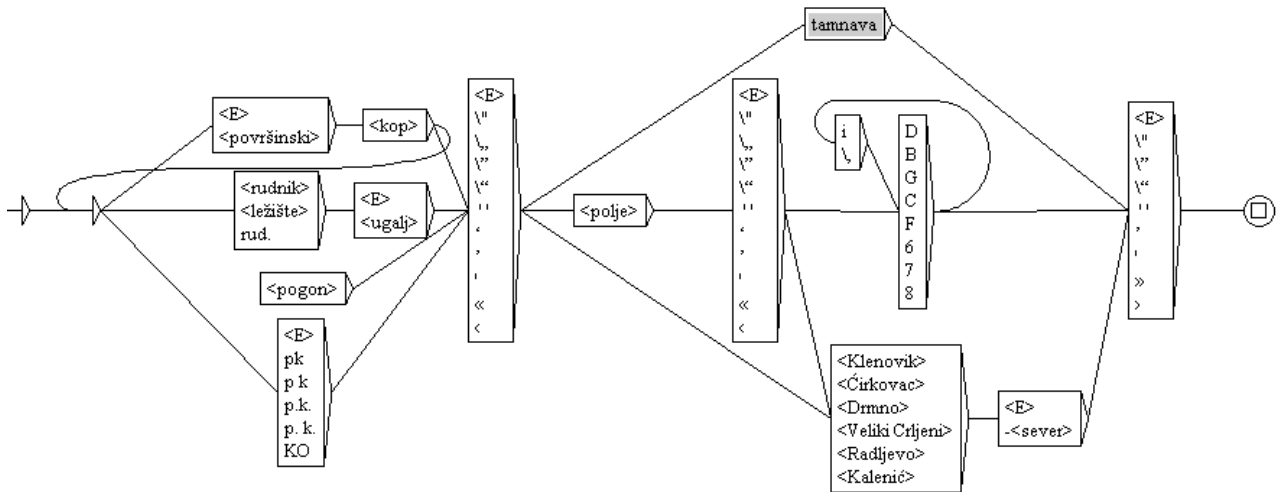
- TE Nikola Tesla
- TE "Nikola Tesla B"
- TE-KO, Kostolac
- PD TE-KO "Kostolac"



Slika 4-19 Podgraf za obeležavanje termoelektrana kojima se isporučuje uglj

Podgraf koji prepoznaje površinske kopove i izvorišta kao mesta sa kojih se transportuje uglj (Slika 4-20) takođe zavisi od detektovanih imena kopova. Uz naziv opciono mogu ići rudnik/ležište (uglja), (površinski) kop, i skraćenice za iste. Eksploataciona polja mogu biti obeležena slovima ili brojevima, tako da ovaj graf prepoznaje entitete:

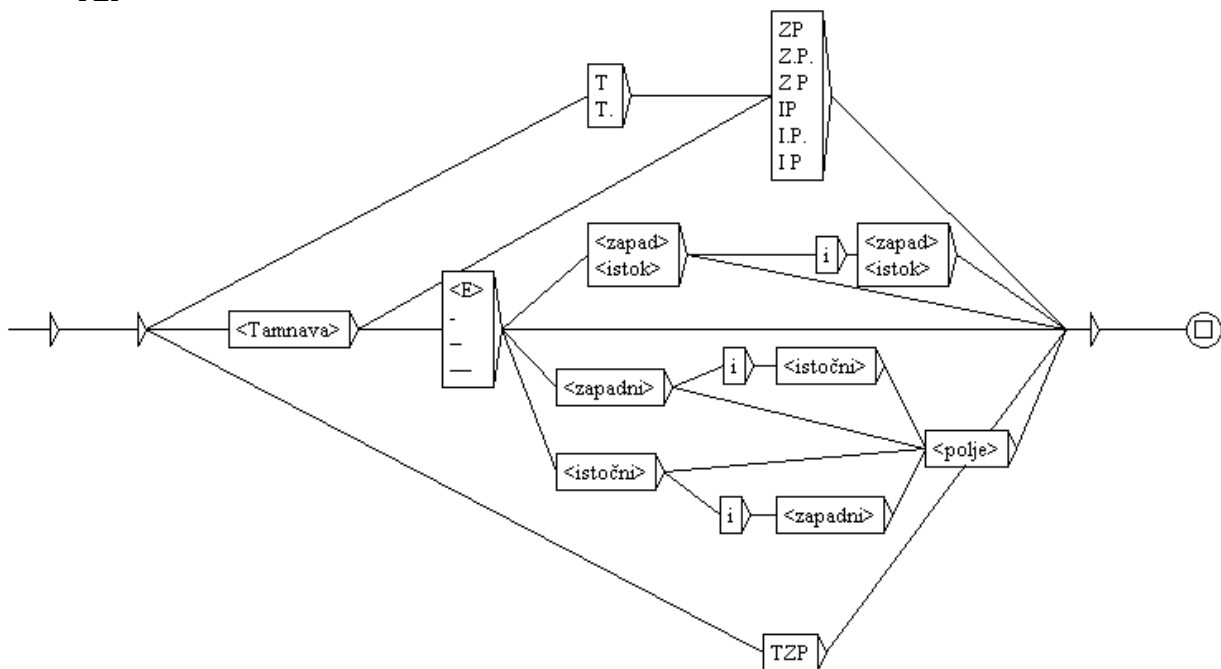
- kop Polje G
- površinski kop „Radljevo-sever“
- PK „Veliki Crljeni“
- Površinskog kopa „Polje C“



Slika 4-20 Podgraf za obeležavanje rudarskih kopova i izvorišta

Poseban je slučaj izvorišta Tamnava zapadno i istočno polje. Zbog velike kompleksnosti u prepoznavanju ovog entiteta, za njega je pripremljen poseban graf (Slika 4-21). Ovaj graf je proširena verzija grafa koji prepoznaje ovaj entitet iz rada (Tomašević 2018) i dopunjava prepoznavanje kopova različitim oblicima entiteta ovog polja, poput:

- Tamnava-Zapadno polje
- „Tamnava Istočno polje”
- TZP



Slika 4-21 Specijalan slučaj - Tamnava i njena polja

4.3.3. Entiteti otkaza opreme i kvarova

Deo korpusa koji sadrži informacije o kvarovima je analiziran, i sačinjen je tekst od liste opisa najčešćih kvarova u dvogodišnjem periodu. Na izdvojenom tekstu urađena je lematizacija i određivanje vrste reči, kao i najčešće kolokacije, kako bi se pronašli „okidači“ u tekstu koji upućuju da se radi o opisu neke neželjene situacije na rudniku.

Prvo je pronađen skup imenica koje mogu da „okinu“ uzrok, a to su: *kvar, ispad, otkaz, zastoј, stajanje, opravka, ...*

- **glagolske imenice:** *stajanje, oštećenje, zavarivanje, centriranje, nameštanje, isključenje, napajanje, zatezanje, spuštanje, odvodnjavanje, dizanje, kanalisanje, kretanje, skraćanje, prebacivanje, spadanje, razvezivanje, zabacivanje, pomeranje, ispumpavanje, izdizanje, varenje, hlađenje, podmazivanje, uključenje, habanje, curenje, vešanje, čišćenje, odsecanje, pražnjenje, podešavanje, punjenje, premeštanje, izmeštanje, usmeravanje, postavljanje, paljenje, grejanje, obaranje, sređivanje, podizanje, merenje, vraćanje, uklanjanje, zatvaranje, povezivanje, upravljanje, osiguranje, formiranje, otvaranje, proklizavanje*
- **imenice:** *buldozer, vedrica, odlagač, zemljospoj, bager, kofica, rolna, reduktor, drobilana, reza, transporter, bunker, spojnica, bubanj, rez, katarka, otprema, kabl, vitlo, koturača, kvar, ležaj, napon, kolica, magla, hod, okret, demontaža, namotavanje, rudnik, ćelija, točak, sonda, osovina, kompresor, gusenica, sajla, zupčanik, sneg, produžetak, planum, spojnik, trafo, pogon, kopča, etaža, drobilica, traka, kiša, motor, graničnik, vetar, trkač, hidraulika, lanac, regulator, blok, gumica, transport, ploča, zavrtnj, ljudstvo, namotaj, deponija, otpornik, osovina, bagerista, gazište, protok, obloga, levak, optički kabl, polumesec, transformator, membrana, homogenizacija*
- **pridevi:** *radni, gusenični, prekostrujni, blokadni, utovarni, pretovarni, uljni, bagerski, hidrodinamički, istovarni, usmeravajući, obrtni, havarijski, stezni, zaptivni, zatezni, pogonski, vremenski, neispravan, čeoni, otpornički, odlagališni, ispao, dalekovodni, srednjenaponski, linijski, potisni, bočni, neispavan, zamrzao, visokonaponski, povratan, komandni, hidrauličan, smeštajni, ugljeni, zglobni, povišen, popucao, desni, prijemni, viseći, strujni, stacionaran, klizni, frekventan, merni, poprečan, pokidan*
- **glagoli:** *bubnjati, rezati, otpremiti, vitlati, kositi, grudvati, kopati, okretati, šinuti, čekićati, prstenovati, presipati, povisiti, kaišati, usloviti, puknuti, popucati, popustiti, zalediti, pokidati, iščupati, upaliti, ukočiti, zavesiti, reagovati, spojiti, pritiskati, stojati, izgoreti, metati, stajati, ograničiti, pomeriti, ispasti, udarati, udariti, produžiti, propasti, oštetiti, spasti, stareti, opremiti, angažovati, stariti, zameniti, vratiti, zaštititi, raditi, voditi, pripremati, izazvati, pasti, pratiti, postaviti*

Primeri specifičnih izraza i kolokacija:

- **kvar** { pogona, dizanja, protoka, uključenja, kompresora, blokade, motora, računara, spuštanja, komunikacije, kretanja, transporta, zatezanja, trkača, reduktora, sonde, pretvarača, transformatora, oštećenja, hidraulike, rasvete, spojnice, }
- **kvar** {-oštećenje, -zamena, -ispad, -kontrola, -sanacija, -montaža, zemljospoj, varenje, opravka, demontaža, -zavarivanje, nameštanje, -spadanje, -detekcija, pretvarač, -vraćanje, transport, -greška, -hlađenje, -podešavanje, }
- **kvar** {spuštanja katarke, protoka ulja reduktora, protoka ulja u reduktoru, komandnog napona, zatezanja trake, motora pogona, pogona vedrica, transporta bagera, uključenja komandnog napona, deformacije sajle, računara na bageru, trkača na gusenici, frekventnog pretvarača, desnog motora, sonde protoka, hidrodinamičke spojnice, pogona regulatora, pretvarača intervencije, kretanja pogona, oštećenja ležaja, uljne pumpe, pogona motora, dizanja kabine, komunikacije kontrola, sonde protoka ulja, pogona vedrica bagera, računara za upravljanje, transporta utovarnih kolica, pogona obrtnog tanjira, motora uljne pumpe, pretvarača intervencije na vedricama}
- **ispad** {ćelije, napajanja, napajanja ćelije, napona, kolica, trkač, zavesica, kontaktora, kablova, ćelija, zavesica}
- radni točak, gusenična ploča, blokadni kabl, uljna instalacija, pretovarna mašina, vulkanizerska kontrola, obrtni transporter, komandni napon, pogonski bubanj, raden element, radni element, radan element, bočna zavesica, obrtni tanjir, uljna pumpa,

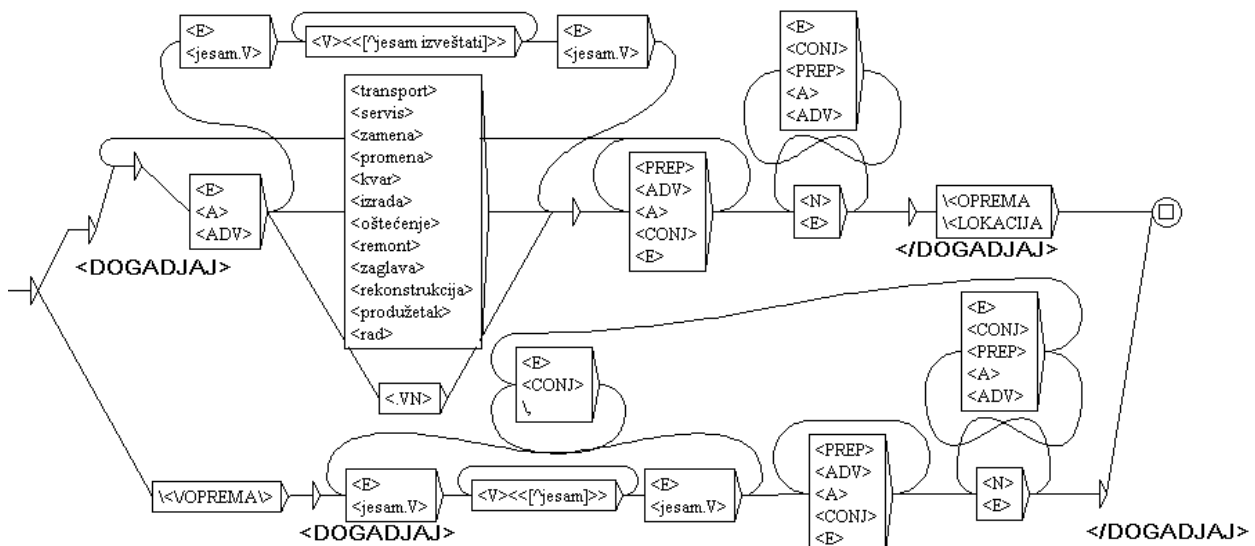
prijemna traka, prijemni trak, havarijsko isključenje, istovarni bubanj, obrtni levak, optički kabl, optički kabao, čeoni bubanj, mašinska opravka, smeštajni kapacitet, ugljena stanica, tehnološko stajanje, proizvodna mehanizacija, bočna zavesa, istovarna katarka,

- transport bagera, trasa transportera, regulacija trase, bager intervencije, promena bloka, bager promena, lanac vedrica, ispad ćelije, servis sistema, izrada sastava, vedrica bagera, pogon drobilane, ploča bagera, rt bagera, transport odlagača, remont bagera, mašina bagera, oštećenje trake, izrada fleke, gradnja bagera, remont sistema

4.3.4. Entiteti događaja na kopovima

Izrada sledećeg grafa je imala za cilj prepoznavanje događaja u kontekstu pojava vezanih za prethodno prepoznate lokacije i opremu (Slika 4-22). Graf pokušava da pronađe glagole ili specifične imenice u blizini pojavljivanja opreme i lokacija, uz opcione druge elemente rečenice.

Kako je neophodno imati tekst u kom je prethodno obeležena oprema i lokacije, za svrhu ovog eksperimenta napravljena je kaskada transduktora (Friburger 2002). Najpre se pronalaze oprema i lokacije, pa se tek onda aktivira graf za događaje.



Slika 4-22 Graf za prepoznavanje događaja vezanih za opremu i lokacije

Ovaj graf se ponaša kao dva odvojena transduktora, od kojih jedan traži pojavljivanje događaja koji u tekstu prethodi opremama i/lokacijama, dok drugi traži događaje koji slede nakon pojave naziva opreme.

Prvi transduktor prepoznaje i obeležava entitete poput:

- upravljanje kvalitetom uglja na <LOKACIJA type=polje>površinskom kopu Tamnava - Zapadno polje</LOKACIJA>
- ugradnju peska sa<LOKACIJA type=polje>PK TZP</LOKACIJA>
- Izvršen je pregled <OPREMA><sistem><tip>BTD</tip>-<naziv>1</naziv></sistem></OPREMA>

dok drugi prepoznaje

- <OPREMA><ostalo>transporter</ostalo></OPREMA> stoji
- <OPREMA><ostalo>Odlagač-2 -</ostalo></OPREMA> gorela tapacirana rolna
- <OPREMA><tip>rotorni bager</tip></OPREMA> je kopao visinski blok duž<OPREMA><ostalo>transportera E1</ostalo></OPREMA>

Za ekstrahovane instance se uvodi relacija ekvivalencije „same as“ ukoliko su u pitanju isti primerci opreme, ukoliko su tako ekstrahovani ili se definišu leksičke varijante za jednu instancu.

Primer izgleda teksta obeleženog XML etiketama, u mašinski čitljivom obliku, donosi Slika 4-23, gde se mogu uočiti obeležene sekvence teksta koje sadrže instance opreme, događaja, mera i lokacija. Može se videti da su dozvoljene i ugnježdene XML etikete.

```
Tehnologija otkopavanja, <DOGADJAJ>transporta i odlaganja do polovine meseca </DOGADJAJ>
<OPREMA><tip>rotorni bager</tip></OPREMA><DOGADJAJ> je kopao visinski blok duž</DOGADJAJ>
<OPREMA><ostalo>transportera E1</ostalo></OPREMA> prvo u zoni povratne stanice, a potom se
<DOGADJAJ>okrenuo i otkopavao visinski blok prema pogonskoj stanici</DOGADJAJ>
<OPREMA><ostalo>transportera E1</ostalo></OPREMA>.
<DOGADJAJ>Potom je vršeno radijalno pomeranje</DOGADJAJ> <OPREMA><ostalo>transportera E1</ostalo></OPREMA>
ka jugu i <OPREMA><ostalo>transporter</ostalo></OPREMA> <DOGADJAJ> je postavljen u novi položaj</DOGADJAJ>.
<OPREMA><tip>Rotorni bager</tip></OPREMA> je <DOGADJAJ>istranspotovan</DOGADJAJ> na plato kod stanice E1.
Nova dužina <OPREMA><ostalo>transportera E1</ostalo></OPREMA> je 1300 m.
<OPREMA><tip>Bager</tip></OPREMA> je iz novog položaja <OPREMA><ostalo>transportera E1</ostalo></OPREMA>
<OPREMA><ostalo>Odlagač</ostalo></OPREMA> je od početka meseca <DOGADJAJ>odlagao visinski
blok sa visinske strane</DOGADJAJ><OPREMA><ostalo> transportera O1</ostalo></OPREMA>.
Polovinom meseca <OPREMA><ostalo>odlagač</ostalo></OPREMA> je preko<OPREMA><ostalo> transportera
O1</ostalo></OPREMA><DOGADJAJ> prešao na dubinsku stranu i odlagao je dubinski blok iz
pune širine duž</DOGADJAJ> <OPREMA><ostalo>odlagališnog transportera O1</ostalo></OPREMA>.
<MERE>Postaviti zaštitnu ogradu na svim mostovima za prelazak vozila</MERE> preko
<OPREMA><ostalo> transportera</ostalo></OPREMA>, a sa obe strane
<MERE>postaviti tablu sa maksimalno dozvoljenom nosivošću</MERE>, kao i širinom mosta.
<DOGADJAJ>Angažovanje Vatrogasne jedinice</DOGADJAJ> <LOKACIJA type="polje">Tamnava</LOKACIJA>
<OPREMA><naziv>Eš-2</naziv></OPREMA> -<DOGADJAJ>gorelo ispod motora pogona</DOGADJAJ>
<MERE>Na mestima na kojima se nalaze uređaji za čišćenje <OPREMA><ostalo>trake</ostalo></OPREMA>
mora se sprečiti nagomilavanje skinutog materijala. Skinuti materijal, ne sme se ručno
nabacivati na <OPREMA><ostalo>traku</ostalo></OPREMA> kada je
<OPREMA><ostalo>traka</ostalo></OPREMA> u pokretu</MERE>.
<OPREMA>Materijal koji <DOGADJAJ>padne sa</DOGADJAJ> <OPREMA><ostalo>transportera</ostalo></OPREMA>,
a naročito materijal koji se <DOGADJAJ>skuplja ispod donjeg kraka</DOGADJAJ>
<OPREMA><ostalo>trake</ostalo></OPREMA>, mora se redovno čistiti.
Pri <OPREMA><ostalo>radu transportera</ostalo></OPREMA> noseći i povratni valjci moraju se okretati.
<MERE>Oštećeni valjci i valjci čiji se ležajevi pri radu zagrevaju moraju se zameniti ispravnim
valjcima ili privremeno izbaciti iz funkcije</MERE>.
```

Slika 4-23 Tekst obeležen XML etiketama

Obeleženi tekst se može vizuelizovati (Slika 4-24) u alatu za anotaciju teksta *brat*⁶⁹ (Stenetorp et al. 2012)



Slika 4-24 Vizuelizacija obeleženog teksta

⁶⁹ <https://brat.nlplab.org/>

4.3.5. Organizovanje pojmova u hijerarhiju

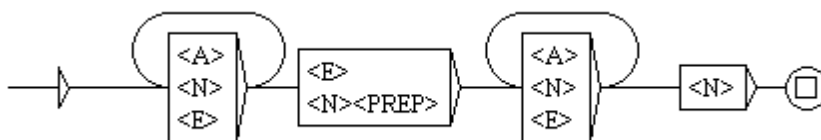
Sinonimi i klasifikacije koncepata koji su postojali u RudOnto, Termi i SrpWN-u su preuzete kao osnovni izvor, a potom je urađena dopuna poluautomatskim pristupom. Za popunjavanje su korišćeni različiti resursi, kao što su različite klasifikacije u knjigama ili drugim relevantnim izvorima.

Za dopunjavanje ontologije se takođe koriste relacije sinonimije, hiponimije i hiperonimije. Na osnovu sinonima se u ontologiji uspostavlja relacija „same_as“. Relacije podređen - nadređen, odnosno hiponim - hipernimi pomažu da se pojmovi organizuju u hijerhijsku strukturu, odnosno u taksonomiju pojmova. Analiza terminoloških rečnika je pokazala da se većina termina sastoji od više reči, ne samo u srpskom već i u drugim jezicima, tako da su već duži niz godina istraživani obrasci koji bi bili dovoljno jednostavni da budu deo opšteg sistema za prepoznavanje relacija među konceptima, ali i dovoljno precizni da prepoznaju upravo imeničke fraze. Jedan od obrazaca koji se pokazao uspešan za engleski jezik daju (Justeson and Katz 1995):

$$((Adj/Noun)+|((Adj/Noun)*(Noun Prep)?)(Adj/Noun)*)Noun,$$

Gde je: Adj pridev, Noun imenica, Prep predlog, | disjunkcija, + označava pojavljivanje jednom ili više puta, ? opciono pojavljivanje jednom, * opciono pojavljivanje više puta.

Grafovi za prepoznavanje sinonima, hiponima i hiperonima koriste NP (eng. noun phrase) podgraf, koji prepoznaje imenice i imeničke fraze. Graf prikazuje Slika 4-25,.



Slika 4-25 Graf koji prepoznaje imeničke fraze

Graf se interpretira na sledeći način: imenica kojoj prethode bilo koji broj uzastopnih imenica i prideva, uz moguću pojavu predloga nakon jedne od imenica. Primeri rezultata grafa su:

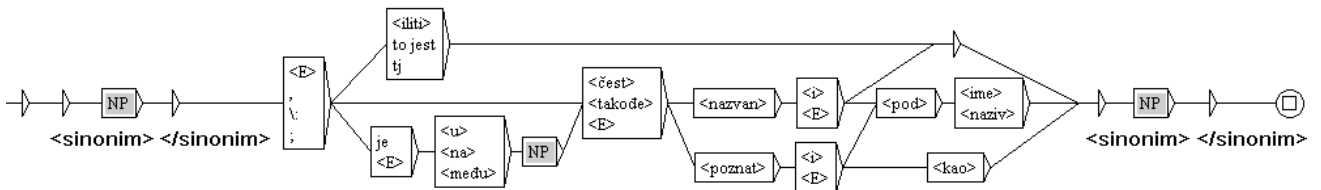
- *analitičkih metoda*
- *havarijske situacije u radu mašine*
- *negativan uticaj na životnu sredinu*
- *otkazima na drobilici*
- *zaposlene radnike*
- *životnu sredinu u blizini rudarskih kompleksa*

Osim dobrih primera, na žalost pronalaze se i nevalidni oblici koji su najčešće posledica pojavljivanja reči sa višeznačnim oblikom, tako da pridružena lema nije željena, poput:

- *dovode do rizičnog događaja* (glagol *dovoditi* prepoznat kao imenica *dovod*)
- *kod Petrovca na Mlavi* (predlog *kod* prepoznat kao imenica *kod*)
- *rade u složenim uslovima* (glagol *raditi* prepoznat kao pridev *rad*)

Prepoznavanje sinonima iz korpusa je inspirisano ranijim istraživanjima za engleski jezik (Hearst 1992; Tovar et al. 2018; Roller, Kiela, and Nickel 2018), potom za srpski (Krstev, Vitas, and Stanković 2015).

Graf za prepoznavanje sinonima (Slika 4-26) prepoznaje dve imenske fraze koje između sebe imaju određenu sponu. Spona počinje opcionim znakom interpunkcije, i nastavlja se konstrukcijom koja naznačava da sledi sinonim, tj. da naredna fraza označava drugo ime ili naziv prethodno navedenog pojma.



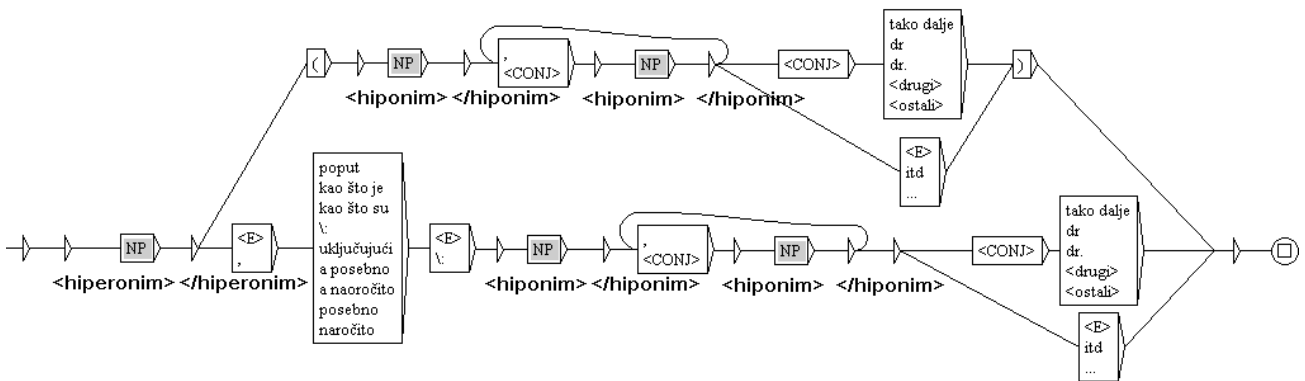
Slika 4-26 Graf koji ekstrahuje sinonimske imeničke fraze iz teksta

Primeri spona su:

- iliti
- takođe nazvan i
- poznat i kao
- je među narodom često poznat i pod imenom

Ovaj graf se nije pokazao preterano dobro, vrativši mnogo loših (preko 90%) i svega nekoliko uspešno prepoznatih parova sinonima, poput:

- *odušne otvore, to jest otvore na zidovima*
- *priručniku, to jest uputstvu za formiranje tarifa*



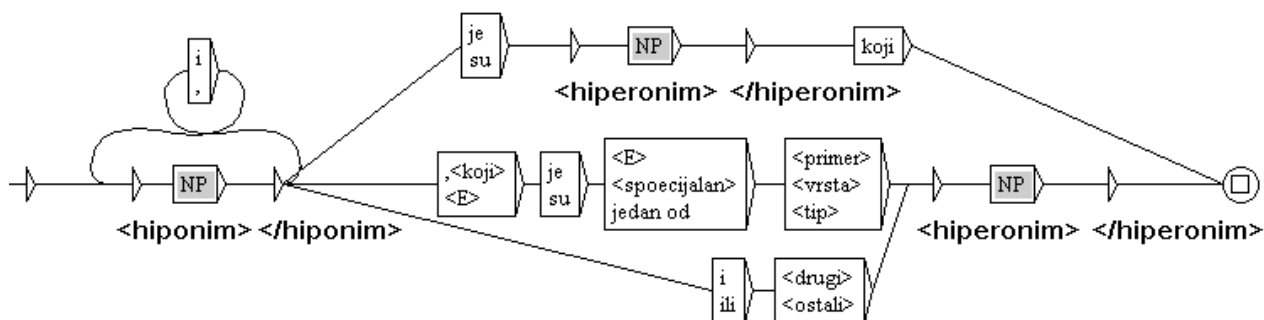
Slika 4-27 Graf koji iz teksta ekstrahuje hiperonime sa listama hiponima

Na osnovu analize obrazaca za engleski jezik (Roller, Kiela, and Nickel 2018), razvijen je graf za prepoznavanje hiperonima sa listom hiponima (Slika 4-27) ima veliki odziv, uz umerenu preciznost. On se odlikuje pojavljivanjem imenske fraze (hiperonim), koja je u određenom okviru praćena jednom ili više imenskih fraza (hiponima). Primeri okvira su :

- **hiperonim**, kao što su *hiponim* i *hiponim*
- **hiperonim** (*hiponim, hiponim, hiponim* itd)
- **hiperonim**, uključujući *hiponim, hiponim* i ostale
- **hiperonim**: *hiponim, hiponim, hiponim...*

Neki od primera koje ovaj graf ekstrahuje su:

- **Kritične posledice**: *ozbiljne povrede, šteta na postrojenju i totalni prekid proizvodnje*
- **kvar** (*otkaz, neželjeni događaj*)
- **Manja povreda** (*modrica, razderotina*)
- **metoda internog smanjenja rizika** (*diverzifikacija i investiranje u informacije*)
- **brojna ostrva** kao što su *Kiribati, Maršalska ostrva i Maldivi, gusto naseljena obale*
- **Vrste štetnog događaja** (*eksplozija, požar, curenje iz cevovoda*)



Slika 4-28 Graf koji ekstrahuje hiponime i njihov hiperonim

Drugi graf za prepoznavanje hiponima i hiperonima polazi sa tačke hiponima (Slika 4-28). On prepoznaje hiponim (ili listu hiponima), koji je u određenom okviru praćen heteronimom. Primeri struktura koje ovaj graf prepoznaje su:

- *hiponim* i drugi **hiperonim**
- *hiponim* je **hiperonim** koji
- *hiponim, hiponim* i ostali **hiperonim**
- *hiponim* je jedan od primera **hiperonim**

Ovaj graf ima manji odziv, ali veću preciznost. Neki od ekstrahovanih primera su:

- *advokatski troškovi* i drugi **troškovi odbrane**
- *banke* i druge **međunarodne organizacije**
- *Elementarni otkaz* je **prost događaj** koji
- *Lekarima, advokatima, računovođama, inženjerima* i drugim **profesionalcima**
- *Protivpožarna vrata, klapne* i druge **prepreke**
- *Mašine* i ostalu **opremu**

4.4. Ekstrakcija relacija između koncepata

Oslanjajući se na obrasce prikazane u odeljku 2.5 istražene su mogućnosti ekstrakcije relacija nad korpusom RudKorp koji je publikovan u razvojnom okruženju SketchEngine⁷⁰, a upiti su postavljeni korišćenjem jezika CQL⁷¹, o kojima je više reči bilo u odeljku 3.2.2.

4.4.1. Okidač „kao što je/su“

Za ekstrakciju okidačem „**such as**“, odnosno na srpskom „**kao što je/su**“, odgovarajuće rečenice bi se pronašle obrascem: „*NP_o such as NP₁, NP₂ ... (and | or) } NP_n*“, gde NP označava *noun phrase*, odnosno imeničiku frazu. Za takve NP_i, 1 < i < n možemo da uspostavimo relaciju hiponimije sa NP_o.

⇒ for all NP_i, 1 < i < n, hyponym(NP_i, NP_o)

Na osnovu primera na engleskom: „*The **bow lute**, such as the **Bambara ndang**, is plucked and has an individual curved neck: for each string*“, se izvodi relacija:

⇒ hyponym ("Bambara ndang", "bow lute")

Primer upita nad korpusom koji bi odgovarao sličnom obrascu je:

```
[lemma="proces"] [word!="\."]{0,3} [word="kao"] [word="što"] [lemma="jesam"]
```

⁷⁰ <https://www.sketchengine.eu/>

⁷¹ <https://www.sketchengine.eu/documentation/corpus-querying/>

Gde se „[word!="\."]“ omogućava da se u tekstu mogu javiti neke reči ali ne i tačka. prikazuje deo rezultata varijante prethodnog upita za NP₀="proces" (Slika 4-29).

ili segmentnim rudničkim procesima kao što je čišćenje uglja, od ključnog je značaja za ostvarivanje očekivanih proizvodnih performansi. Pažnje usmerene ka procesima kao što su homogenizacija, usitnjavanje i klasiranje, čišćenje uglja i flotacijski procesi. I poraglasiti da modeli složenih procesa kao što je flotacijska koncentracija nikada ne mogu u potpunosti preslikavati stvarni proces. Šteta u praksi to uključuje različite procese kao što su briketiranje uglja i drugih sitnozrnatih materijala, formiranje agregata pod delovanjenjem i sinterovanja. Uključuje i procese kao što su peletizacija koncentrata obojenih metala bakra, olova, cinka itd. Okrupnjavanje se odnosi na kritične poslovne procese kao što je ispunjenje porudžbine kupca ili razvoj proizvoda. 4. 3. 7 Nivo organizovanja i upravljanja rizikom, kod npr. složenih procesa kao što su velika hemijska postrojenja • Padati negde između gornje dve krajnosti. Za očekivano

Slika 4-29 Konkordance za upit „<proces>⁷² kao što <jesam>“

Na osnovu prikazanih rezultata upita možemo dodati u ontologiju sledeće relacije :

hyponym ("proces", "čišćenje uglja").
 hyponym ("proces", "homogenizacija uglja")
 hyponym ("proces", "usitnjavanje uglja")
 hyponym ("proces", "klasiranje uglja")
 hyponym ("proces", "flotacija")

Primer prigodne klasifikacije koja je izvedena iz rezultata za NP₀="rizik", odnosno upita [lemma="rizik"][word!="\."]{0,5}[word="kao"][word="što"][lemma="jesam"] daje Slika 4-30. Gornji primer upućuje na klasifikaciju rizika, dok donji govori o opsegu štete, tako da je jasno da ovakav pristup ne sme da potpuno automatizuje već je potrebna validacija, filtriranje i korekcija automatski izvučenog sadržaja, pa se stoga govori o poluautomatskoj ekstrakciji (Slika 4-30).

resursi, kvalitet i tako dalje. Takođe, rudarski projekti, posebno projekti otvaranja površinskih kopova uglja, nose veliki broj nepoznanica koje sa sobom nose velike rizike, kao što su prirodni, geološki, ekološki, sociološki, tehničko-tehnoški, ekonomski i istovremeno su karakteristični po velikom obimu, širokoj strukturi, dugom vremenu trajanja i velikom broju učesnika u njihovoj realizaciji. Ovakva složenost i stohastičnost rudarskih projekata u fazi otvaranja površinskog kopa, praćena velikim rizicima

Korisno je razmotriti mogući opseg štete od nekog uočenog rizika kao opseg posledica kao što su : • Manja šteta • Incident bez povrede • Manja povreda (modrica , razderotina) • Ozbiljna povreda (fraktura , amputacija , hronična bolest) • Smrt • Slučaj višestruke smrtnosti I verovatnoća moguće štete se može pojaviti u sledećem nizu mogućnosti : • Neverovatna • Moguća (ali ne vrlo) • Moguća • Neizbežna (u određenom vremenskom razdoblju) 4 . Odluka o uvođenju novih mera (ako je potrebno) za

Slika 4-30 Konkordance za upit „rizik ... kao što...je“

Iterativnim postupkom se može doći do odgovarajućih upita, naime upit se može prilagoditi da izvuče štete okidačima, recimo „što inicira|prouzrokuje|...“. Dakle potrebne su takve rečenice i odgovarajući upiti kojima će se precizno pronaći koncepti i uspostaviti relacije među njima.

4.4.2. Okidač „je vrsta“

Za ekstrakciju okidačem „is a (kind of)“, odnosno na srpskom „je vrsta“, za obrazac:

⇒ An L0 is a (kind of) L1

relacija u ontologiji bi se uspostavila

⇒ hypernym(L1, L0)

⁷² Reč u uglastim zagradama predstavlja sve gramatičke oblike reči, na primer: „<proces>“ zamenjuje „proces, procesa, procesu, procesima, ...“

Prilagođavanjem izraza za srpski, primer upita nad korpusom koji bi odgovarao sličnom obrascu je:

```
[lemma="jesam"] [word!="\." ] {0,3} [lemma="vrsta"] [tag="N"]
```

Analizom rezultata (Slika 4-31) je uočeno da dosta primera za nabranjanja idu sa zagradom tako da se može upit modifikovati u tom smeru.

sistematizovana su saznanja u ovoj oblasti kao i rezultati dosadašnjih istraživanja . Za potrebe analize i definisanja željene metodologije izvršena su ispitivanja u konkretnim uslovima radne sredine na Površinskom kopu Gacko , gde **su prvenstveno ispitivani uticaj vrste reza** (**horizontalni** , **vertikalni** , **kombinovani**) i promene parametara odreska (debljina , širina i visina odreska i njihovi međusobni odnosi) na otpore kopanju i dinamičko ponašanje bagera , za rotorni bager ER - 1250 16 / 1 , 5 sa

Slika 4-31 Konkordance za upit „je ... vrsta ...“

Za ovaj primer kandidati relacija za ontologiju su:

```
hyponym („rez“, „horizontalni rez“)  
hyponym („rez“, „vertikalni rez“)  
hyponym („rez“, „kombinovani rez“)
```

Analizom rezultata je uočeno da dosta primera za nabranjanja idu sa zagradom tako da se može upit modifikovati u tom smeru.

4.4.3. Okidač „drugi, ostali, “

Za ekstrakciju izrazom „NP { NP}*{ } { and | or } other NP“ u engleskom tekstu „Bruises, wounds, broken bones or other injuries ...“ se mogu ekstrahovati sledeće relacije:

- hyponym("bruise", "injury"),
- hyponym ("wound", " injury "),
- hyponym("broken bone", "injury")

Iz teksta „... temples, treasuries, and other important civic buildings.“ Se mogu ekstrahovati:

- hyponym("temple", "civic building"),
- hyponym("treasury", "civic building")

U srpskom bi trigeri bili „drugi“ i „ostali“, ali se u praksi pokazalo da se oni mogu zaobići, a oblik NP često nije kanonski. Analizirajmo sledeći upit i njegove rezultate (Slika 4-32).

```
( [pos="A|N"] * [pos="N"] [word=" , |i|i|i"] ) + [word="i|i|i"] [lemma="drugi|ostali"  
] [pos="A|N"] * [pos="N"]
```

dinamike . Za potrebe eksploatacije mineralnih sirovina , ili tokom izvođenja podzemnih radova za **potrebe izrade podzemnih skladišta , skloništa , i drugih vrsta podzemnih objekata** , naročito na većim dubinama , remeti se postojeće naponsko stanje u okolnoj stenskoj masi . U slučaju postojanja raseda ili sličnog mehaničkog (ili litološkog) diskontinuiteta u širem području

napreznja na različitim dubinama stijenskog masiva . Te razlike napona u zemljinoj kori u različitim regionima i dubinama zavise od tektonskih aktivnosti koje su se dešavale u zemljinoj usled **pojave zemljotresa , djelovanja vulkana i formiranja planinskih vjenaca , i drugih aktivnosti** . Da bi se mogle približno utvrditi i pratiti te promjene koristi se instrumentalno praćenje i proračuni

Slika 4-32 Konkordance za upit „ ... |i|i|i...drugi|ostali...“

Sličani upiti takođe mogu dati relacije:

```
[pos="A|N"] * [pos="N"] [word="uključujući|podrazumevajući|naročito|specijalno  
|posebno"] ( [pos="A|N"] * [pos="N"] [word=" , |i|i|i"] ) +  
[pos="A|N"] * [pos="N"] [lemma="poznat|imenovan|naveden|poput|kao"] [word="kao"  
] [pos="A|N"] * [pos="N"]
```

likvefakcija i dr .) . Pored toga , tu se detaljno moraju razmatrati **uslovi lokacije gradilišta uključujući klimatske prilike ,** geologiju , hidrogeologiju , hidrologiju , topografiju , presek terena i dr . Iz pomenutih podloga moraju se definisati uslovi

kvaliteta pojedinih struktura jalovine - pedološke degradacije . Ovo će uticati na **formiranje tipa zemljišta poznatog kao deposol** , tip zemljišta karakterističan za sve površinske kopove uglja , kod kojih se ne vrši selektivno otkopavanje gornjih

Slika 4-33 Konkordance za upite sa trigerima „uključujući“ i „poznat kao“

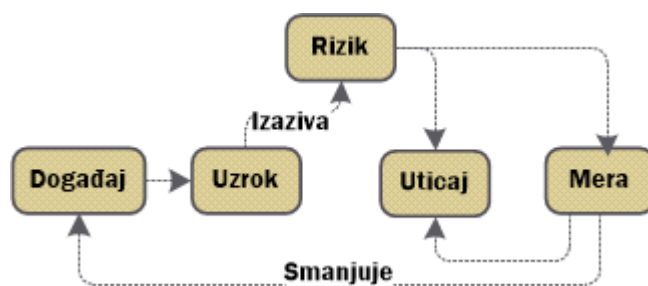
Što se tiče automatskog prepoznavanja definicija i relacija iz teksta, postoji vrlo mali broj istraživanja do sada. Postupak pronalazjenja akronima i njihovih definicija za srpski jezik zasnovan na elektronskim rečnicima srpskog jezika i konačnim automatima, koji je prikazan u (Krstev, Vitas, and Stanković 2015). Aktuelna istraživanja tima su usmerena ka ekstrakciji definicija i prigodnih primera za ilustraciju konteksta upotrebe korišćenjem servisa GDEX⁷³ koji se bazira na ekstrakciji fičera iz teksta ⁷⁴ (Stanković et al. 2019).

4.4.4. Uzročne veze

Uzročni proces predstavlja odnos između dva događaja: prvi je uzrokujući, a drugi uzrokovani događaj, znači odnos između uzroka i posledice. Ružin-Ivanović (2015) navodi sledeće vrste rezultata uzrokovanja: 1) promena stanja ili svojstva; 2) postojanje/nepostojanje; pojavljivanje; 3) promena zvanja ili statusa; 4) kretanje (ili promena mesta/položaja); 5) događaj, fizička pojava; 6) radnja, akcija; 7) afektivna/fiziološka reakcija; 8) intelektualna radnja (mišljenje, verovanje i sl.) (Ružin 2015).

Svaki događaj ima vremensku i prostornu koordinatu, kojima se određuje položaj tela koje trpi promenu. Koncept *RADNJE* ili *AKCIJE* (delovati, učiniti, uraditi nešto) odnosi se na ono što neko lice čini, i razlikuje radnju od događaja. Radnje je ono što činimo, događaj odnos što se dešava. Argumente koje posmatamo u sintaksičkim strukturama ograničavamo u odnosu na krovne koncepte ontologije rizika, koristeći klasifikaciju glagola uzorkovanja i njihovih sintaksičkih struktura.

Slika 4-34 prikazuje ključne koncepte upravljanja rizikom i njihove relacije čije leksikalizacije u tekstu želimo pronaći i povezati.



Slika 4-34 Ključni koncepti upravljanja rizikom i njihove relacije

Glagoli promene stanja imaju kauzativno značenje i mogu se prepoznati pomoću sintaksičkih i morfosintaksičkih markera:

- imaju tranzitivni oblik;
- ulaze u sastav tranzitivne konstrukcije (subjekat predikat objekat) čime kodiraju kauzalni događaj;
- odnose se na UZROK, kao i na REZULTAT;
- kodiraju događaje koji imaju SPOLjAŠNjI UZROK.

⁷³ <http://gdex.jerteh.rs/>

⁷⁴ <http://features.jerteh.rs/>

Ružin-Ivanović (2015) navodi sledeće grupe glagola, daje njihove sintaksičke strukture, klase uzročnika i posledica, od kojih navodimo samo neke:

- promene stanja ili svojstva
- stvaranje i uništavanje
- imenovanje ili karakterizacija
- kretanja uzrokovanje kretanja
- uzrokovanje događaja, pojave
- uzrokovanje radnje
- uzrokovanje afektivne reakcije
- uzrokovanje intelektualnih radnji

U okviru grupe *promene stanja ili svojstva*, tako se mogu dalje posmatrati grupe, kao na primer:

- lomljenja: zdrobiti, izlomiti, izmrviti, nakrnjiti, nalomiti, okrnjiti, polomiti, pocepati, prasnuti, prelomiti, razbiti, razdrobiti, rasprsnuti, slomiti; lomiti, drobiti, krnjiti, lupati.
- savijanja: iskriviti, pokriviti, presaviti, saviti, poviti, uviti; iskrivljavati, kriviti.
- promene u opštem smislu: izmeniti, obrnuti, preinačiti, preokrenuti, pretvoriti, promeniti; izmenjivati, menjati.

Za uništavanje su recimo karakteristični sledeći glagoli: istrošiti, likvidirati, potrošiti, razvejati, razoriti, rastočiti, rasturiti, uništiti; pustošiti, uništavati.

Kao opšti glagoli uzrokovanja događaja se navode: dovesti (do nečega), dovući, doneti, izazvati, izroditi (boleštine), naneti (štetu, smrt), načiniti, povući, pričiniti (štetu, zadovoljstvo), proizvesti, uneti (zbrku, nemir), uroditi, usloviti, učiniti (utisak); izazivati, navlačiti, prouzrokovati, uzrokovati.

Istraživanje reči koje mogu biti okidači za pronalaženje uzročno posledičnih veza krenulo se od primera na engleskom gde su korišćeni sledeći koncepti: „provoke, evoke, cause, explode, result, fire“ i njihovi sinonimi. Za ovo istraživanje su korišćeni sledeći ekvivalenti na srpskom:

- *provoke*: provocirati, izazivati, dražiti, zametnuti,
- *evoke*: evocirati, probuditi izazivati prizivati,
- *cause*: izazvati prouzrokovati zadati usloviti pričiniti,
- *explode*: eksplodirati,
- *result*: rezultirati, poticati, proisticati,
- *fire*: otpustiti pucati opaliti paliti zasuti podstaći pući raspaliti.

Poredeći tekstove na engleskom jeziku i konstrukcije: *because (of), if/then, so, of, as, when, after, present continues* tražene su odgovarajuće strukture i u srpskom:

- Predložno-padeški adverbijali: *zbog, zbog toga što, zato što, usled, jer, od,...*
- Vremenski veznici: *posle, nakon, kada, dok,*
- Adverbijali: *vredno radeći stekao je mnogo...*

Dakle posle uslovne fraze sledi genitiv (može jedna ili više reči). Leksička realizacija kauzativnosti odnosi se na kodiranje uzroka u okviru leksičkog značenja glagola (*cause, force, compel, affect; izazvati, nagnati, nagovoriti, pobuditi*), imenica (*uzrok, razlog, motiv, povod*), prideva (*uslovan, razložan*), priloga (*uslovno, razložno, ...*)

Analizirajući tekstove, pokazalo se da treba posmatrati sledeće reči: *podstaći, prizivati, provocirati, prouzrokovati, probuditi, isprovocirati, izazivati, dražiti, zametnuti, usloviti, pričiniti, rezultirati, poticati, proisticati.*

Traženi su sledeći obrasci u tekstu: **Uzrok** **triger** **posledica**, pa ih tako i obeležavamo u sledećim primerima.

Probijanje nasipa i prodiranje reke Kolubare u kop, kao i prelivanje jezera u Paljuvima, zbog čega je došlo do i do urušavanja i pucanja brane na retenziji Kladnica su prouzrokovale **slabljenje kopa** do kote +76,5mrv.

Voda koja dospe u dublje delove terena, preko razvijenog planuma, izaziva prepakovanje zrna peska i promenu zapremine peskovitih masa. {S} Promenom zapremine peskovitog sočiva izaziva se **sleganje površinskih slojeva alevrita** i **njegovo blokovsko pucanje**. {S}

Dinamika bušenja nije bila zadovoljavajuća zbog čestih kvarova bušaće garniture koji su prouzrokovali **prekide rada** u trajanju od 10 dana.

utvrđivanje postojanja i visine štete prouzrokovane **geološkim istraživanjima**

Sa kvar se mogu takođe promaći zanimljive kolokacije:

Bager je stao sa radom 16.06.2013.godine, zbog **kvara na odlagaču (požar u elektro postrojenju)**.

Zbog **kvara bušaće garniture radovi su prekinuti** od 22.7.2016. godine;

Zbog **kvara na bušačoj garnituri i loših vremenskih prilika istražni radovi su prekinuti**

bager je prestao sa radom došlo je do **kvara na bageru - oštećenje kugle na fremu transporta**

Bunar BL5'-1 je u kvaru od 25.7.2018. godine (nema napona)

Konstrukcije kojima se iskazuju neke uslovne ranje tako mogu biti od interesa:

Reaktiviranje bunara na etaži V BTO sistema, **ukoliko se stvore uslovi**, zavisno od ovodnjenosti etaže i raspoloživosti mehanizacije.

Montaža magistralnog cevovoda ŠLA linije na više mesta, **nakon završetka** transporta rudarske mehanizacije VI BTO sistema. Puštanje u rad svih bunara na ŠLA liniji.

U slučaju da sporazum o naknadi za eksproprisane nepokretnosti sa ranijim vlasnikom ne bude zaključen, odnosno **ukoliko odluka** nadležnog suda o visini ...

4.5. UML model ontologije

Pojam rizika je široko rasprostranjen u različitim poslovnim i društvenim oblastima, što za posledicu ima veliki broj različitih definicija. U osnovi svih definicija sadržana je neizvesnost, odnosno verovatnoća događaja i stepen uticaja njegovih posledica na postavljene ciljeve.

U nastavku teksta prvo će biti opisani koncepti koji se koriste u identifikaciji, analizi i proceni rizika, a zatim će biti predstavljeno proširenje sistema za potrebe upravljanja operativnim rizicima.

U procesu upravljanju rizikom, faza identifikacije rizika i rizičnih događaja je veoma bitna, krucijalni zadatak jeste identifikacija svih potencijalnih događaja, uključujući i one za koji izgleda malo verovatno da će se zaista i dogoditi. Potrebno je obuhvatati rizične događaje koji se mogu desiti u uslovima uobičajenih, kontrolisanih uslova rada, kao i moguće situacije koje ne spadaju u domen uobičajenih aktivnosti, kao što su na primer vanredne situacije.

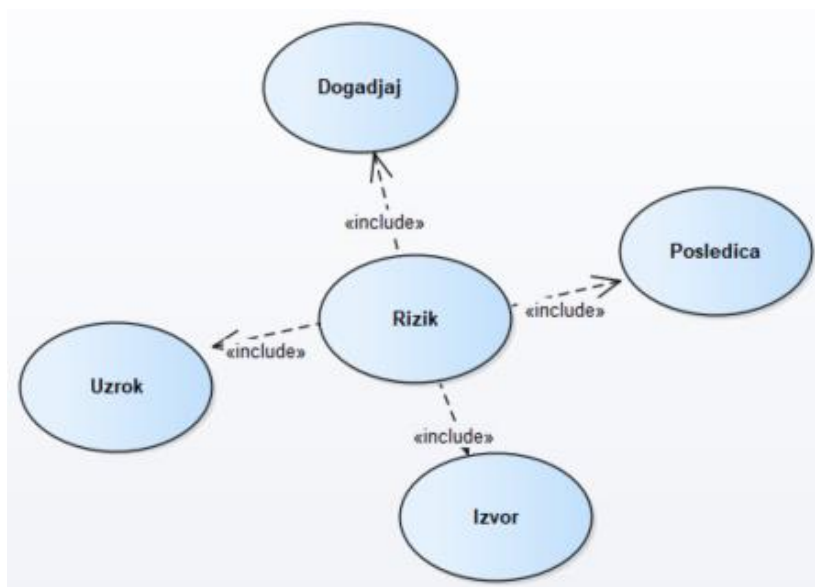
Rizik predstavlja centralni koncept modela. Postoje brojne podele i klasifikacije rizika, ali su najčešće povezane sa određenim aspektom posmatranja. Poslovanje rudarskog preduzeća se u velikoj meri može posmatrati kao i poslovanje svakog drugog industrijskog preduzeća, ali pored klasičnog poslovnog sistema karakterišu ga veoma složeni sistemi, visoka međuzavisnost pojedinačnih procesa i podsistema, specifični uslovi rada, tako da se i mnogi rizični događaji mogu naći u većem broju kategorija. U okviru ovog pristupa napravljena je podela na strateške, projektne i operativne rizike.

Strateški rizici odnose se na događaje koje preduzeće kao celovit sistem ima u odnosu na okruženje ili sa unutrašnjim elementima ali se moraju posmatrati kroz duži vremenski period. To mogu biti rizici povezani za nova nalazišta uglja, dostupnost kreditnih aranžmana, proširenje kapaciteta, nova tržišta, inovacije, licence za rad, starosnu strukturu radnika. Starosna struktura i odliv radne snage u druge grane industrije se i u svetu i kod nas pokazuje kao značajan izvor rizika, s obzirom da rudarska proizvodnja zapošljava dosta raznorodnih profesija, a činjenica je da je za određene poslove neophodno višegodišnje iskustvo.

Projektni rizici vezani su za projekte ili poslove koji ne spadaju u redovno poslovanje, kao što su eksproprijacija zemljišta, izmeštanje infrastrukturnih objekata, projekti koji se realizuju radi unapređenja pojedinih segmenata poslovanja itd.

Operativni rizici odnose se na događaje koji remete dostizanje operativnih ciljeva i oni ređe mogu biti prilike, već uglavnom predstavljaju opasnosti.

Osnovni koncepti koji se koriste u identifikaciji i opisu rizika su događaj, uzrok, izvor i posledica. Slika 4-35 prikazuje UML model osnovnih koncepata za identifikaciju rizika.



Slika 4-35 UML model osnovnih koncepata za identifikaciju rizika

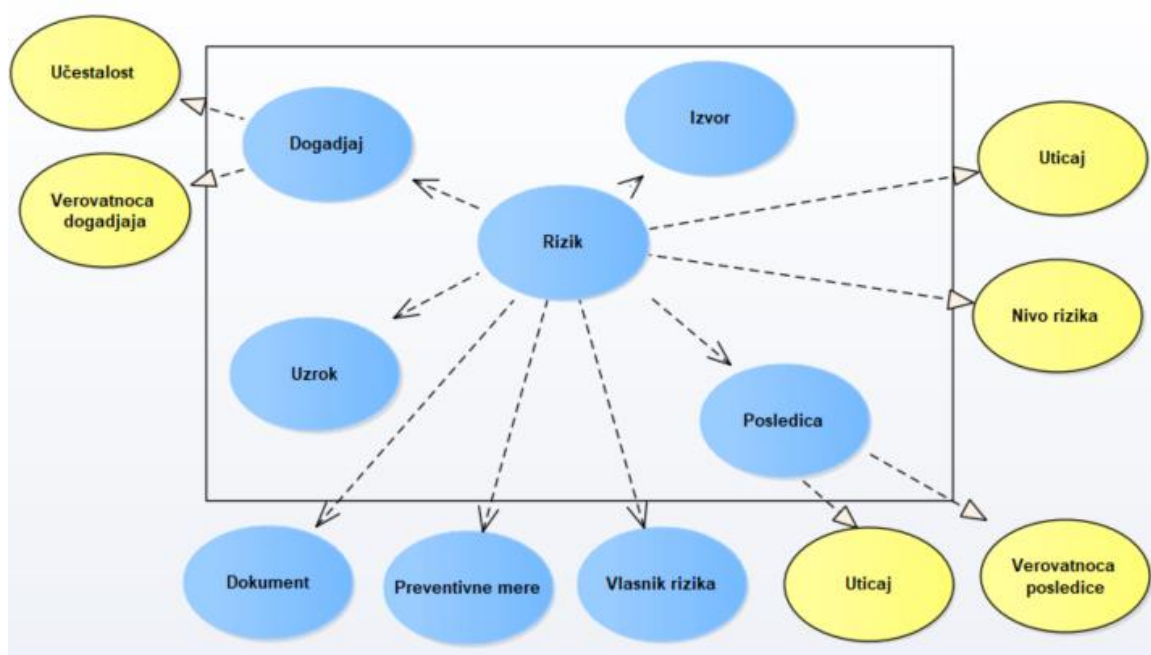
Događaj predstavlja katalog pojava ili spleta okolnosti koje mogu uticati na promenu okolnosti i ostvarenje poslovnih ciljeva. Događaj može biti izazvan sa jednim ili više spoljnih uzroka ili unutrašnjih faktora. Kada iniciraju negativne efekte nazivamo ih opasnostima ili pretnjama, a kada dovode do pozitivnih promena predstavljaju prilike. Prilike se takođe moraju pažljivo razmatrati, jer prilika sa jednog aspekta može izazvati opasnost za neki drugi aspekt sistema. Događaj se dalje razvrstava u taksonomiju različitih vrsta događaja.

Izvor rizičnog događaja su jedan ili više materijalizovanih činilaca koji samostalno ili u sadejstvu sa drugim činilcima mogu dovesti do rizika. To mogu biti akteri na radnom mestu, oprema i alat za rad, radni objekti i materijali.

Uzrok rizika predstavlja određenu radnju ili stanje koje dovodi do neplanirane situacije, praktično je okidač događaja. Klasu čine elementi koje smo u ekstrakcijama termina nalazili povezane sa rizikom, a mogu se koristiti u nekim budućim pretraživanjima teksta kao termini koji ukazuju na rizik. Primeri su: lom, kvar, kidanje. Uzroci se dalje razvrstavaju u taksonomiju različitih vrsta uzroka. Uzroci, kao i izvori, su važan faktor za identifikovanje rizika, ali su takođe veoma bitni u planiranju mera za smanjenje rizika.

Posledica definiše ishod događaja i promenu okolnosti koje se mogu odraziti pozitivno ili negativno na ciljeve. Jedan događaj može imati više posledica, a posledice mogu izazvati i lančanu reakciju. Posledica može biti predstavljena opisno ili u numeričkim vrednostima. Posledice se najčešće mere u finansijskim vrednostima, ali to nije uvek moguće. U ovom modelu usvojicemo podelu na finansijske, ljudske (zdravlje i bezbednost na radu), posledice na životnu sredinu, reputaciju, zakonsku regulativu, prekid rada i kvalitet. Prekid rada i kvalitet se mogu finansijski reprezentovati, ali pošto u rudarstvu imaju poseban značaj, naročito u operativnom upravljanju, izdvajaju se kao posebne kategorije. Koncept Posledica se dalje razvrstava u taksonomiju različitih vrsta posledica.

Analiza i procena rizika i prilika su procesi shvatanja prirode rizika i prilika, vrednovanje i kvantifikovanje značaja, odlučivanje o postupanju sa rizikom i definisanje mera. Koncepti koji pripadaju analizi, proceni i tretmanu rizika su verovatnoća događaja, verovatnoća posledice, uticaj, tretman rizika, vlasnik rizika, učestalost, preventivne mere. Slika 4-36 prikazuje UML model osnovnih konceptata za analizu i procenu rizika.



Slika 4-36 UML model osnovnih konceptata za analizu i procenu rizika

Za svaki rizičan događaj definiše se **verovatnoća događaja**, kao mera šanse da se takav događaj desi. Može biti bazirana na objektivnoj proceni ili je posledica subjektivnog verovanja da se može dogoditi određeni ishod. Termin verovatnoća u domenu upravljanja rizicima treba razlikovati od matematičkog poimanja, gde se vrednost iskazuje brojčano od 0 do 1, dok u teoriji rizika on ukazuje na mogućnost, šansu neke pojave, ali kako se u praksi učvrstila upotreba termina u ovom kontekstu i u ovom radu će biti korišćen kao takav. **Učestalost** se odnosi prevashodno na događaje koji su se već odigrali u prošlosti, a eventualno na potencijalna buduća dešavanja ukoliko je moguće izvršiti takvu procenu.

Uticaj predstavlja kvantifikovanu vrednost posledice rizičnog događaja. Uticaj se određuje za svaku posledicu ponaosob, a zatim i za rizik u celini. Skala uticaja definiše se za svaku kategoriju posledica posebno, npr. prema potencijalnim gubicima, uticaju na zdravlje i bezbednost ljudi ili reputaciju firme.

Kako rizični događaji često utiču na više različitih aspekata sistema, što je u rudarskoj proizvodnji posebno izraženo, uticaj treba tretirati različito sa svakog aspekta, odnosno treba za svaku posledicu izvršiti procenu uticaja koji ona može imati. Takođe, neke od posledica će se

desiti sa većom ili manjom izvesnošću, pa je potrebno pojedinačno za svaku od identifikovanih odrediti i verovatnoću njene pojave.

Verovatnoća i uticaj definišu matricu za rangiranje rizika. Dimenzije matrice mogu se menjati i biti zadate prema potrebama sistema, a u ovom modelu usvajamo dvodimenzionalnu matricu 5x5. **Nivo rizika** predstavlja presek uticaja i verovatnoće u matrici, odnosno njihov proizvod.

Tretman rizika definiše dalje postupanje sa rizikom. Idealno bi bilo svakako eliminisati svaki rizik ili pospešiti ga ukoliko on sa sobom nosi prilike, ali se nameću tehničko-tehnološka, ljudska i finansijska ograničenja. Kroz ovaj koncept beleži se da li će se rizik prihvatiti kao takav i da li ga je potrebno nadgledati ili će se preduzeti potrebne mere za njegovo umanjeње ili potpuno sprečavanje.

Preventivne mere su aktivnosti koje se preduzimaju za smanjenje ili sprečavanje rizika, odnosno pospešivanje kada se radi o prilikama. Mere mogu biti strateške, taktičke i operativne. Kod operativnih rizika to su npr. obuke izvršilaca na radnim mestima sa povećanim stepenom rizika.

Dokument može biti uputstvo, propis, procedura za prevenciju, zaštitu ili kontrolu.

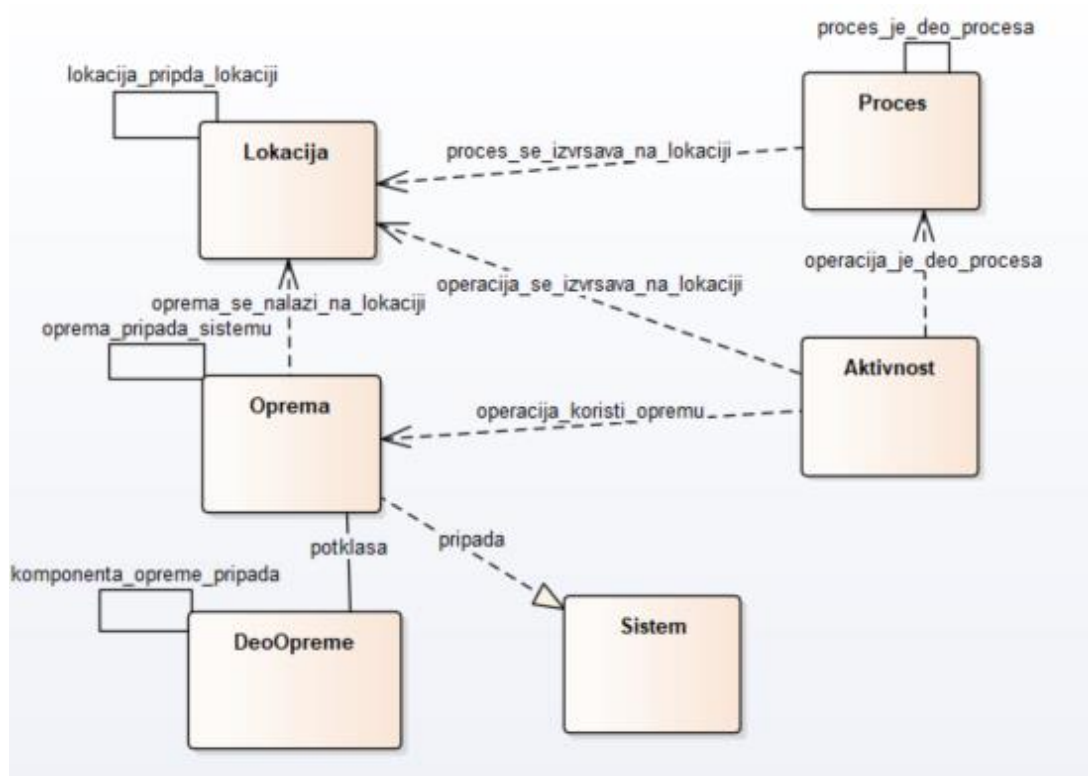
Vlasnik rizika je pozicija ili služba koja je zadužena i odgovorna za upravljanje rizikom.

Upravljanje operativnim rizicima predstavlja poseban segment menadžmenta rizikom koji ima za cilj zaštitu organizacije od unutrašnjih i spoljnih rizika, koji bi mogli negativno uticati na dostizanje operativnih ciljeva organizacije. Fokusirano je na poslovne procese i procedure i ima za cilj da predvidi i spreči bilo koji događaj koji bi mogao prouzrokovati poremećaj u planiranim aktivnostima. Primeri takvih događaja su najčešće povezani sa ljudskim greškama i kvarovima sistema.

Koncepti **Proces, Aktivnost, Lokacija, Izvršilac radnog mesta, Oprema** vezani su za organizaciju rada i predstavljaju entitete koji su usko povezani sa operativnim rizicima. Rizični događaj može se desiti na određenoj lokaciji, tokom određene aktivnosti, prilikom rukovanja opremom, može biti izazvan neadekvatnim radom izvršioca radnog mesta. Ova grupa entiteta može se izdvojiti i kao posebna celina ili više pojedinačnih ontologija koje se mogu koristiti u svim ostalim upravljačkim sistemima, osim u menadžmentu rizikom.

Proces predstavlja niz povezanih aktivnosti koje čine jednu celinu i odvijaju se na jednoj ili više lokacija. Procesi se mogu dekomponovati na potprocese kako bi se lakše vršilo praćenje rizika.

Rudarski proizvodni sistemi i procesi nose sa sobom veliki stepen složenosti i visoku međuzavisnost podsistema, što zahteva stalni nadzor i sinhronizaciju procesa. Osnovni procesi u rudarskoj organizaciji su proizvodnja, održavanje opreme, procesi vezani za upravljanje rudnikom, investicijama, kao i procesi podrške proizvodnji. Kada je u pitanju rudnik uglja proizvodni proces čine potprocesi: proizvodnja uglja, eksploatacija jalovine, drobljenje uglja, proizvodnja komadnog uglja, upravljanje pomoćnom mehanizacijom, pripremni radovi, odvodnjavanje kopova, geološka istraživanja. Za neometano funkcionisanje proizvodnje zadužen je i proces održavanja opreme, čiji su osnovni potprocesi mašinsko održavanje, elektro održavanje, održavanje pomoćne mehanizacije i vulkanizersko održavanje.



Slika 4-37 UML model organizacije rada

Aktivnost je atomska operacija koja se izvršava ili nezavisno ili u okviru jednog procesa. Na jednoj aktivnosti može biti angažovano više različitih izvršilaca, u različitim trenucima i na različitim lokacijama.

Kompleksna rudarska oprema, sistemi i postrojenja zahtevaju da se klasa **Oprema** posebno modelira kroz odnose između sistema i njegovih delova klasa, uz specifikaciju tipa opreme. Klasa **DeoOpreme** je hijerarhijska struktura i čini sastavnicu za jedan komad opreme. U kontekstu upravljanja rizikom dekompozicija opreme može biti i podskup cele sastavnice, sa onim elementima koji imaju značaj razmatranja i praćenja. Oprema može biti inicijator rizika ili pak oštećena usled nekog događaja.

Koncept **Izvršilac radnog mesta** predstavlja u suštini katalog radnih mesta čiji izvršioци mogu biti ugroženi rizičnim događajem, ali isto tako mogu biti izvor rizika tako što prilikom obavljanja neke aktivnosti i rukovanjem određenom opremom mogu inicirati opasnost. Takođe ovoj klasi pripadaju i pozicije odgovorne za određeni rizik, takozvani vlasnici rizika. Dakle, ova klasa je povezana sa konceptima posledica, izvor i vlasnik rizika.

Klasa **Lokacija** predstavlja katalog svih mesta na kojima može doći do rizičnog događaja, to su takođe i mesta na kojima se izvršavaju procesi i aktivnosti. Lokacija može sadržati manje lokacije u svom okviru, npr. basen čini više eksploatacionih polja.

Klasu **Materijal** čine materijali koji mogu biti uzrok rizičnog događaja, kao što su eksploziv, uglj (samozapaljenje), voda, gas, prašina, toksični materijali koji mogu izazvati hemijske udes.

4.6. Razvoj ontologije u Protégé

4.6.1. Koncept razvoja ontologije

Protégé je alat za izgradnju ontologija, koji je otvorenog koda i ima veliku primenu u praksi, a razvija se na Univerzitetu Stanford u Americi (Horridge et al. 2004). Koristi se generalno za razvoj sistema zasnovanih na znanju, sa razvojnim okruženjem koje pruža mogućnost rada istovremeno sa klasama i instancama. Pripada kategoriji ontoloških editora koji podržavaju kreiranje različitih tipova i formata ontologija. Za potrebe ovog rada odabran je editor Protégé 5.5.0.

Razvoj ontologije odvijalo se kroz nekoliko faza:

- kreiranje kostura ontologije, koji sadrži ključne koncepte kao čvrstu osnovu za dalju nadogradnju,
- definisanje ograničenja, atributa, relacija, svojstava,
- paralelno se vršio razvoj alata za prikupljanje znanja (eng. knowledge-acquisition),
- prikupljeni podaci koji čine bazu znanja su uvezeni u ontologiju,
- ručni unos dodatnih potrebnih instanci,
- dopuna novim potklasama i novim ograničenjima,
- definisanje aksioma, mehanizama zaključivanja, i testiranje.

Razvojno okruženje Protégé pruža mogućnost povezivanja različitih ontologija i korišćenje pojedinih klasa iz jedne ontologije kao domena u drugoj ontologiji, što je još jedan od razloga za upotrebu upravo ovog alata.

Osnovni elementi svake ontologije su klase (eng. Classes), atributi (eng. Data Properties) i relacije između klasa (eng. Object Properties).

Jedna od najsnažnijih karakteristika ontoloških modela, koja ih čini izuzetnim u odnosu na klasične sisteme za upravljanje relacionim bazama, ogleda se upravo u bogatstvu načina za definisanje raznovrsnih i višestrukih odnosa između entiteta. Ontologije omogućavaju da dve klase, odnosno dve instance tih klasa mogu imati više različitih relacija, što važi i za instance u okviru jedne iste klase. Primer bi bio da izvršilac radnog mesta može biti pokretač rizika, može biti posledica u slučaju povrede ili uticaja na zdravlje, ali isto tako može biti i osoba koja je odgovorna za monitoring rizika. Tipičan primer za relaciju između instanci iste klase jeste kada hoćemo da ukažemo na ekvivalenciju između dva pojma, npr. *sistem BTO 1* je isto što i *BTO I*, *prvi BTO sistem*.

Protégé nudi ugrađene tipove relacija kao što su potklasa (eng. SubClass Of), ekvivalencija (eng. Equivalent To), aksiomi za povezivanje pojmova sa drugim ontologijama ili rečnicima (eng. General class axioms), različito od (eng. Disjoint With), inverzna relacija (eng. Inverse Of) itd.

U okviru ovog modela u velikoj meri je zastupljena upotreba inverznih relacija, jer one omogućavaju lakše kretanje kroz mrežu podataka, kao i pretraživanje podataka. Za relaciju izvršilac je izazvao rizik, inverzna relacija bi bila rizik je izazvan od strane izvršioca. U praktičnoj upotrebi to nam obezbeđuje da kada posmatramo rizik lako dolazimo do izvršioca, a isto tako kada posmatramo izvršioca možemo lakše da dođemo do svih rizika koje je on izazvao. U relacionim bazama podataka ova inverznost je podrazumevana, međutim SWRL jezik se razlikuje od SQL jezika i bazira se na kaskadnom povezivanju klasa, relacija i atributa, gde smer povezivanja igra bitnu ulogu.

Relacije, osim ugrađenih osobina koje nose sa sobom kada su iz predefinisano opsega, imaju mogućnost pridruživanja dodatnih karakteristika. Jedna od njih je funkcionalnost (eng. Functional) koja ukazuje da se instanci jedne klase može pridružiti samo jedna vrednost druge

klase (npr. događaj može imati samo jednu verovatnoću), a kada to nije naznačeno instanca jedne klase se može povezati sa nebrojeno instanci druge klase. Ova karakteristika postoji i kod atributnih objekata (eng. Data propertis), tako da možemo primera radi reći da rizik može imati više instanci atributa opis i samo jednu instancu atributa nivo rizika. U nastavku teksta takvi odnosi nazivaju se funkcionalna relacija ili funkcionalni atribut.

Pravila kardinalnosti definišu sa koliko instanci (primeraka) jedne klase se može povezati instanca druge klase. Kod klasičnih relacionih modela imamo mogućnost da se jedna instanca može povezati sa jednim ili više instanci druge klase. Ontologija pruža šire mogućnosti, pa može da se za instancu jedne klase navede sa koliko tačno, odnosno maksimalno ili minimalno instanci druge klase može da se izvrši povezivanje. Ranije pomenuta funkcionalnost je jedan od oblika kardinalnosti.

Slika 4-38 (u donjem desnom delu) pokazuje primere različitih veza, pa se vidi da rizik može da ima samo i najviše jednu vrednost za uticaj, može da pripada najviše jednom procesu, ali može da ima maksimalno dva vlasnika rizika. Osobina *tačno* (eng. exactly) ukazuje da mora imati tačno jedan događaja, što znači da mora biti povezan sa događajem, dok sa ostalim klasama i ne mora. Možemo i za svaku potklasu definisati posebne uslove, npr. operativni rizik ima dva vlasnika, a strateški jednog. Nivo rizika je atributni objekat koji može imati najviše jednu celobrojnu vrednost.

Slično je i kada su u pitanju atributi, može se definisati tip podatka, skup mogućih vrednosti ili se vrednosti mogu birati iz instanci neke druge klase. Na attribute se takođe odnose ista pravila kardinalnosti, što znači da klasa može imati više instanci jednog atributa. Kod relacionih modela, jedna instanca može imati samo jednu vrednost atributa.

Svaka klasa, kao i svaka instanca jedne klase može se povezati sa drugim jezičkim resursima, domaćim ili stranim, a na slici je prikazano referenciranje rizika na ontologije i rečnike Wikidata, Mesh, Stanford, Brilliant (Slika 4-38).

The screenshot displays the 'Rizik' class page in a web ontology editor. On the left, a class hierarchy tree shows 'Rizik' as a subclass of 'Rang_verovatnoce', with further subclasses 'Operativni_rizik', 'Projektni_rizik', and 'Strateski_rizik'. The main area is divided into 'Annotations' and 'Description'. The 'Annotations' section lists four 'rdfs:isDefinedBy' relationships with external URIs. The 'Description' section shows 'Equivalent To' and 'SubClass Of' relationships. The 'SubClass Of' section lists four relationships with cardinalities: 'rizik_ima_uticaj max 1 Rang_uticaja', 'rizik_ima_vlasnika max 2 IzvrsilacRadnogMesta', 'rizik_je_rezultat_dogadjaja exactly 1 Događaj', and 'rizik_pripada_procesu max 1 Proces'.

Slika 4-38 Klasa Rizik povezana sa drugim ontologijama

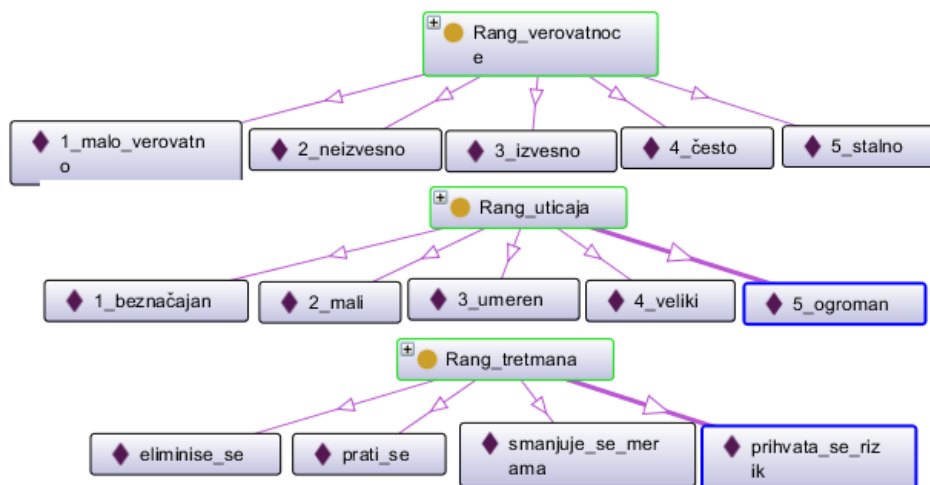
4.6.2. Izgradnja ontologije

Koncepti i pojmovi koji su opisani u okviru UML modela u okviru implementacije prevode se u odgovarajuće objekte. Osnovne klase koje su se izdvojile u modelu prikazuje Slika 4-38. Sve klase imaju atribute Naziv i Definicija, pri čemu se može dati više definicija za jedan pojam.

Rizik predstavlja centralni entitet modela i sadrži tri potklase: Strateški rizici, Projektni rizici i Operativni rizici. Klase se tematski mogu podeliti u tri grupe:

- klase kroz koje se vrši opis rizika su **Rizik, Događaj, Uzrok, Posledica, PreventivneMere, Dokument**.
- klase kojima se preciznije definiše radno okruženje i koje povezuju rizike sa konkretnim činiocima poslovanja su **Projekat, Proces, Aktivnost, Lokacija, Izvršilac radnog mesta, Oprema, DeoOpreme, Materijal**.
- klase za kategorizaciju rizika su **Rang_verovatnoce, Rang_uticaja i Rang_tretmana**.

Klase **Rang_uticaja** i **Rang_verovatnoce** koriste se za definisanje matrice rizika, odnosno kao klasifikatori u proceni rizika. U modelu (Slika 4-39) je usvojena matrica 5x5, kao najčešće korišćena, ali rešenje dozvoljava izbor dimenzija matrice. Elementi klase **Rang_verovatnoce** su: *malo verovatno, neizvesno, izvesno, često, stalno*, a klase **Rang_uticaja**: *beznačajan, mali, umeren, veliki, ogroman*. Klasa **Rang_tretmana** ima kategorije: *prihvata se, prati se, eliminiše se u potpunosti i smanjuje se na nivo prihvatljivog*, čime se definiše način tretiranja rizika.



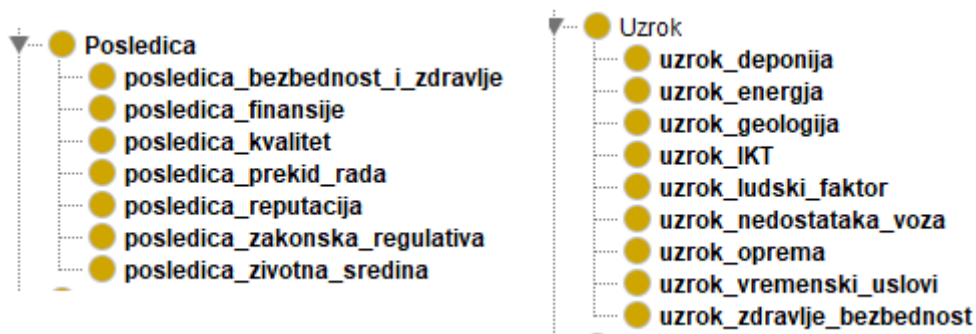
Slika 4-39 Klase matrice rizika

Značenje klasa i njihovi osnovni sadržaji su detaljno opisani u okviru UML modela, pa će nastavak teksta biti posvećen načinu implementacije klasa, odnosima između klasa, kao i odnosima između instanci koje pripadaju jednoj klasi.

Relacija *rizik_je_rezultat_dogadjaja* (**Rizik->Događaj**) opisuje odnos gde je rizik posledica neizvesnog događaja koji može izazvati poremećaj u postizanju zadatih ciljeva. Relacija je funkcionalna jer je rizik posledica samo jednog događaja. Inverzna relacija je *dogadjaj_implicitira_rizik* (**Događaj->Rizik**).

Klasa Uzrok sadrži posebnu terminologiju koja sadrži u sebi ključne inicijatore koje mogu izazvati rizik. Ujedno, njen sadržaj se može koristiti za dalju ekspanziju upita u iterativnim fazama pronalaženja rizika ekstrakcijom. Može se izdvojiti i kao posebna ontologija ili povezati sa rečnikom Termini. Odnos između uzroka i rizika opisuje se funkcionalnom vezom *ima_uzrok* (**Rizik->Uzrok**) i inverznom vezom *uzrok_inicira_rizik* (**Uzrok ->Rizik**). Uzroci se između sebe povezuju relacijom ekvivalencije *uzrok_isto_kao* (**Uzrok -> Uzrok**) kada želimo da ukažemo na isto značenje. Sa aspekta rudničke proizvodnje uzroke delimo na one koji su povezani sa

opremom, vremenskim uslovima, ljudskim faktorom, zdravljem i bezbednošću, stanjem deponije, energetskim snabdevanjem, stabilnošću dolaska vozova za transport uglja, geologijom, informaciono-komunikacionom strukturom. Slika 4-40 prikazuje potklase uzorka, kao i potklase posledica.



Slika 4-40 Potklase uzroka i potklase posledica

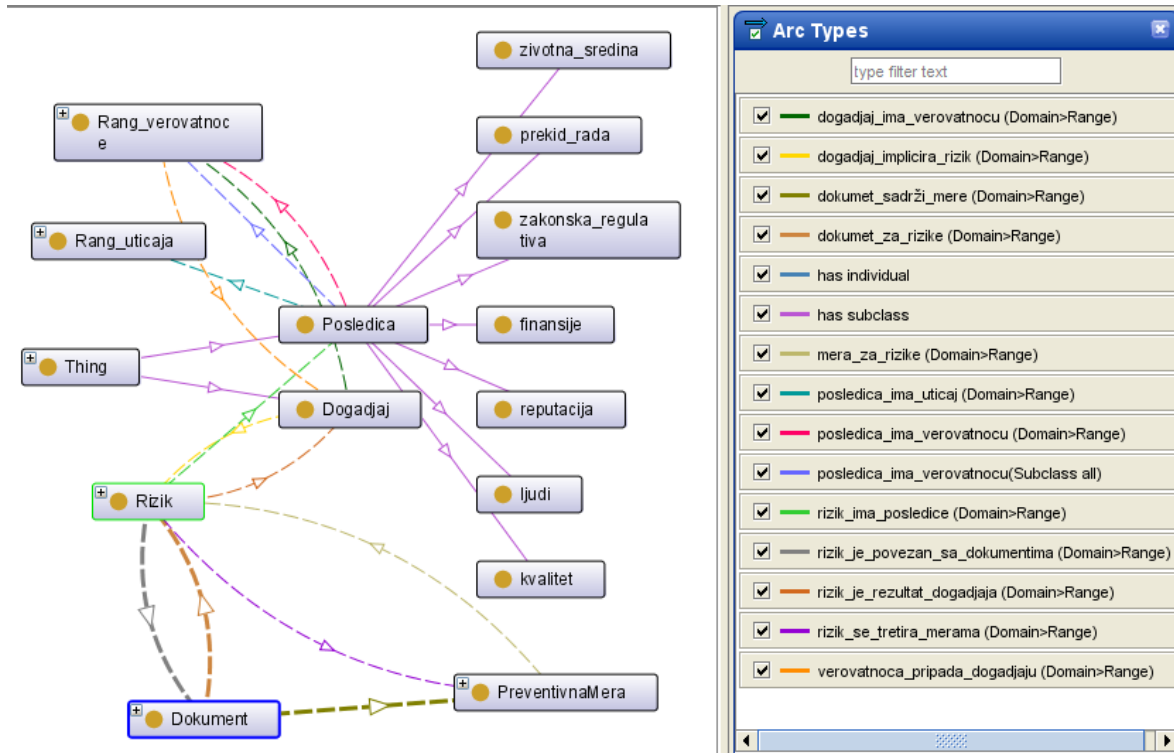
Klasa posledica predstavlja katalog svih mogućih posledica koje se mogu povezati sa odgovarajućim rizikom. Veza *rizik_ima_posledice* (**Rizik**->**Posledica**) je relacija između rizika i posledica koje on može da ima, a inverzna relacija bi bila *posledica_je_rizika* (**Posledica** -> **Rizik**). Posledica ima potkategorije sa kojima je povezana odnosom superklasa -> potklasa. To su: finansije, ljudi (zdravlje i bezbednost na radu), životna sredina, reputacija, zakonska regulativa, prekid rada, kvalitet (Slika 4-41). Kako je ranije rečeno, rizični događaj u rudarskoj industriji često ima posledice koji se sagledavaju sa više aspekata, pa je za svaku posledicu potrebno odrediti verovatnoću njene pojave i uticaj koji ima na ciljeve. Vrednosti se biraju iz predefinisanih lista **Rang_verovatnoce** i **Rang_uticaja**, a implementacija je realizovana preko funkcionalnih relacija *posledica_ima_verovatnocu* (**Posledica**->**Rang_verovatnoce**) i *posledica_ima_uticaj* (**Posledica**->**Rang_uticaja**).

Rangiranje uticaja se vrši za svaku posledicu ponaosob, a zatim se daje ukupan nivo uticaja rizika. Rangiranje, odnosno procena rizika vrši se kroz verovatnoću događaja i stepena uticaja iz zadate matrice rizika, što se realizuje kroz relacije *dogadjaj_ima_verovatnocu* (**Događaj**->**Rang_verovatnoce**) i *rizik_ima_uticaj* (**Rizik** ->**Rang_uticaja**). Obe relacije su funkcionalne, odnosno za jedan rizični događaj može se definisati samo jedna verovatnoća i jedna visina uticaja. Inverzne relacije su *dogadjaji_sa_verovatnocom* (**Rang_verovatnoce**->**Događaj**) i *uticaj_je_na_rizike* (**Rang_uticaja**->**Rizik**), koje omogućavaju identifikaciju svih događaja koji nose određenu verovatnoću pojavljivanja, odnosno rizika sa zadatim nivoom uticaja. Učestalost događaja predstavlja funkcionalni atribut klase **Događaj**. Nakon procene rizika vrši se i procena daljeg postupanja u odnosu na rizik. Veza *rizik_ima_tretman* (**Rizik**-> **Rang_tretmana**) definiše način tretiranja rizika, sa inverznom relacijom *tretman_se_primenjuje_na_riziku* (**Rang_tretmana** ->**Rizik**).

Rizicima se upravlja kroz delovanje odgovarajućim merama. Mere za prevenciju su u najvećem broju sadržane u dokumentima kao što su procedure, planovi i uputstva, a kada to nije slučaj definišu se kroz plan delovanja. Klasa Preventivne mere sadrži katalog svih mogućih mera, a prevencija rizika definiše se kroz relacije *rizik_se_tretira_merama* (**Rizik**->**Preventivna_mera**) i *rizik_je_povezan_sa_dokumentima* (**Rizik**->**Dokument**). Preventivnih mera može biti više, a jedna mera može se odnositi na više rizika. Isto važi i za dokumenta, a inverzne relacije su *mera_za_rizike* (**Preventivna_mera** ->**Rizik**) i *dokument_za_rizike* (**Dokument** ->**Rizik**).

Dokument koji sadrži instrukcije za sprečavanje neželjenih ishoda se može razložiti na pojedinačne mere prevencije, što se obezbeđuje kroz relaciju *dokument_sadrzi_mere* (**Dokument** -> **Preventivna_mera**) je.

Slika 4-41 prikazuje deo modela kojim se formalizuju događaji i njihove posledice.

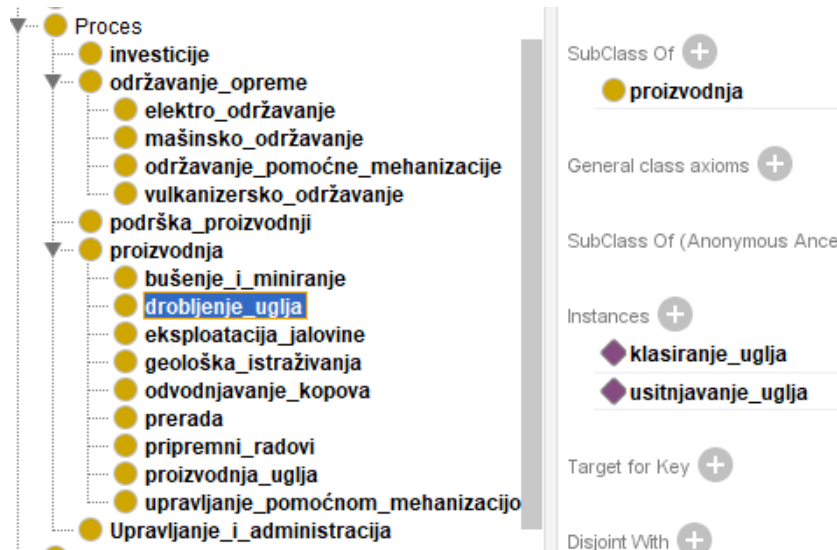


Slika 4-41 Klase kojima se modeliraju događaj, posledice i preventivne mere

Klase kojima se proširuje model za potrebe preciznijeg definisanja projektnih i operativnih rizika su **Projekat**, **Proces**, **Aktivnost**, **Lokacija**, **Izvršilac radnog mesta**, **Oprema**, **Sistem**, **DeoOpreme**, **Materijal**.

Klasom **Projekat** evidentiraju se svi rizični projekti u preduzeću, a relacijom *projekat_ima_rizik* (**Projekat**->**ProjektniRizik**) povezuju se svi potencijalni rizici sa projektom, gde projekat može imati veći broj rizika.

Klase **Proces** i **Aktivnost** definišu sve procese u preduzeću i aktivnosti u okviru njih. Osnovni procesi se mogu dekomponovati na potprocese kako bi se lakše vršilo praćenje rizika i to je omogućeno vezom *proces_je_deo_procesa* (**Proces**-> **Proces**), a aktivnosti u okviru procesa ili potprocesa se povezuju relacijom *aktivnost_je_deo_procesa* (**Aktivnost**-> **Proces**) i aktivnost može biti deo samo jednog procesa. Dekompoziciju procesa na potklase prikazuje Slika 4-42.



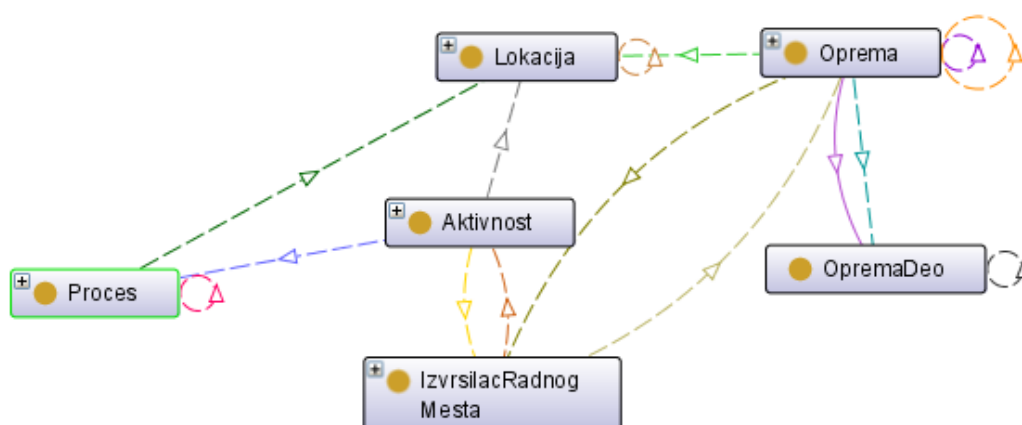
Slika 4-42 Dekompozicija rudarskih procesa

Lokacije su mesta na kojima je moguće da dođe do rizičnog događaja i na kojima se izvršavaju procesi i rizične aktivnosti, što se implementira relacijama *proces_se_izvršava_na_lokaciji* (**Proces->Lokacija**), odnosno *aktivnost_se_izvršava_na_lokaciji* (**Aktivnost-> Lokacija**). Lokacije se mogu dekomponovati i kreirati u hijerarhijsku strukturu što obezbeđuje veza *lokacija_pripada_lokaciji* (**Lokacija->Lokacija**). Proces proizvodnje se odvija na eksploatacionom polju, a recimo otkopavanje rude na određenoj etaži.

Oprema predstavlja ključnu komponentu sistema, jer od njenog stanja, ispravnosti, starosti, rezervnih delova u velikom stepenu zavisi celokupan proizvodni proces. Kod kontinualne proizvodnje se pojačava ovaj uticaj, jer prekid rada jedne opreme može lančano ugroziti ceo proces. Opremu u rudarstvu čine izuzetno kompleksni sistemi, pa je potrebno izvršiti dekompoziciju opreme, kako bi se preciziralo tačno koji delovi predstavljaju rizična mesta. Sa aspekta upravljanja rizikom nije neophodno kreirati kompletne sastavnice, niti preslikati tačnu hijerarhiju sistema, već identifikovati rizične komponente svake opreme. Sastavnica se realizuje kroz relaciju *komponenta_pripada_opremi* (**Oprema->DeoOpreme**) i *komponenta_je_deo* (**DeoOpreme ->DeoOpreme**). Više pojedinačnih komada opreme se povezuje u jedan proizvodni lanac, kao što je slučaj sa utovarno transportnim sistemima koji čine jedinstvenu zaokruženu celinu. Rizik se može posmatrati sa aspekta sistema, pojedinačne opreme ili njene komponente. Sisteme ćemo takođe evidentirati kroz klasu **Oprema** a povezanost više oprema u jedan sistem implementiran je involucijom *oprema_pripada_sistemu* (**Oprema-> Oprema**).

U toku obavljanja aktivnosti učestvuju izvršioци radnog mesta, koji koriste određenu opremu, što se definiše vezama *uposljava* (**Aktivnost->IzvršilacRadnogMesta**) i *koristi_opremu* (**IzvršilacRadnogMesta-> Oprema**). Inverzne relacije su *izvršava* (**IzvršilacRadnogMesta->Aktivnost**) i *opremu_koristi* (**Oprema ->IzvršilacRadnogMesta**).

Slika 4-43 prikazuje klase organizacije operativnih aktivnosti.



Slika 4-43 Klase organizacije operativnih aktivnosti

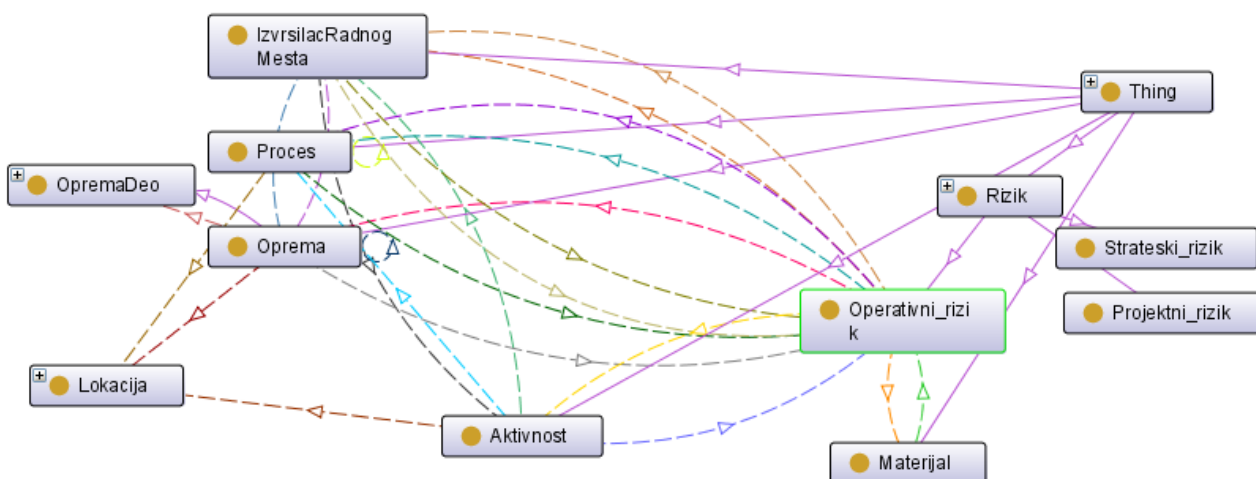
Događaj se dešava na lokaciji prilikom obavljanja određene aktivnosti, a odnos između klasa predstavljen je vezama *dogadjaj_se_desava_na_lokaciji* (**Dogadjaj->Lokacija**).

Rizik se može povezati za obavljanje određene aktivnosti ili za proces u celini. Veze za implementaciju odnosa su *rizik_pripada_procesu* (**OperativniRizik ->Prosec**) i *rizik_pripada_aktivnosti* (**OperativniRizik ->Aktivnost**) Slika 4-44. Rizik može pripadati samo jednom procesu, odnosno aktivnosti, a proces i aktivnost mogu imati više rizika i njihov odnos definiše se inverznim relacijama *proces_ima_rizik* (**Prosec->OperativniRizik**) i *aktivnost_ima_rizik* (**Aktivnost-> OperativniRizik**).

Nepovoljan događaj u okviru jednog procesa koji izaziva prekid rada u uslovima velike međuzavisnosti rudarskih procesa neminovno će uticati i na druge procese, što implementiramo vezom *rizik_utice_i_na* (**OperativniRizik** ->**Prosec**).

Izvor rizičnog događaja može biti izvršilac radnog mesta, nepravilno rukovanje, kvar opreme ili npr. incident usled toga da se otkači neki deo opreme, materijal (eksploziv, samozapaljenje uglja, toksične materije). Izvorce rizika povezujemo sa pripadajućim elementima preko relacija *rizik_je_iniciran_izvršilac* (**OperativniRizik**->**IzvršilacRadnogMesta**), *rizik_je_iniciran_oprema* (**OperativniRizik**->**Oprema**) i *rizik_je_iniciran_materijal* (**OperativniRizik**->**Materijal**). Inverzne relacije obezbeđuju da se za svakog izvršioca, materijal ili opremu definišu potencijalni rizici koje oni mogu izazvati, to su *izvršilac_inicira_rizik* (**IzvršilacRadnogMesta**->**OperativniRizik**), *materijal_inicira_rizik* (**Materijal**->**OperativniRizik**) i *oprema_inicira_rizik* (**Oprema**->**OperativniRizik**).

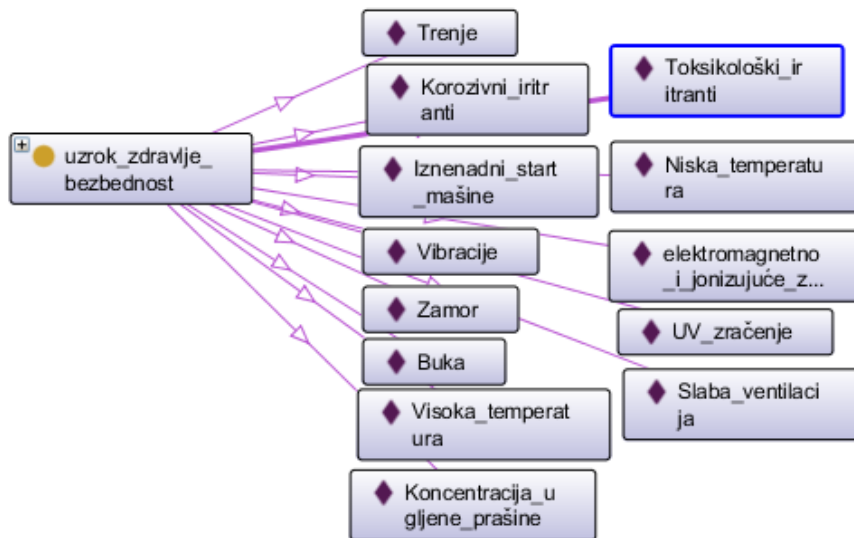
Vlasnik rizika jeste radna pozicija u organizaciji i bira se iz domena izvršilaca radnih mesta, tako da se ovaj koncept realizuje vezom *rizik_ima_vlasnika* (**OperativniRizik**->**IzvršilacRadnogMesta**). Vlasnik rizika može biti zadužen za više rizičnih situacija i relacijom *je_vlasnik_rizika* (**IzvršilacRadnogMesta**->**OperativniRizik**) se opisuje taj odnos.



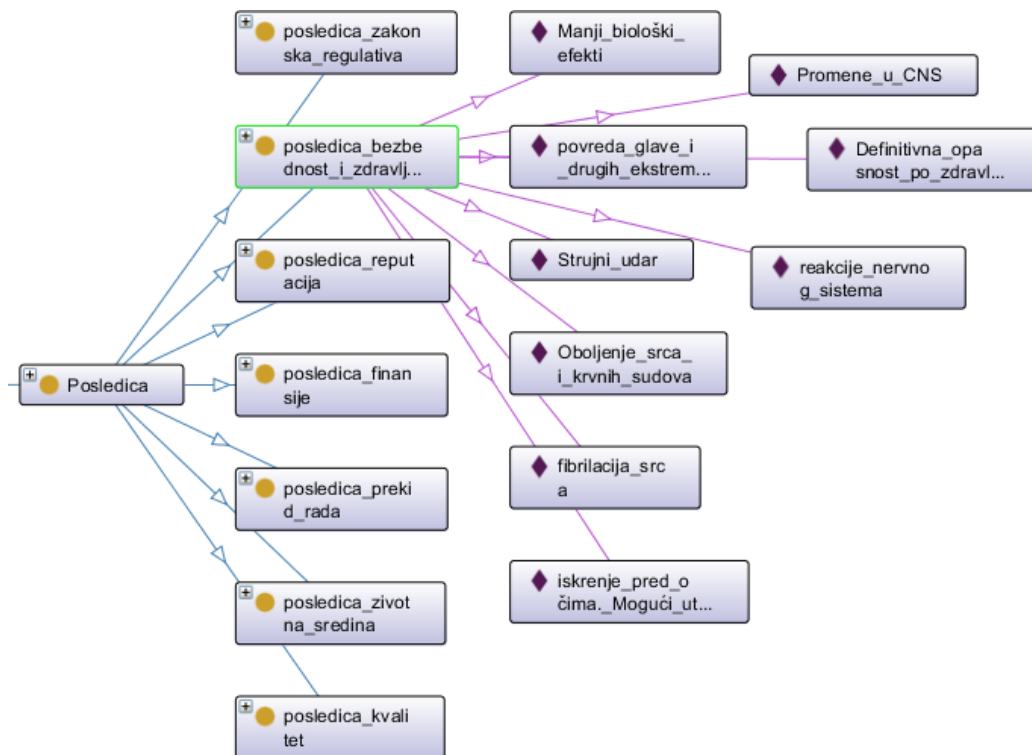
Slika 4-44 Operativni rizici

4.6.3. Mehanizam zaključivanja, upiti i popunjavanje ontologije

Paralelno sa izgradnjom ontologije vršena je i ekstrakcija termina, razvrstavanje u odgovarajuće klase i uvoz podataka. Slika 4-45 prikazuje neke od instanci vezane za uzroke koji utiču na zdravlje operatera, a Slika 4-46 prikazuje instance koje pripadaju posledicama na bezbednost i zdravlje ljudi.



Slika 4-45 Primer instanci potklase Uzroci rizika po zdravlje



Slika 4-46 Primeri instanci potklase Posledice na zdravlje i bezbednost

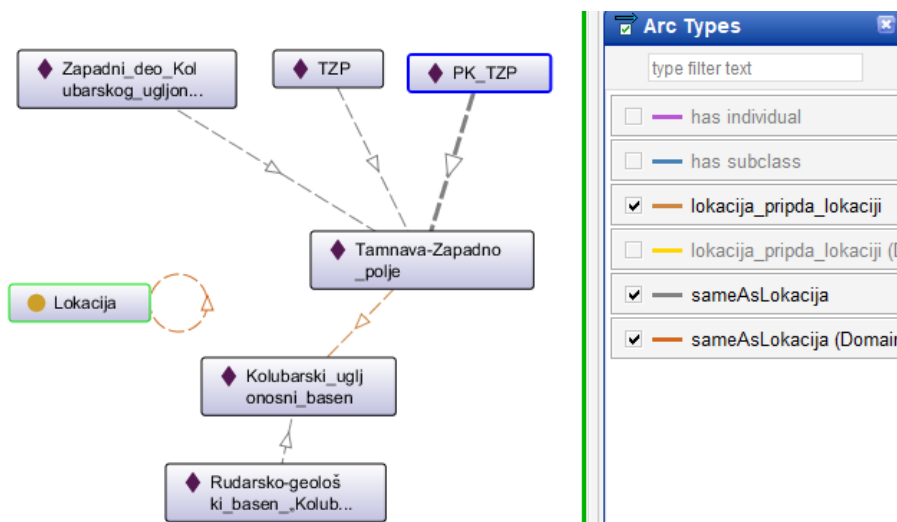
Klasa Lokacije je na početku istraživanja bila osmišljena kao hijerarhijska struktura, u kojoj su navedene sve lokacije u rudniku sa pripadajućim manjim lokacijama. Tokom rada uočeno je da postoji velika heterogenost u upotrebi različitih naziva za jednu istu lokaciju, sa čim smo se i ranije susretali u praksi. Iz tih razloga model je proširen vezom *sameAsLokacija* koja treba da poveže sinonime. Slična problematika je uočena i kod naziva opreme, pa je naknadno implementirana i veza *sameAsOprema*

Primenom mehanizma zaključivanja delimično formalizovani podaci mogu se dodatno formalizovati. Za kreiranje mehanizama koristi se SWLR jezik, koji je ranije opisan, a u nastavku teksta demonstriran je primer sa lokacijama.

Opis primera: ako je lokacija L1 isto što L2, a L2 pripada lokaciji L3, onda i L1 pripada L3.

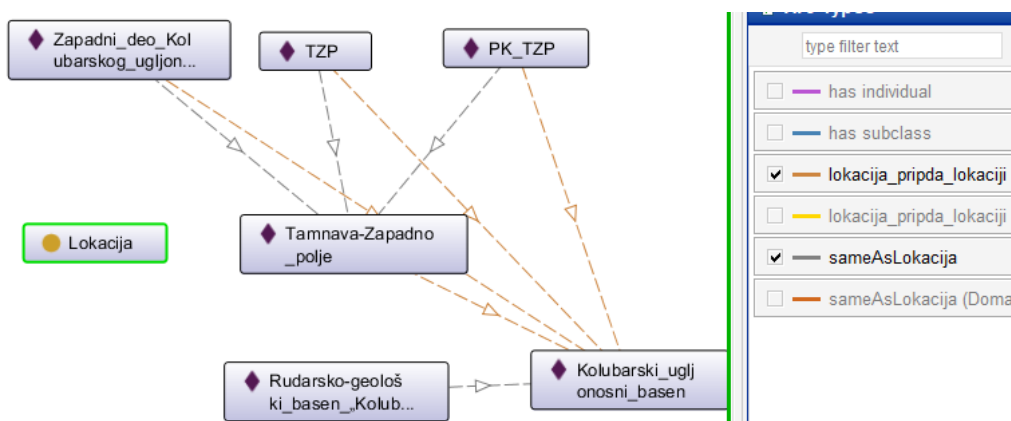
Pravilo: Lokacija(?l1) ^ sameAsLokacija(?l1, ?l2) ^ jeMasterNaziv(?l2, true) ^ lokacija_pripda_lokaciji(?l2, ?l3) -> lokacija_pripda_lokaciji(?l1, ?l3)

Slika 4-47 prikazuje primer odnosa lokacija pre primene mehanizma zaključivanja. Lokacija Tamnava Zapadno polje koja pripada Kolubarskom ugljonošnom basenu, a ima sinonime TZP, PK TZP, Zapadni deo Kolubarskog ugljonošnog basena.



Slika 4-47 Lokacije pre primene pravila zaključivanja

Slika 4-48 prikazuje nove veze koje su nastale kao rezultat primene mehanizma zaključivanja. TZP, PK TZP, Zapadni deo Kolubarskog ugljonošnog basena pripadaju Kolubarskom ugljonošnom basenu.



Slika 4-48 Lokacije posle primene mehanizma zaključivanja

Sledeći primer ilustruje dodatno formalizovanje događaja koji su delimično opisani kroz ontologiju, a koje možemo primenom pravila mehanizma zaključivanja dodatno definisati ili kategorizovati povezivanjem sa odgovarajućim instancama. Primer su događaji kao što su „Došlo je do pucanja lanca gusenice transportnog mehanizma“ ili „Desio se kvar na vratilu reduktora radnog točka“, koji je evidentiran u ontologiji kroz klase Događaj, Uzrok i DeoOpreme.

Opis pravila: Ukoliko je događaj D povezan sa delom opreme DO i uzrok događaja je U koji pripada kategoriji uzrok_kvar(sadrži uzroke kao što su kidanje, lom...) onda je neminovno došlo do Posledice zastoje.

```
Pravilo: Rizik(?r) ^ rizik_je_iniciran_oprema(?r, ?o) ^
rizik_ima_uzrok(?r, ?u) ^ swrlb:member(?u, uzrok_kvar) ->
rizik_ima_posledice(?r, zastojs)
```

SWLR jezik je namenjen prvenstveno definisanju pravila zaključivanja, ali se može koristiti i za selekciju i prezentaciju željenih podataka, što je demonstrirano kroz naredna dva primera.

Neželjeni događaji mogu biti inicirani delovima opreme (npr. kvar određenog dela), što se kroz ontologiju upravo i vezuje za deo opreme. Naredni primer (Slika 4-49) prikazuje kako da izvršimo pregled svih kvarova koji su vezani za jednu instancu opreme, npr. Bager 1.

Neželjeni događaj povezan je za strelu radnog točka koja predstavlja deo opreme. Zahvaljujući hijerarhijskoj organizaciji klase Oprema znamo da je strela deo radnog točka, a opet radni točak je deo Bagera 1, tako da kada želimo da pogledamo sve događaje koji su povezani za određenu opremu i sve njene komponente, to možemo učiniti preko odgovarajućih veza. Primer upita i dobijenih rezultata prikazan je na slici ispod, a prikazuje zastoje koji su se desili na Bageru 1.

p	d1	d2	opis	count(?opis)
Bager1	:Radni_tocak	:vratilo_reduktora	"Desio se kvar na vratilu reduktora radnog točka... :zastoj	1
Bager1	:Radni_tocak	:Strela_radnog_tocka	"Već na početku smene došlo je do kvara na re... :zastoj	4
Bager1	:Radni_tocak	:Strela_radnog_tocka	"Gusenica transport. meh. se pokidala"^^rdf:Plai... :zastoj	1
Bager1	:Radni_tocak	:Strela_radnog_tocka	"Tokom I smene došlo je do kvara na transpotn... :zastoj	1
Bager1	:Radni_tocak	:Strela_radnog_tocka	"Došlo je do pucanja lanca gusenice transportn... :zastoj	1

Slika 4-49 Primer SWLR upita za selekciju podataka o opremi

Naredni primer ilustruje pretraživanje, grupisanje i prebrojavanje podataka (Slika 4-50).

Ukoliko želimo da izračunamo frekvenciju zastoja koje su izazvale pojedinačne komponente bagera, to možemo učiniti upitom prikazanim na slici ispod, gde kao rezultat možemo videti da je vratilo reduktora izazvalo jedan incident, a strela radnog točka četiri.

o	d2	count(?opis)
:Bager1	:vratilo_reduktora	"1^^xsd:int
:Bager1	:Strela_radnog_tocka	"4^^xsd:int

Slika 4-50 Primer SWLR upita za grupisanje i prebrojavanje podataka

Za pretraživanje i selekciju željenih podataka iz ontologije koristi se SPARQL jezik, koji je pominjan ranije u radu, a u nastavku su prezentovani primenjeni primeri nad rudarskom ontologijom. Slika 4-51 prikazuje instance iz klase Posledice koje su povezane sa rizicima, gde možemo videti da ih ima ukupno 48. Slika 4-52 dok prikazuje sve lokacije koje pripadaju jednom rudarskom basenu.

The screenshot shows a SPARQL query in a web interface. The query is as follows:

```

1 PREFIX : <https://rgf.bg.ac.rs/MiRis#>
2 PREFIX dc: <http://purl.org/dc/elements/1.1/>
3 PREFIX grddl: <http://www.w3.org/2003/g/data-view#>
4 PREFIX MiRis2: <https://rgf.bg.ac.rs/MiRis2#>
5 PREFIX owl: <http://www.w3.org/2002/07/owl#>
6 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
7 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
8 PREFIX swrl: <http://www.w3.org/2003/11/swrl#>
9 PREFIX swrla: <http://swrl.stanford.edu/ontologies/3.3/swrla.owl#>
10 PREFIX swrlb: <http://www.w3.org/2003/11/swrlb#>
11
12 SELECT *
13 from <https://rgf.bg.ac.rs/MiRis2>
14 WHERE {
15   ?s ?p ?rizik .
16   ?s a MiRis2:Rizik.
17   ?s MiRis2:rizik_ima_posledice ?rizik.
18 } LIMIT 1000

```

The interface includes a 'Submit' button, a 'Clear' button, and a checkbox for 'Include inferred' (checked). The execution time is shown as 'Done in 48 millisecond'. The results table has three columns: 's', 'p', and 'rizik'. The results show various instances of 'Rizik' and their associated 'Posledice' (consequences).

Slika 4-51 Primer SPARQL upita za selekciju posledica povezanih sa rizikom

The screenshot shows a SPARQL query in a web interface. The query is as follows:

```

1 PREFIX rdo: <http://www.semanticweb.org/RuDokOnto#>
2 PREFIX dc: <http://purl.org/dc/elements/1.1/>
3 PREFIX grddl: <http://www.w3.org/2003/g/data-view#>
4 PREFIX owl: <http://www.w3.org/2002/07/owl#>
5 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
6 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
7 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
8
9 Select ?rudnik ?tr
10 Where {
11   ?ps rdo:upravljaRudnikom ?rudnik.
12   ?rudnik rdo:imaStatus ?rs.
13   ?rudnik rdo:jeTipRudnika ?tr.
14   ?rudnik rdo:pripadaBasenu ?b.
15 }
16

```

The interface includes a 'Submit' button, a 'Clear' button, and a checkbox for 'Include inferred' (checked). The execution time is shown as 'Done in 66 millisecond'. The results table has two columns: 'rudnik' and 'tr'. The results show various instances of 'Rudnik' and their associated 'TipRudnika' (mine types).

Slika 4-52 Primer SPARQL upita za izbor lokacija

4.6.4. Predlog interfejsa za unos rizika

Protégé platforma nudi snažne alate za kreiranje raznovrsnih tipova i mnogostrukih veza između entiteta ontologije, kao i alate za kreiranje novih znanja na osnovu postojeće baze, međutim slabost ovog sistema ogleda se u delu za unos pojedinačnih instanci i njihovo opisivanje, što iziskuje nelagodan rad koji oduzima i dosta vremena.

Iz tog razloga, a za potrebe istraživanja, kreirana je pilot aplikacija *MIR* za evidentiranje podataka koja značajno doprinosi daljoj dopuni ontologije. Kroz prvi deo panela (Slika 4-53) unosi se opis događaja, jedinstveni identifikator i broj ponavljanja. Ukoliko se događaj već ranije dešavao evidentira se broj ponavljanja. Kategorija, verovatnoća, uticaj, tretman i vlasnik rizika odabiraju se sa predefinisanih lista, koje su pohranjene u ontologiji.

The image shows a web-based form for entering risk data. It consists of several input fields and buttons. The first row contains 'Identifikator' with the value 'ID5412' and 'Kategorija' with the value 'Operativni rizik'. The second row contains 'Opis rizičnog događaja' with a text area containing the description: 'Došlo je do pucanja vratila reduktora radnog točka, zbog čega je bager bio u zastoju celog dana'. Below this are four buttons: 'Verovatnoća' (3 izvesno), 'Broj ponavljanja' (empty), 'Ukupan uticaj' (4 veliki), and 'Tretman rizika' (Izaberi...). The final row contains 'Vlasnik rizika' (empty).

Slika 4-53 Panel za unos događaja i procenu rizika

U sledećem koraku evidentiraju se uzroci, izvori i posledice. Uzroka može biti više i oni se biraju sa predefinisane liste. Ukoliko se potreban uzrok ne nalazi na listi onda ga je potrebno dodati u ontologiju. Dugme *Import* omogućava uvoz većeg broja uzoraka. Izvori, odnosno pokretači izvora biraju se sa predefinisanih lista skladištenih u ontologiji, materijal, oprema ili deo opreme i/ili izvršilac na radnom mestu. Posledica se definiše kroz kategoriju kojoj pripada, a izbor konkretne instance bira se sa liste. Kao i kod uzroka, ukoliko se instanca ne nalazi na listi, potrebno ju je dodati u ontologiju, a dugme *Import* omogućava uvoz u ontologiju većeg skupa definisanih posledica. Za svaku posledicu, odabira se sa liste verovatnoća njene pojave i nivo uticaja.

Slika 4-54 Panel za evidentiranje uzroka i posledice

Postupak evidentiranja preventivnih mera je analogan prethodno opisanim procedurama. Posledica se bira sa padajuće liste, a može se izvršiti masovan uvoz u ontologiju.

Slika 4-55 Panel za unos preventivnih mera

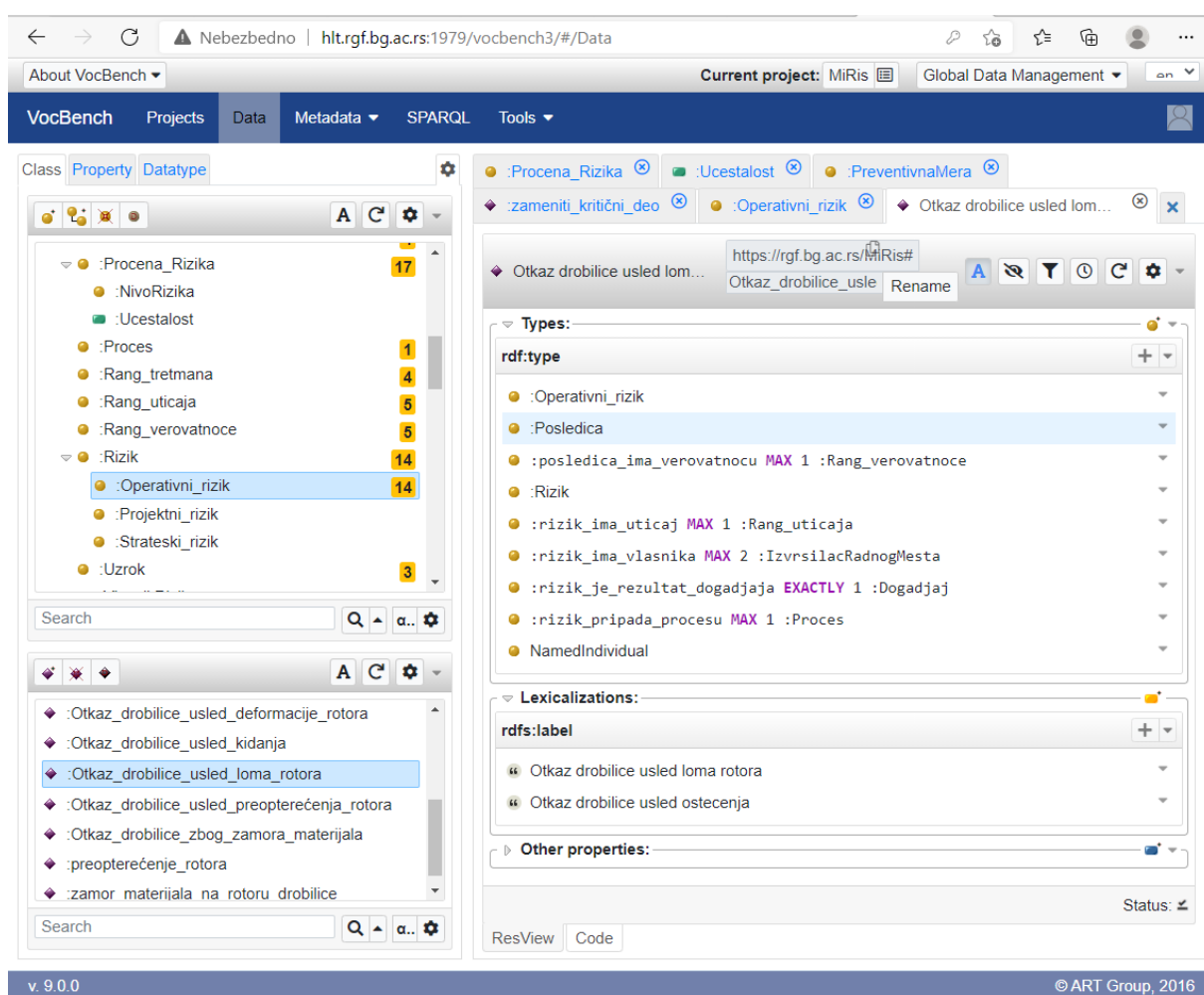
Kroz poslednji odeljak vrši se dodatna formalizacija rizika kroz povezivanje za proces rada, aktivnost i lokaciju.

Slika 4-56 Panel za definisanje procesa i lokacije događaja

4.7. Publikovanje ontologije alatom VocBench

Rudarska ontologija publikovana je upotrebom VocBench⁷⁵ alata, na adresi <https://rgf.bg.ac.rs/MiRis>. VocBench je veb platforma otvorenog koda, namenjena razvoju i publikovanju jezičkih resursa, u skladu sa standardima semantičkog veba, (Stellato et al. 2020). Sistem je interoperabilan i može da vrši razmenu podataka sa drugim standardnim tehnologijama. Prednosti u odnosu na Protégé su veb orijentacija, mogućnost rada više korisnika istovremeno i bolji interfejs za unos.

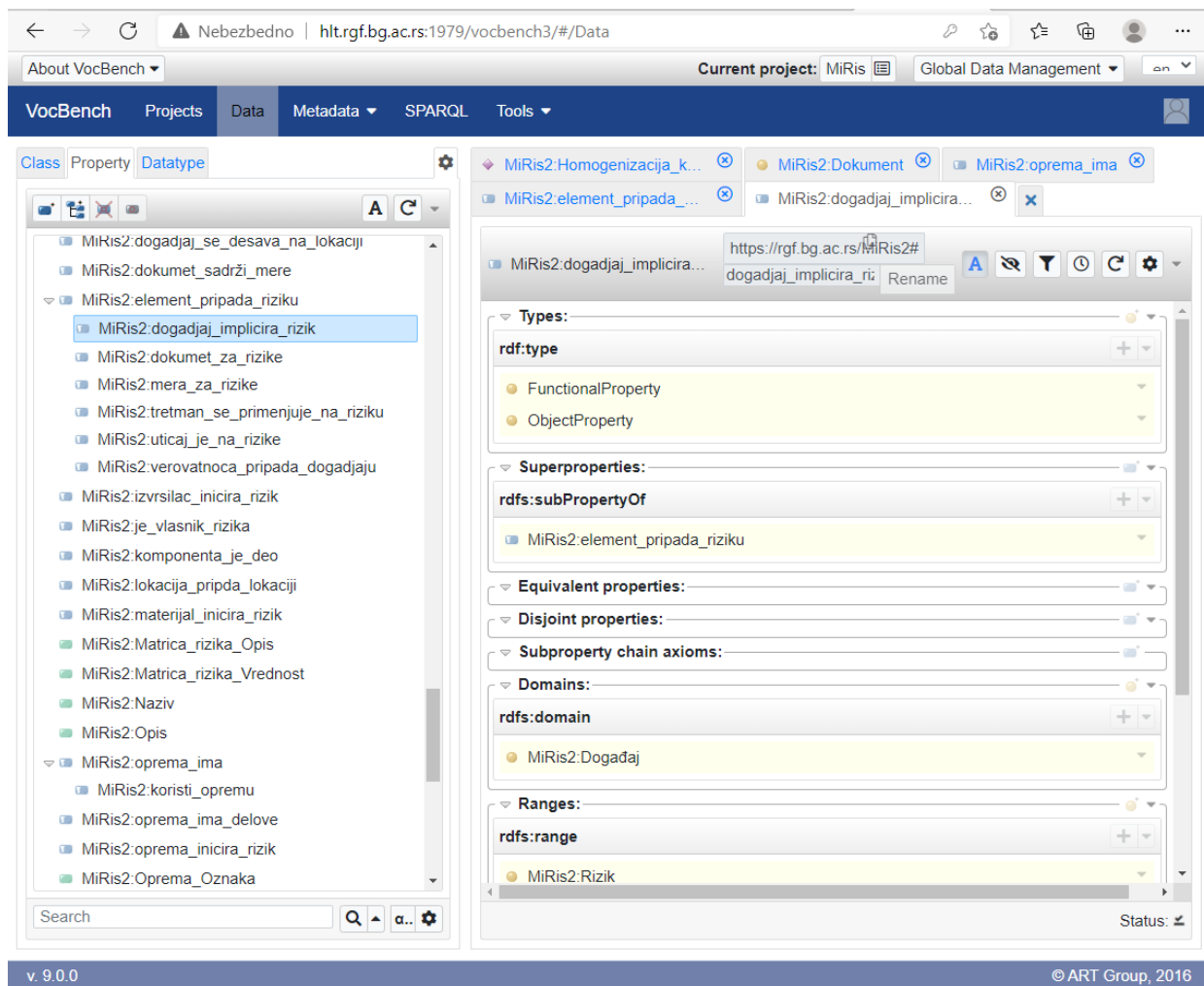
Za prikaz pojedinih delova sistema i grafova je već korišćena veb publikacija, a na slici ispod prikazujemo samo jedan od pogleda. Na prozoru gore levo je prikazana hijerarhija klasa, u okviru ispod se prikazuju primerci selektovane klase, dok su na desnoj strani prikazane informacije o semantičkim vezama sa drugim klasama ili instancama, u zavisnosti od izbora tipa objekta. U konkretnom primeru je prikazana instanca *Otkaz_drobnice_usled_loma_rotora* klase *Operativni_rizik*, koja je potklasa klase *Rizik*.



Slika 4-57 VocBench sa primerima klasa i primeraka publikovane ontologije

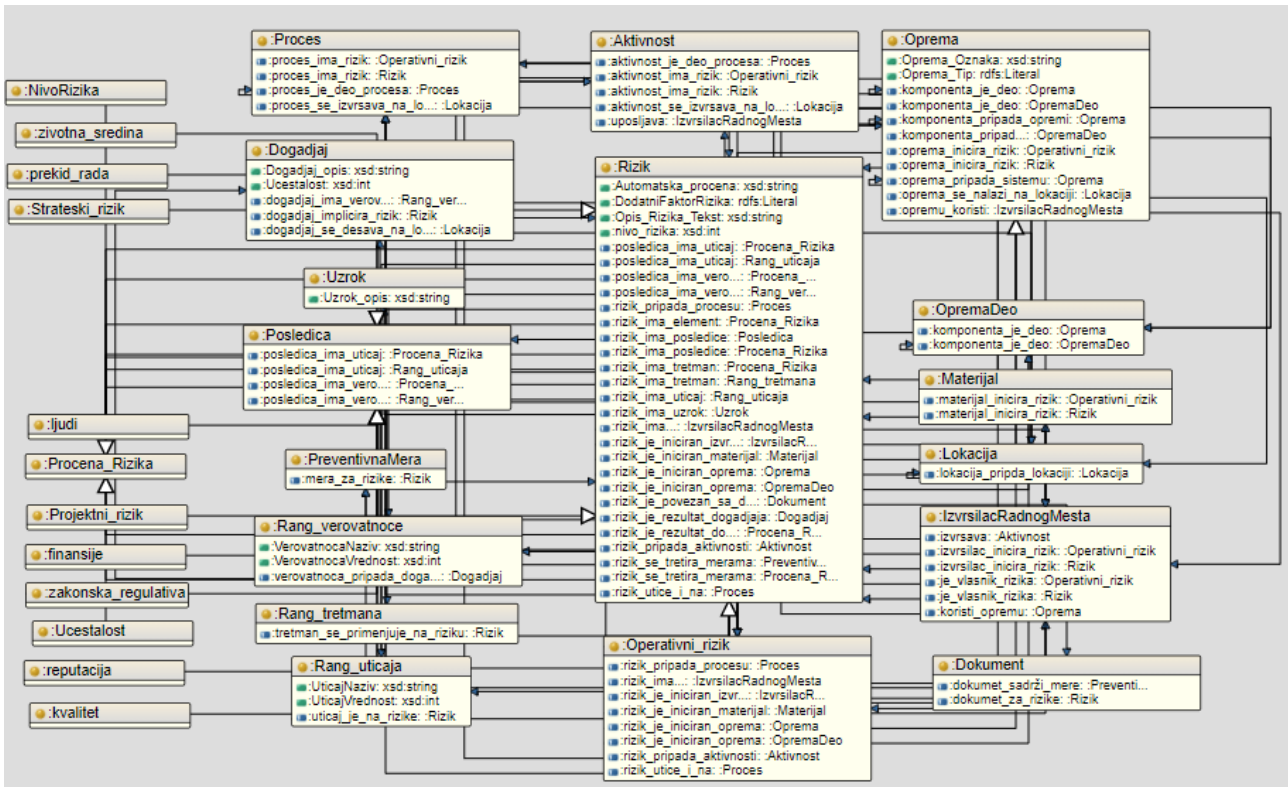
⁷⁵ <http://vocbench.uniroma2>

Slika 4-58 prikazuje na levoj strani hijerarhiju svojstva definisanih u ontologiji, dok su na desnoj strani prikazane informacije za selektovanu relaciju. U prikazanom primeru je selektovano svojstvo `dogadjaj_implicira_rizik`, kojim je uspostavlja funkcionalna relacija između klase `Događaj` i klase `Rizik`. Sa slike se može videti da je svojstvu `dogadjaj_implicira_rizik` nadređen `element_pripada_riziku`.



Slika 4-58 VocBench sa svojstvima publikovanje ontologije

Pogled na deo dijagrama klasa implementiranog ontološkog modela upravljanja rizikom u rudarstvu prikazuje Slika 4-59. Za svaku klasu su ispod naziva taksativno navedena svojstva, i to najpre atributivna, a zatim relacije ka drugim klasama. Relacije su dodatno predstavljene usmerenim strelicama ka povezanim klasama. Sa leve strane možemo videti potklase, koje nasleđuju sve osobine svojih nadređenih klasa i nemaju dodatne specifičnosti. Za razliku od njih na primeru operativnog rizika vidimo da on, iako je potklasa klase Rizik i nasleđuje sva njena svojstva, ima i svoje sopstvene osobine koja važe samo za tu potklasu, kao što je npr. veza za opremom, izvršiocem radnog mesta i slično.



Slika 4-59 VocBench sa svojstvima publikovanje ontologije

5. EVALUACIJA REZULTATA NA PRIMERIMA IZ PRAKSE

5.1. Evaluacija ekstrakcije informacija

5.1.1. Evaluacija pokrivenosti domena

Terminologija iz digitalizovanih rečnika rudarske terminologije na srpskom je proverena rečnicima SrpMD i rudarskim korpusom, radi dopunjavanja SrpMD novim rečima koje se nalaze i u terminološkim rečnicima i u korpusu. Od odrednica srpskog dela Rudarskog rečnika je kreiran tekst koji je potom analiziran SrpMD. Od 12.655 različitih pojedinačnih reči pronađenih u tekstu iz rečnika, SrpMD je prepoznao 9758. Među 2897 (23%) koji nisu prepoznati, bilo je akronima (npr. „pH“, „RR“, „LD“, „TV“), imena (npr. „Western“, „Bets“, „Reni“), arhaizmi (npr. „abanje“ umesto „habanje“, „bolcn“ umesto „zavrtnj“ itd.), kao i određenih grešaka u OCR-u (uprkos ručnoj proveru). Na osnovu ove analize pripremljen je skup kandidata za nove unose u SrpMD (npr. „degazacija“, „eksploatabilan“, „sabirnik“ itd.). Svaki kandidat je dalje proveren prema rudarskom korpusu i ako je rezultat (u osnovi, njegova učestalost) bio zadovoljavajući, dodat je u SrpMD.

Isti postupak je primenjen i na druge rečnike sa srpskim zapisima. Iako su sveobuhvatni terminološki rečnici sadržali puno prostih reči koje su nedostajale u SrpMD, manji rečnici su, kako se očekivalo, uključivali često korišćene izraze koji su uglavnom već bili u SrpMD. Tako, na primer, u Elektropediji⁷⁶ SrpMD nije prepoznao 13% reči, dok u srpskom delu englesko-srpskog rečnika terminologije u oblasti upravljanja otpadom nije prepoznato 6% reči. U svim ostalim rečnicima procenat neprepoznatih reči bio je između 3% -5%, ali da li će biti uključene u SrpMD zavisilo je od njihove učestalosti u rudarskom korpusu.

Pored digitalizovanih rečnika, srpski korpus i srpski deo dvojezičnog korpusa iz rudarskog domena bili su još jedan izvor novih termina domena sirovina koji nisu postojali u SrpMD. Izdvajanje jednostavnih reči bilo je relativno jednostavno, naime, reči koje SrpMD nije prepoznao bile su pažljivo proučene i ako su bile dovoljno česte, postale su kandidati za dodavanje u SrpMD. Pored toga, SrpMD nije prepoznao manje od 4% reči u jednojezičnom rudarskom korpusu, pri čemu je približno 1,3% od ovih 4% odgovarajućih kandidata za dodavanje u SrpMD, a preostale neprepoznate reči su promenljive iz jednačina (0,7%), akronimi (1%), retke reči (hapaksi i slovne greške - 0,5%), strana imena i reči (0,5%).

Međutim, kada je reč o terminima u obliku terminoloških fraza, njihovo izdvajanje iz narativnog teksta postaje mnogo komplikovanije. Automatska ekstrakcija kandidata višočlanih termina za srpski jezik oslanja se na postupak prikazan u (Krstev et al. 2015), a zasniva na otkrivanju niski reči koje prepoznaje neki od 23 specifična sintaksička obrasca za najčešće imeničke pojmove (AN pridev-imenica, NNg imenica-imenica u genitivu, AAN, ...). Prvi korak u ovom zadatku je prepoznavanje i izdvajanje srpskih terminoloških fraza iz korpusa pomoću sintaksičkih obrazaca i izračunavanje njihove učestalosti. Frekvencija je bila glavni parametar za određivanje ranga terminološke fraze kao kandidata za obradu za SrpMD. Međutim, korišćene su i druge mere pridruživanja, poput T-Score, Keiness-a, Log-verovatnoće, kao što je detaljno opisano u (Stanković, Krstev, Obradović, et al. 2016). Zadatak se zatim nastavlja lematizacijom terminoloških fraza kandidata, uklanjanjem višeznačnosti za terminološke fraze gde se može proizvesti više lema, a završava se izradom konačne leme, koja omogućava proizvodnju svih flektivnih oblika za verifikovane terminološke fraze.

Kao i u slučaju pojedinačnih pojmova, učestalost terminoloških fraza takođe je uzeta kao važan kriterijum, pri čemu je uvedena pomoćna mera koja izračunava prosečne frekvencije

⁷⁶ Electropedia <https://www.electropedia.org/iev/iev.nsf/Welcome?OpenForm&Seq=1>

komponenti višočlane reči. Imajući u vidu slobodan redosled reči u terminološkim frazama, tražili smo meru koja je labavija od tačnog podudaranja. Za svaku terminološku frazu čuvaju se sledeće informacije: minimalna, prosečna i maksimalna učestalost njegovih komponenti, broj „poznatih“ komponenti-reči koje prepoznaje SrpMD. Učestalost u korpusu i broj rečnika koji potvrđuju pojam glavni su kriterijumi za prioritet izraza nakon uređivanja.

Za ovaj rad izvršeno je izdvajanje srpskih terminoloških fraza sa pragom učestalosti 10, a proizvedeno je 12.632 fraza kandidata u lematizovanom obliku. (Kitanović et al. 2021) Učestalost svake terminološke fraze izračunata je kao zbir frekvencija svih njenih promenljivih oblika. Na primer, „kvalitet uglja“ (kvalitet uglja) ima frekvenciju 1110 kao zbir frekvencija njegovih oblika: „kvalitet uglja“ (172), „kvaliteta uglja“ (587), „kvalitetom uglja“ (284), „kvalitetu uglja“ (53), „kvaliteta uglja“ (8), „kvaliteti uglja“ (2), „kvalitetima uglja“ (4). Šest najproduktivnijih obrazaca, koji su dali 92% kandidata, navedeni su sa primerima i njihovim frekvencijama (Kitanović et al. 2021):

- NNgi (32%), N2X—imenica praćena rečju koja se u terminološkoj frazi ne menja. Obično je ova reč imenica u genitivu ili u instrumentalu; na primer ‘procena rizika’ (514).
- AN (29%), AXN— pridev iza koga sledi imenica; pridev i imenica moraju da se slažu u sve četiri gramatičke kategorije; na primer: ‘površinski kop’ (5738).
- NprepNp (11%), N4X—imenica za kojom slede dve reči koje se ne menjaju – čine predložku frazu; na primer ‘zdravlje na radu’ (1323).
- N-N (10%), NXN— imenica iza koje dolazi imenica koja se s njom slaže po broju i padežu, pri čemu separator može biti crtica; primeri su ‘bager vedričar’ (174). Ova klasa je imala najveći broj prepoznatih fraza za odbacivanje, odnosno onih čije su malo drugačije leme već bile obuhvaćene drugim obrascem, pa bi ovaj obrazac trebalo da bude stavljen sa nižim prioritetom u višeznačnosti (Kitanović et al. 2021).
- X-N (6%), 2XN— imenica kojoj prethodi reč koja se u terminološkoj frazi se ne menja. Obično se radi o reči koja se koristi samo u jednoj ili nekoliko terminoloških fraza, prefiksu ili prilogu izvedenom od prideva, dok separator može biti crtica; primer je ‘bto sistem’ (1728).
- NNgiNgi (4%), N4X— imenica za skojom slede dve reči koje se ne menjaju u terminološkoj frazi gde su ove dve reči pridevi / imenice u genitivu ili instrumentalu; na primer: ‘zaštita životne sredine’ (668).

Sledeći korak je evaluacija ekstrakcije, gde se proverava: da li je kandidat terminološka fraza, kom domenu pripada (rudarski, tehnički itd.) i eventualno kom poddomenu pripada (podzemna ili površinska eksploatacija, priprema mineralnih sirovina,...). Ako su domen ili poddomen identifikovani, terminološkoj frazi se dodeljuju odgovarajući semantički markeri: da li je u pitanju proces, materijal, profesija, oprema, mašina,... Nakon procesa evaluacije, sve pravilno procenjene terminološke fraze su importovane u terminološku bazu podataka Termi.

Preuzeto je ukupno 10.059 zapisa termina iz Rudarskog rečina, s tim da su skupovi engleskih pojmova upareni sa srpskim. Većina njih je naknadno obeležena domenom (24 različita), poddomenom (15) i semantičkim markerima (35).

Jednojezične liste izdvojenih pojmova dalje su proširene pojmovima preuzetim iz digitalizovanih rečnika koji daju 94.539 engleskih i 48.096 srpskih izraza.

Kao i u slučaju višejezičnih pojmova i terminoloških fraza, nakon procesa ocenjivanja, svi pravilno procenjeni dvojezični pojmovi su pripremljeni za ubacivanje u terminološku bazu podataka Termi. Do sada je ubačeno više od 3000 parova termina. U ovom procesu su objedinjeni da bi se formirali sinonimni skupovi (sinseti) koristeći informacije iz postojećih rečnika i jednostavnih pravila, kao što su: ako su dva engleska izraza prevedena istim srpskim izrazom, oni su kandidati za sinonime.

5.1.2. Evaluacija prepoznavanja domenskih entiteta

5.1.2.1. Preliminarni test

Preliminarnu evaluaciju transduktora za obeležavanje rudarskih entiteta je uradio ekspertskeg tim sastavljen od stručnjaka iz oblasti rudarstva. Pojedinačno (ručno) su pregledana sva obeležavanja opreme u tekstu i uočeno je da od ukupno 898 pojavljivanja entiteta opreme nisu prepoznata ukupno 4 (i to 3 različita). U slučaju obeležavanja lokacija, od ukupno 210 entiteta lokacija nisu prepoznata 2 (jedan koji se dva puta ponavlja). Takođe, samo je delimično ili sa greškama prepoznato 38 (25 različitih) entiteta opreme i 17 (4 različita) entiteta lokacija.

Dakle, rezultat prepoznavanja opreme daje preciznost i odziv su $894/898=99.55\%$, jer nije bilo pogrešno prepoznate opreme., a 4 od 898 nisu prepoznata i to:

- EŠ-evima (2 pojavljivanja)
- EŠ-eva
- BTO II

Ukoliko uračunamo i delimično prepoznavanje opreme kao negativan rezultat onda je preciznost 95.77%, tako da bi balansirana F-mera za opremu bi bila: 97.62%. U nastavku dajemo tipove parcijalnih obeležavanja:

- <OPREMA><tip>Bageri</tip></OPREMA> **dreglajni**
- <OPREMA><tip>bager dreglajn</tip><naziv> EŠ-6/45</naziv></OPREMA>-**II** (3 puta)
- <OPREMA><tip>bager dreglajn</tip><naziv> EŠ-6/45</naziv></OPREMA>-**III** (2 puta)
- <OPREMA><tip>bager dreglajn</tip><naziv> EŠ-6/45</naziv></OPREMA>-**I** (2 puta)
- <OPREMA><ostalo> **novopostavljenog** transporterera O5</ostalo></OPREMA> (3 puta)
- <OPREMA><tip>Bager vedričar</tip></OPREMA> (**V-1**) (2 puta)
- <OPREMA><tip>dreglajn</tip><naziv> EŠ-6/45</naziv></OPREMA>-**II** (2 puta)
- <OPREMA><ostalo>transportera O1 -</ostalo></OPREMA>
- <OPREMA><ostalo>transportera E3 -</ostalo></OPREMA> (3 puta)
- <OPREMA><ostalo>veznog transporterera -</ostalo></OPREMA>
- <OPREMA><ostalo>transportera O2 -</ostalo></OPREMA> x2
- <OPREMA><ostalo>transportera U5 -</ostalo></OPREMA>
- <OPREMA><ostalo>transportera M2 -</ostalo></OPREMA>
- <OPREMA><sistem><tip>BTD</tip></sistem></OPREMA>-**1** (3 puta)
- <OPREMA><sistem><tip>BTD</tip></sistem></OPREMA>-**2** (2 puta)
- <OPREMA><ostalo>**rada** transporterera</ostalo></OPREMA>
- <OPREMA><ostalo>Odlagač-2 **5CO2**</ostalo></OPREMA>
- <OPREMA><ostalo>Odlagač 3 -</ostalo></OPREMA>
- <OPREMA><ostalo>radi odlagač 3</ostalo></OPREMA>
- <OPREMA><tip>bager dreglajn</tip><naziv> EŠ 6/45</naziv></OPREMA> //**III**
- <OPREMA><ostalo>transportera</ostalo></OPREMA> **Omj** (2 puta)
- <OPREMA><tip>bager dreglajn</tip> (<naziv>EŠ 3)</naziv></OPREMA> (2 puta)

Na sličan način je evaluirano prepoznavanje lokacija, gde su 193 od 210 lokacija uspešno prepoznate, što je dalo odziv od 91.90%:

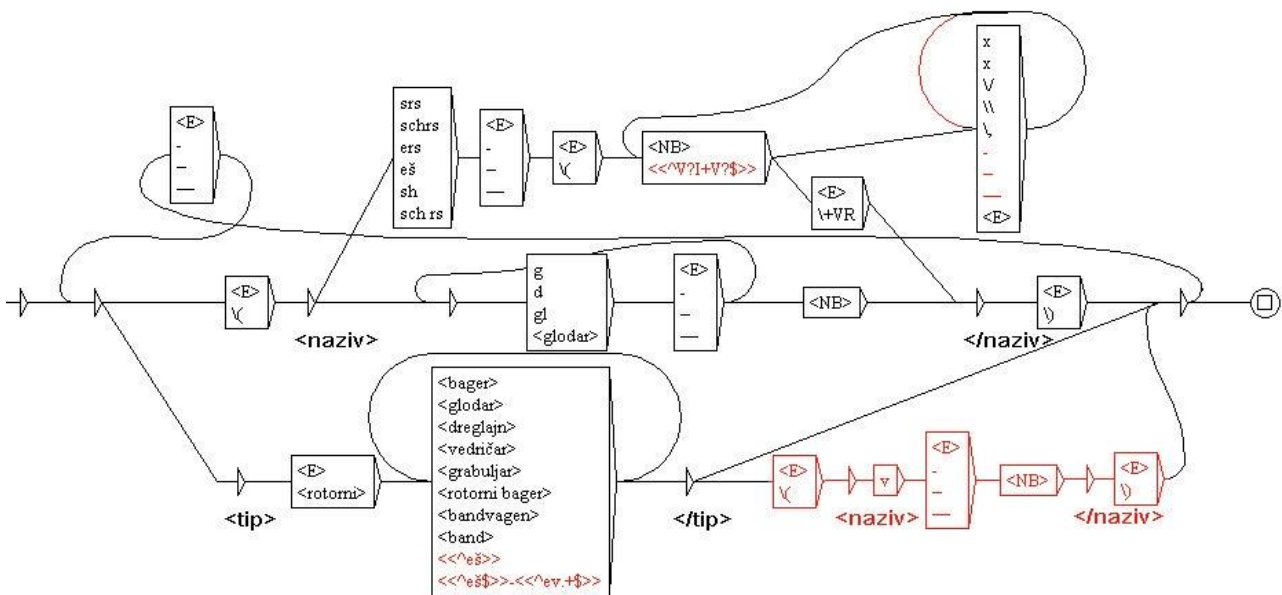
- ogranka „Kolubara-Površinski kopovi“ x9
- p.k. „Kolubara-Površinski kopovi“ x4
- Površinski kopovi – Baroševac x2
- Tamnavskih kopova x2

Delimično su prepoznate 2 lokacije, odnosno jedna lokacija navedena dva puta, što bi onda dalo preciznost 99.05%, a balansirana F-mera za lokacije bi bila: 95.27%

- <LOKACIJA type=polje>PK „RADLjEVO</LOKACIJA> SEVER“ x2

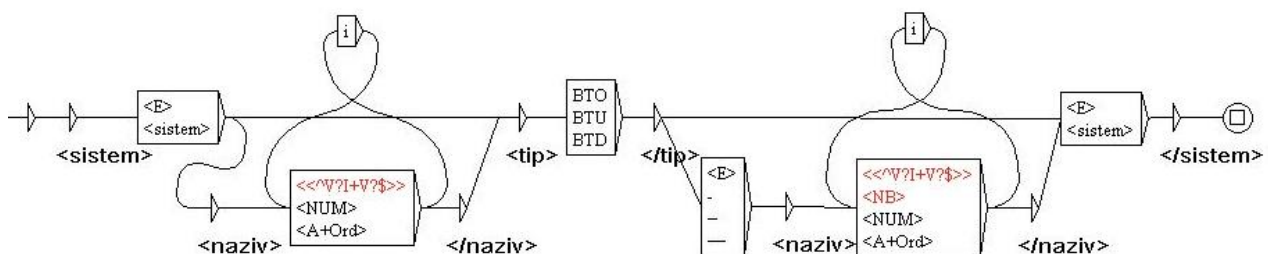
5.1.2.2. Ispravka transduktora za entitete rudarske opreme na osnovu rezultata evaluacije

Ispravke na glavnom transduktoru za prepoznavanje opreme svele su se na ispravljanje prepoznavanja entiteta rudarske opreme koji ne predstavljaju bagere (Slika 5-1). Kako bi se izbegli slučajevi ranije pomenutog parcijalnog prepoznavanja poput: *novopostavljenog transportera O5*, *rada transportera*, *radi odlagač 3*, leme rad, nov i novopostavljen su dodate u listu izuzetaka. Dodata je i mogućnost prepoznavanja tipova opreme koji se ne ispisuju isključivo velikim slovima (na primer: *transportera Omj*), i ukinuta je rekurzija nad slovima i brojevima kako bi se izbeglo prethodno pomenuto parcijalno prepoznavanje *Odlagač-2 5CO2*.



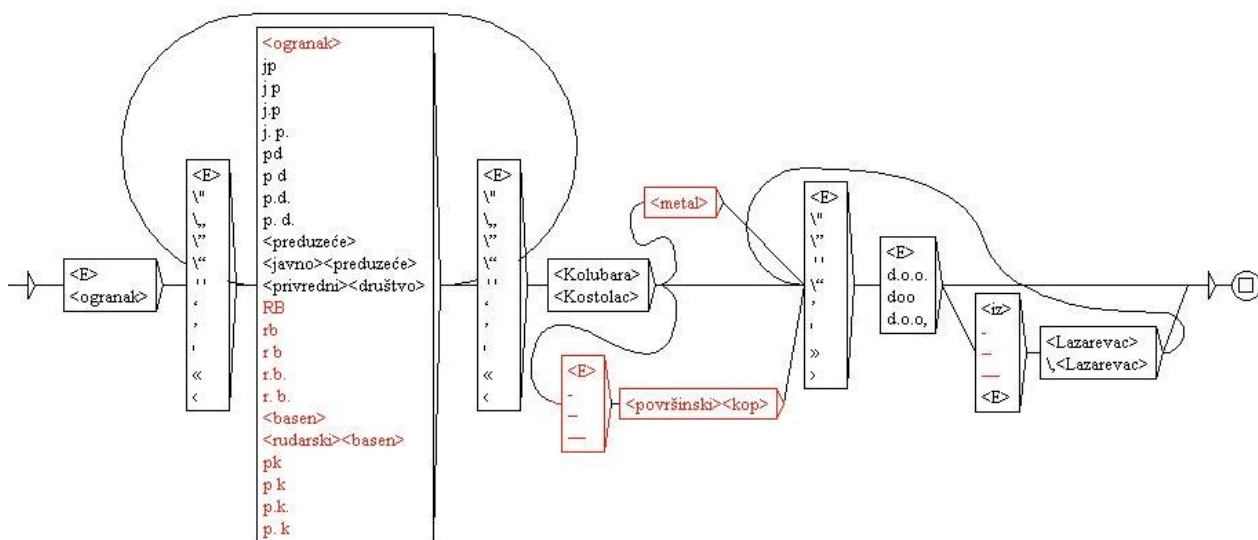
Slika 5-1 Ispravljeni transduktor za prepoznavanje opreme

Kod prepoznavanja proizvodnih sistema dorađeno je prepoznavanje brojnih oznaka (Slika 5-2) čime se sada prepoznaju primeri *BTD-1* i *BTD-2*.



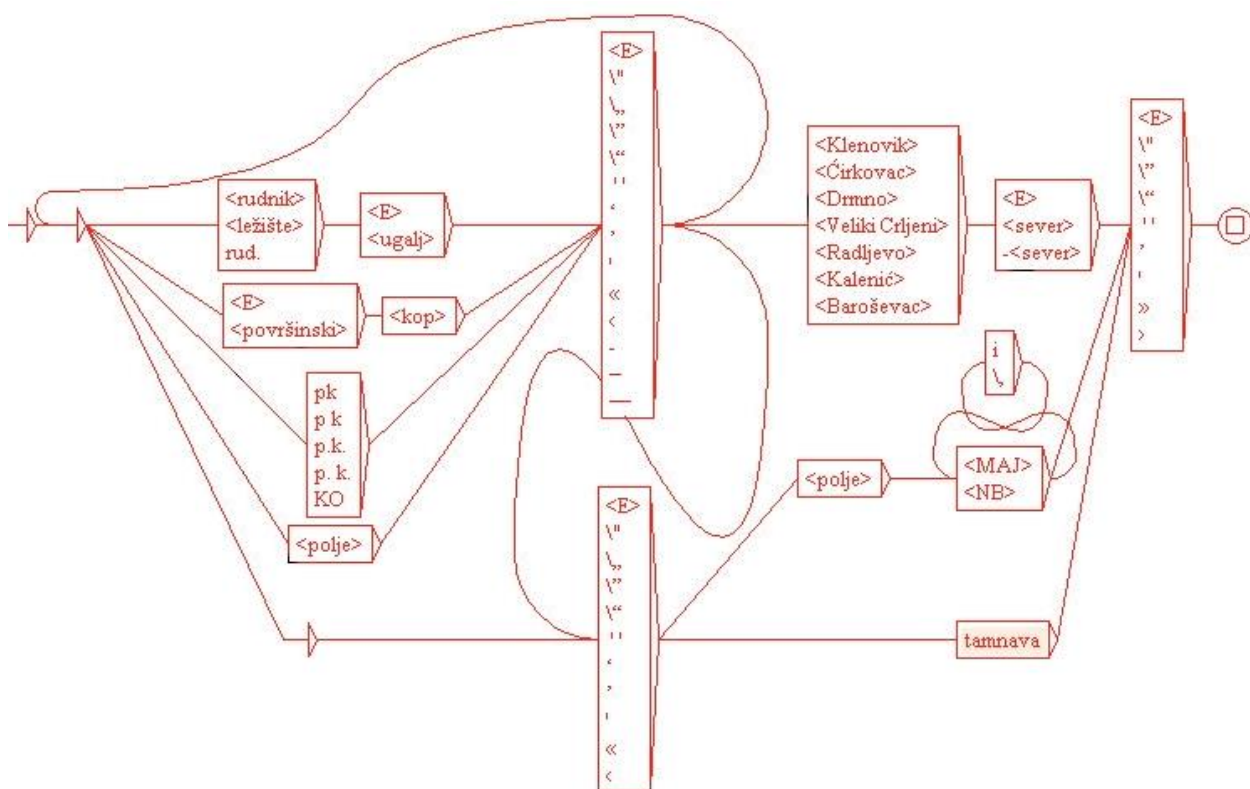
Slika 5-2 Ispravljeni podgraf za prepoznavanje bagerskih Sistema

Najveće izmene pretrpeo je podgraf koji prepoznaje bagere (Slika 5-3). Najpre je, kao i kod grafa za prepoznavanje sistema, dorađeno prepoznavanje brojeva, kako bi se ispravno prepoznali primeri: *bager dreglajn EŠ-6/45 -II*, *bager dreglajn EŠ-6/45 -III*, *bager dreglajn EŠ-6/45 -I*, *dreglajn EŠ-6/45 -II*, *bager dreglajn EŠ 6/45 //III*. Uklonjena je dupla zagrada na kraju prepoznavanja naziva kako bi se izbeglo parcijalno prepoznavanje naziva kod: *bager dreglajn*



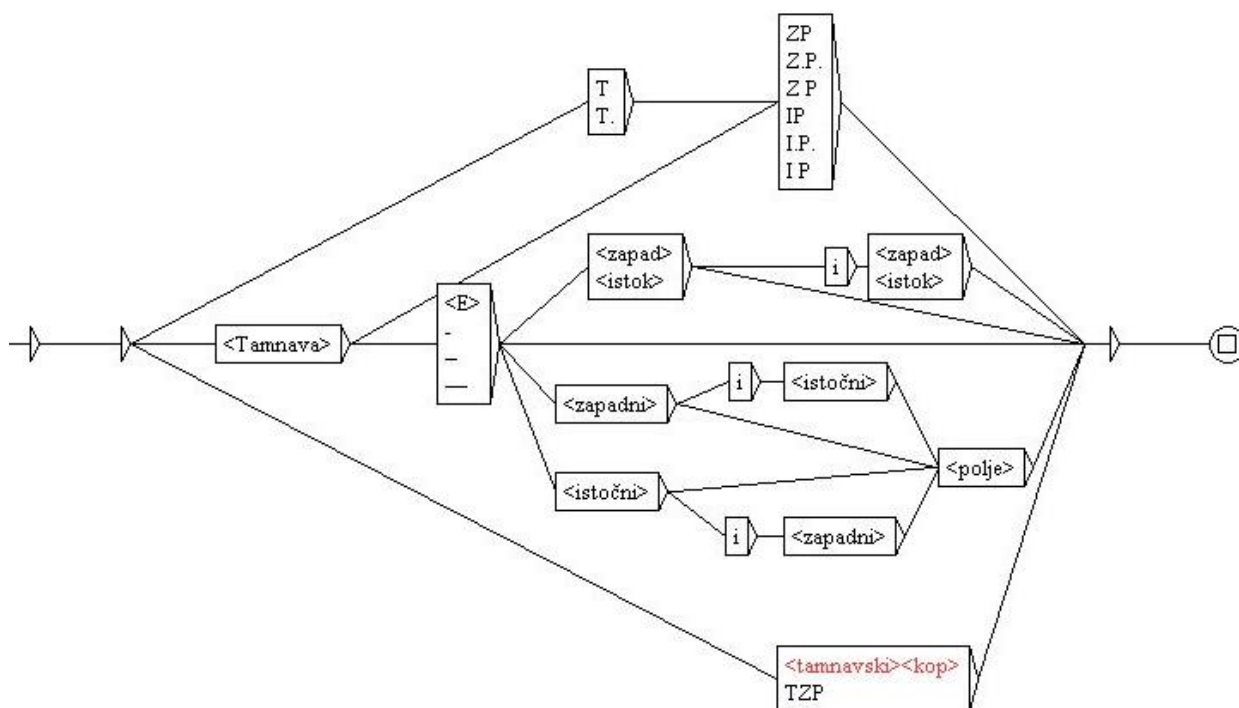
Slika 5-4 Ispravljeni podgraf za prepoznavanje ogranaka

Svi ostali problem su se odnosili na prepoznavanje kopova i polazišta ili odredišta transporta. Dodat je Baroševac kao vlastito ime polazišta kako bi se prepoznao primer *Površinski kopovi – Baroševac*, kao i opcija za sever bez prethodne crtica, kako bi se u potpunosti prepoznao entitet *PK „RADLjEVO SEVER“* iz evaluacije. Graf je u potpunosti prerađen kako bi se izbeglo prepoznavanje toponima (Slika 5-5), kao i sa Kolubarom u prethodnom primeru.



Slika 5-5 Ispravljeni podgraf za prepoznavanje kopova i polazišta ili odredišta

Takođe, kako bi bili prepoznati Tamnavski kopovi, izmenjen je i podgraf koji pronalazi pojavljivanje Tamnavskih polja (Slika 5-6).



Slika 5-6 Ispravljeni graf za pronalaženje entiteta tamnavskih polja

Nakon ispravki podgrafova za prepoznavanje lokacija, na test skupu više nije bilo grešaka i F-mera se dostigla 100% na test skupu. Svakako treba imati na umu da će se u nekim drugim dokumentima pojaviti obrasci koji nisu predviđeni implementiranim sistemom, ali je moguće prilagođavanje saglasno agilnim metodama razvoja.

5.1.2.4. Evaluacija prepoznavanja rudarskih događaja

Prepoznati događaji u tekstovima se, prema mišljenju ekspertskeg timu, nisu pokazali od koristi za automatsko unapređenje ontologije rizika? Prevelik broj vraćenih događaja. Odziv preko 95%, ali preciznost ispod 50% zajedno daju lošu balansiranu F meru i zato je odlučeno da se oni ekstrahuju podeljeni u grupe, pa da se ručno obrade i izaberu primeri koji mogu biti od značaja. Za tu namenu je istraživanje usmereno ka kauzalnim glagolima, koji se i u drugim istraživanjima (Todorascu, Romary, and Bekhouche 2002; Makki, Alquier, and Prince 2009; Leidner and Schilder 2010) koriste za „okidači“ („trigeri“) za uspostavljanje relacija.

5.2. Evaluacija na primeru upravljanja homogenizacijom uglja

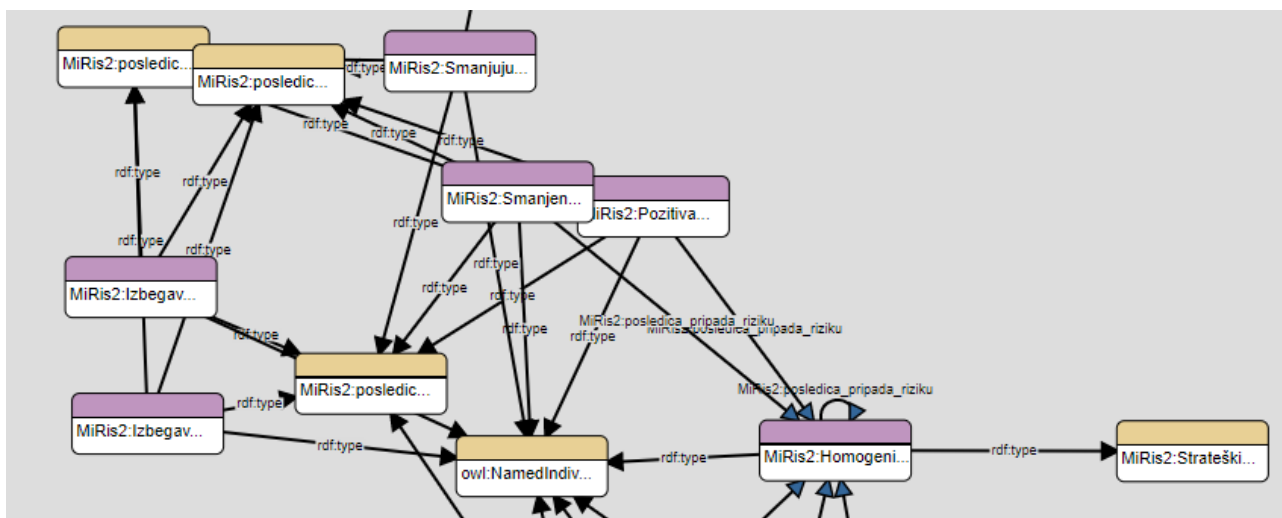
U okviru ove studije sprovedena je evaluacija modela podataka. U prvom delu razmatrano je uvođenje procesa homogenizacije kao prilike, odnosno rizika sa pozitivnim ishodom. Drugi deo je posvećen problemima, odnosno rizicima koji mogu negativno uticati na implementaciju sistema planirane i kontrolisane homogenizacije.

Homogenizacija uglja je proces mešanja uglja sa ciljem ujednačavanja njegovog kvaliteta. Osnovni cilj rudnika je da obezbedi dovoljne količine uglja zadovoljavajućeg kvaliteta za potrebe termoelektrane, uz minimalizaciju proizvodnih troškova. Problemi nastaju kada se eksploatacija vrši na delovima ležišta koja su slabijeg kvaliteta, što je iz perspektive rudnika opravdano, jer se ostvaruje veći stepen iskorišćenja ležišta, produžava se životni vek rudnika i eliminiše rizik od procesa samozapaljenja odbačenog uglja na odlagalištima jalovine. Sa aspekta termoelektrane prijem lošeg uglja iziskuje veću potrošnju mazuta, povećane troškove održavanja i slično.

Planskim pristupom homogenizaciji uglja i uspostavljanjem integralnog sistema upravljanja kvalitetom moglo bi se doći do pomirenja suprotstavljenih zahteva. Za potrebe istraživanja korišćena je monografija „Upravljanje kvalitetom uglja“ (Ignjatović et al. 2007). Znanje koje je pohranjeno u monografiji je izanalizirano i pretočeno u formalizovani oblik pogodan za model ontologije. Izolovani podaci pretočeni su u pozitivne posledice i importovani u ontologiju. Svaka posledica je kategorizovana, pri čemu neke pripadaju većem broju kategorija:

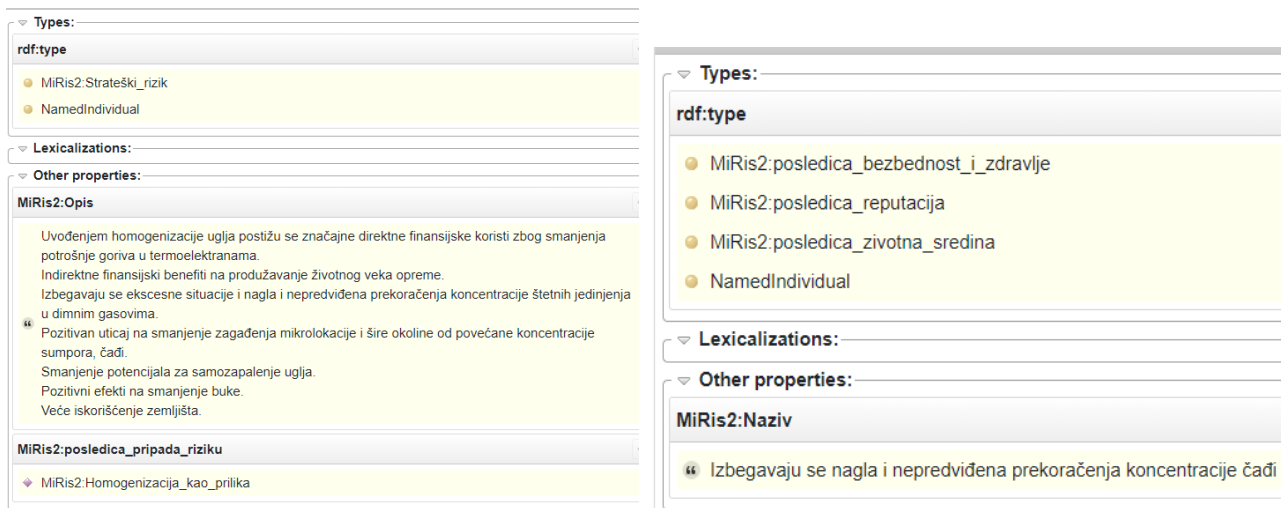
- ogromne direktne finansijske koristi zbog smanjenja potrošnje maziva i drugih troškova održavanja u termoelektranama, (finansijski),
- indirektni finansijski benefiti na produžavanje životnog veka opreme, (finansijski),
- smanjuju se incidenti, nagla i nepredviđena prekoračenja koncentracije štetnih jedinjenja u dimnim gasovima, kao što su sumpor, čađ koja utiču na aero zagađenje šire okoline, pojavu kiselih kiša. Izazivaju i probleme sa stanovništvom zbog povećanog nivoa oboljevanja disajnih puteva, (zakonska, životna sredina, reputacija, zdravlje),
- pozitivan uticaj na smanjenje zagađenja mikro lokacije prašinom, jedinjenjima sumpora, (životna sredina, reputacija, zdravlje),
- pozitivni efekti na mogućnosti samozapaljenja uglja, koje utiče na zagađenje šire okoline dimom, ugljenmonoksidom i sumpor dioksidom, (životna sredina, reputacija, zdravlje),
- pozitivni efekti na smanjenje buke, koja utiče na profesionalna oboljenja radnika i životnu sredinu okolnih naselja, (reputacija, zdravlje),
- veće iskorišćenje zauzetog zemljišta za odlaganje uglja, otkrivke i pepela, (životna sredina).

Slika 5-7 prikazuje deo grafičke reprezentacije pozitivnih efekata homogenizacije. Svaki čvor se može dodatno razgranati dopunjujući mrežu ostalim elementima za koje je povezan.



Slika 5-7Pozitivni uticaj uvođenja planirane homogenizacije

O svakom elementu mreže mogu se pregledati sve informacije koje su pohranjene u ontologiji, uključujući i veze koje ima sa ostalim instancama, kao što prikazuje Slika 5-8.



Slika 5-8 Opis elementa u mreži

Odlučivanje o uvođenju sistema za planiranje i upravljanje kontrolom kvaliteta uglja može se finansijski vrednovati, a na ovom primeru možemo videti koliko je visok uticaj i na sve ostale aspekte poslovanja rudnika, počevši od životne sredine, mikro i makro lokacije, zdravlje ljudi, reputaciju rudnika do zakonske regulative u slučaju prekoračenja koncentracije štetnih gasova.

U drugom delu studije izvršena je analiza i formalizacija znanja vezanih za potencijalne rizike za uspešnu implementaciju ovakvog projekta. Ovaj deo studije je baziran na implicitnim i eksplicitnim znanjima eksperata iz rudnika i sa Rudarsko-geološkog fakulteta. Tabela 5-1 Detektovani rizici koji ugrožavaju implementaciju. Tabela 5-1 prikazuje neke od karakterističnih primera formalizovanog implicitnog znanja.

U oba dela studije pokazalo se da je model prihvatljiv i da može da apsorbuje izdvojene podatke.

Tabela 5-1 Detektovani rizici koji ugrožavaju implementaciju homogenizacije

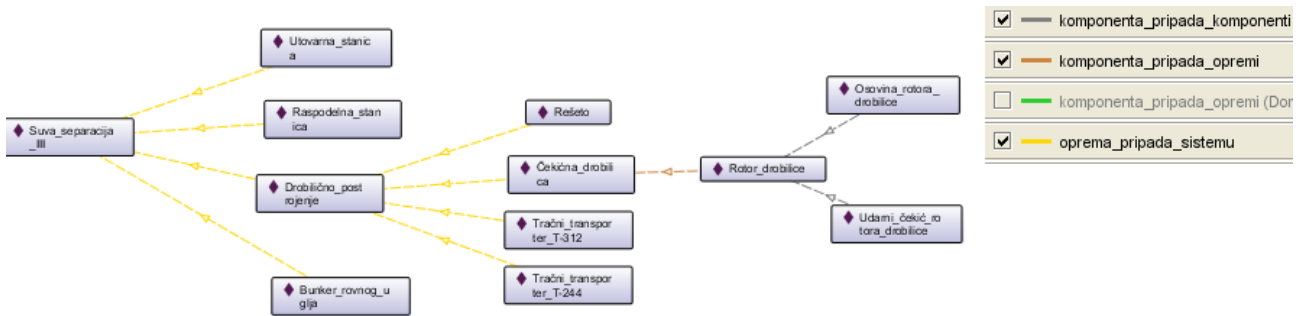
Proces	Rizik	Uzrok	Izvor	Vlasnik	Posledica
Geološka istraživanja	Neprecizan geološki model	Greške u laboratorijskim podacima; Greška u izradi modela	Laborant; Geolog	Laborant; Geolog	kvalitet
Planiranje proizvodnje	Neadekvatan plan	Neprecizan geološki model; Greška u izradi plana	Geolog; Planer	Geolog; Planer	kvalitet
Proizvodnja	Neizvršenje plana	Kvar i ispadanje delova sistema; Nepažnja operatera	Oprema; Bagerista;	Šef održavanja; Bagerista; Šef kopa	kvalitet; prekid rada
Deponovanje uglja	Rizik da nema mesta za odlaganje ili da je prazna	Greška u planiranju	Planer	Planer	kvalitet; prekid rada
Transport železnicom	Nedolazak vozova na vreme	Loša organizacija		Uprava	kvalitet; prekid rada

5.3.

5.4. Evaluacija na primeru procene rizika na procesa zaštite na radu

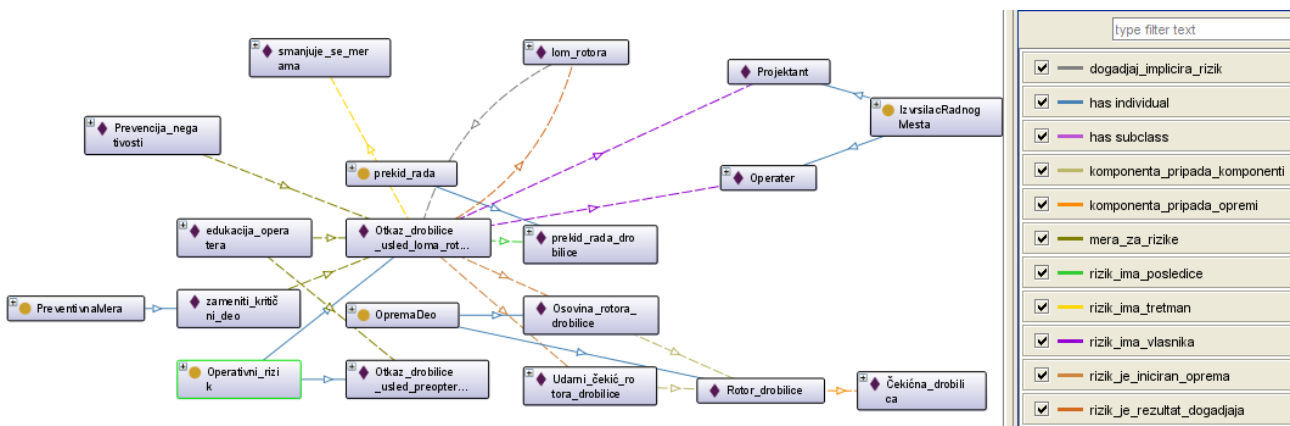
Sledeće istraživanje imalo je za cilj evaluaciju modela, kvaliteta ranije ekstrahovanih podataka i primenu pravila zaključivanja. Za potrebe evaluacije korišćeni su rezultati doktorske disertacije Slobodan D. Radosavljevića „Model procene rizika procesa zaštite na radu u pogonu suva separacija, "Kolubara prerada", Vreoci“, u okviru koje je izrađena analiza i procena rizika u okviru pogona za suhu separaciju (Radosavljević 2009).

Preslikavanje dekompozicije postrojenja za preradu izvršeno je u deo ontologije koji se odnosi na procese, opremu i komponente opreme. Slika 5-9 prikazuje samo deo klasifikacije koji je predmet analize rizičnih ponašanja.



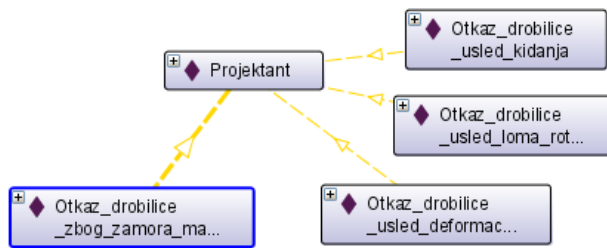
Slika 5-9 Dekompozicija postrojenja za preradu

Nakon toga su uneti rizici, događaji, uzroci i posledice povezani sa odgovarajućim komponentama postrojenja. Ocena rizika izvršena je kroz procenu verovatnoće da se može desiti neželjeni događaj i procene uticaja koji on može imati na rad sistema. U procenu je uključen i dodatni indeks koji služi proceniteljima da dodatno kvantifikuju rizik u situacijama kada je potrebno dodatno istaći opasnost. Na kraju su rizicima pridružene i preventivne mere. Slika 5-10 prikazuje opis jednog rizika sa svim njegovim elementima.



Slika 5-10 Primer vizuelizacije rizika sa njegovim elementima

Kroz ontološku mrežu se možemo kretati i dobijati dodatna saznanja. Na primer, pregledajući rizik vidimo da je za njega odgovoran projektant, proširivanjem čvora koji predstavlja projektanta, otvoriće se dodatni čvorovi za koje je on povezan, a u ovom slučaju to su svi rizici čiji je on vlasnik, što ilustruje Slika 5-11.



Slika 5-11 Primeri instanci vlasnik rizika

U studiji za većinu rizika imamo datu procenu verovatnoće i procenu uticaja, pa možemo primenom SWRL pravila izračunati nivo rizika.

Primer: svaki rizik koji imaj pridružen događaj i verovatnoću događaja i ima zadat uticaj i dodatni indeks pojačanja, preko vrednosti iz skale verovatnoću i skale uticaja izračunati nivo rizika:

```

Rizik(?r) ^ rizik_je_rezultat_dogadjaja(?r, ?d) ^
dogadjaj_ima_verovatnocu(?d, ?v) ^ VerovatnocaVrednost(?v, ?Vvalue) ^
rizik_ima_uticaj(?r, ?u) ^ UticajVrednost(?u, ?Uvalue) ^
DodatniFaktorRizika(?r, ?Fvalue) ^ swrlb:multiply(?rez, ?Vvalue, ?Uvalue) ^
swrlb:multiply(?rez2, ?rez, ?Fvalue) -> nivo_rizika(?r, ?rez2)
  
```

Kako su u podacima već sadržani parametri procene rizika, za potrebe evaluacije dodali smo novi atribut *Automatska procena* u klasi Rizik i primenili sledeće SWRL pravilo.

Primer: ukoliko u opisu rizika postoji reč lom i verovatnoća događaja je preko 4 reći ćemo da je rizik „ogroman“.

```

Pravilo: Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb: contains (?opis,
"lom") rizik_je_rezultat_dogadjaja(?r, ?d) ^ dogadjaj_ima_verovatnocu(?d,
?v) ^ VerovatnocaVrednost(?v, 5) ->Automatska_procena(?p, "ogroman")
  
```

SWRL pravila ne podržavaju „ili (eng. or)“ operator, pa se kao rešenje primenjuje veći broj srodnih pravila. Slična pravila definisana su i za pronalazke rizika sa rečima kvar, preopterećenje, zamor, gde je automatska procena „visok“.

Name	Rule	Comment
S1	Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb:contains(?opis, "lom") -> Automatska_procena(?p, "visok")	
S2	Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb:startsWith(?opis, "lom") -> Automatska_procena(?p, "kvar")	
S3	Rizik(?r) ^ rizik_je_rezultat_dogadjaja(?r, ?d) ^ dogadjaj_ima_verovatnocu(?d, ?v) ^ VerovatnocaVrednost(?v, 5) -> Automatska_procena(?p, "visok")	
S4	Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb:startsWith(?opis, "lom") -> Automatska_procena(?p, "visok")	
S5	Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb:startsWith(?opis, "kvar") -> Automatska_procena(?p, "visok")	
S6	Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb:startsWith(?opis, "kidanj") -> Automatska_procena(?p, "visok")	
S7	Rizik(?p) ^ Opis_Rizika_Tekst(?p, ?opis) ^ swrlb:startsWith(?opis, "preopter") -> Automatska_procena(?p, "visok")	

Control Rules Asserted Axioms Inferred Axioms OWL 2 RL

OWL axioms successfully transferred to rule engine.
 Number of SWRL rules exported to rule engine: 1
 Number of OWL class declarations exported to rule engine: 31
 Number of OWL individual declarations exported to rule engine: 45
 Number of OWL object property declarations exported to rule engine: 50
 Number of OWL data property declarations exported to rule engine: 14
 Total number of OWL axioms exported to rule engine: 664

Slika 5-12 Primeri SWRL pravila

U okviru studije uspešno je izvršeno preslikavanje strukturiranih podataka u odgovarajuće klase ontologije. Uočeno je da se u opisima rizika koriste ranije ekstrahovani termini kao što su preopterećenje, vibracije, udari, deformacije, lom, zamor, buka, oštećenje, habanje, grejanje.

6. ZAKLJUČAK

6.1. Doprinos

Rad u složenim uslovima i zahtevnim procesima zahteva brzo donošenje odluka, pa se vremenska komponenta u rešavanju problema izdvaja kao važan činilac. Cilj ovog istraživanja bio je da se pokaže da je moguće formalizovati implicitna i eksplicitna znanja koja mogu dati doprinos u upravljanju rizicima u rudarstvu, naročito sa aspekta bržeg donošenja kvalitetnijih odluka. Poseban je značaj u obuhvatanju takozvanog „skrivenog“ znanja, koje se nalazi u glavama stručnjaka, naročito kada se zna da rudarstvo zapošljava kadrove heterogenih zanimanja i da je za sticanje veština i sposobnosti jednog eksperta u rudniku potreban veliki broj godina. Ishod obuhvatanja ovog znanja jeste kreiranje baze znanja koja se mogu ponovo upotrebiti. Zabeleženi primeri dobre prakse mogu se ponovo iskoristiti za rešavanje konkretnih problema.

Upravljanje rizicima u rudarstvu uz pomoć ontologija do sada nije razmatrano, a ovakav pristup odabran je imajući u vidu sve veću primenu ontologija za rešavanje sličnih problema, te su obavljena potrebna istraživanja u cilju pronalaženja adekvatnog rešenja. Nakon pregleda i analize literature iz oblasti rizika generalno i rizika u rudarstvu, razmatrana su rešenja vezana za ontologije, sa posebnim osvrtom na one koje su povezane sa rizicima. Na osnovu prikupljenih saznanja kreiran je model ontologije za upravljanje rizicima u rudarskom domenu.

Paralelno sa ovim aktivnostima prikupljeni su tekstualni resursi, kao što su dokumenata iz zakonske regulative, standardi, primeri iz prakse, relevantna literatura i ostali resursi iz oblasti rudarstva i upravljanja rizikom. Od prikupljenih dokumenata kreiran je korpus tekstova, koji je integrisan u već postojeći rudarski korpus.

Izrada sistema za ekstrakciju termina obuhvatila je sledeće komponente:

- razvoj modela za automatsko označavanje (anotaciju) entiteta relevantnih za rizik (događaji, oprema, lokacija, akteri, kvarovi, štete, mere),
- izgradnju sistema za ekstrakciju informacija i znanja vezanih za rizik,
- razvoj sistema za upravljanje resursima, integrisanje razvijenih modela za anotaciju, ekstrakciju, analitičku obradu podataka i istraživanje podataka.

Kreirani su i demonstrirani primeri mehanizma zaključivanja kojima se na osnovu delimično formalizovanih podataka mogu izvesti zaključci, kao što je npr. predviđanje da se za određenu klasu uzroka i vrstu izvora dešava sasvim izvesno određena posledica. Pokazano je da se delimično strukturirani podaci putem zadatih mehanizama mogu dodatno dopuniti konkretnim informacijama, kao i da ontologija može obezbediti automatsku pomoć u klasifikaciji i proveriti konzistentnosti podataka.

Evaluacija ekstrakcije informacija je pokazala da je pokrivenost domena razvijenim resursima dovoljno dobra da se na elektronskim morfološkim i terminološkim resursima mogu razvijati lokalne gramatike za ekstrakciju koncepata, putem relacija: sinonimije, podređeni, nadređeni, kao i pomoću specifičnih kauzalnih odnosa. Posebna pažnja je posvećena evaluaciji kandidata za primerke (instance) klasa ontologije. Izuzetno uspešna ekstrakcija je postignuta za primerke opreme, lokacija i kvarove, dok sa događajima nije postignuto opšte rešenje. Kroz dve studije izvršena je evaluacija i testiranje ontološkog modela, gde se pokazalo da se novi podaci mogu formalizovati kroz implementirani ontološki model. Prilikom unosa podataka iz studije baza je već sadržala neke od pojmova, koji su dobijeni u procesu automatske ekstrakcije termina. Na nove podatke primenjena su pravila SWRL mehanizama za sticanje dodatnih znanja.

Ontološki model pruža mogućnost za pregled pojmova, definicija, kretanje kroz mrežu funkcionalnih odnosa i na taj način omogućava dolaženje do dodatnih saznanja. Jedan od

benefita jeste i uspostavljanje standardizacije terminologije iz oblasti upravljanja rizikom u rudarstvu i povezivanje sa istoznačnim odrednicama u drugim stranim ili domaćim elektronskim resursima, npr. klasa rizik je povezana sa terminom rizik u bazama znanja (WikiData, MeSH, Brilliant,...). Sa druge strane to ne predstavlja ograničenje da se za potrebe daljih ekstrakcija i pronalaženje informacija vezanih za rizike prihvate i neki lokalni izrazi u vidu sinonima.

Model upravljanja rizikom opisan u ovoj disertaciji može se dinamički proširiti novim ontologijama, odnosima, ograničenjima i pravilima, pri čemu se ranije postavljeni mehanizmi zaključivanja automatski primenjuju i na nove elemente, kada je to moguće. Kreirana ontologija se može dalje koristiti u ekstrakciji novih termina za sopstvenu dopunu, kao i za dopunu ostalih jezičkih resursa koji se godinama u nazad razvijaju na Rudarsko-geološkom fakultetu.

Ontologija je percipirana kao baza opšteg znanja, ali se može preuzeti i koristiti za potrebe konkretnih rudnika.

Implicitni uticaj publikovanja baze znanja i njene dostupnosti većem broju ljudi ogleđa se u promociji i podizanju svesti aktera u rudarskoj proizvodnji vezano za rizike, kada se ima u vidu da je najveći broj incidenata u rudarstvu izazvan ljudskom nepažnjom i nepoštovanjem preventivnih mera, kao i da su incidentalne situacije u najvećem procentu mogle da se predvide i preduprede.

6.2. Pravci za dalji rad

Pravci daljeg razvoja usmereni su na terminološku dopunu izgrađene ontologije i dopunu rudarskog korpusa, proširenje u pojedinačnim domenima rudarstva i izradu softverskih komponenti kojima bi se unapredio postojeći sistem.

Kontinuirano dopunjavanje rudarskog korpusa novim sadržajima je aktivnost koja bi svakako trebala da se nastavi u budućnosti. Na dopunjene resurse primenjivaće se mehanizmi za ekstrakciju razvijeni u okviru ove disertacije i izdvajati novi kandidati za odgovarajuće klase ontologije. Komponenta prepoznavanja događaja u tekstu zahteva značajno unapređenje, u prvom redu razlaganje na vrste događaja, te onda dalju implementaciju.

U okviru istraživanja, izgrađena je opšta rudarska ontologija, a u budućnosti bi se mogla primeniti na pojedinačne domene rudarstva, različite tipove rudnika sa različitim primenjenim tehnološkim procesima i okruženjem, pri čemu bi se izvršila dopuna novim konceptima specifičnim za sam domen primene. Primena modela razvijenog u ovoj disertaciji mogla bi se testirati i u drugim industrijama, kao što je recimo metalska.

Protégé alat pruža snažne mogućnosti za modeliranje, kao i za primenu mehanizma zaključivanja, ali mane su što nije višekorisnički orijentisan i interfejs za unos konkretnih podataka nije nimalo komforan za rad. Platforma VocBench prevazilazi delimično probleme jer nudi mogućnost rada više korisnika istovremeno i nešto poboljšan interfejs za unos instanci, ali je to opšti alat te na žalost ni on nema udoban korisnički interfejs za upravljanje podacima. Bilo bi značajno nastaviti izgradnju prezentovanog pilot rešenja te napraviti adekvatno softversko rešenje čime bi se obezbedila dopuna ontologije od strane raznih korisnika, koji ne moraju imati informatičko iskustvo. Dalja istraživanja će se fokusirati na nastavak razvoja veb aplikacije koja bi imala hibridno skladište podataka: deo bi bio u relacionoj bazi podataka, a deo u RDF skladištu, u svakom slučaju višekorisnički orjentisana platforma sa različitim profilima korisnika. Kreiranim mehanizmima podaci bi se preveli u ontologiju, gde bi se primenili postojeći ili kreirali novi mehanizmi obrade i zaključivanja. Nova verzija dopunjene ontologije bi se publikovala na veb, vodeći računa o kompatibilnosti sa prethodnim verzijama.

Predstavljeno rešenje u okviru ove disertacije orijentisano je na izradu opšte baze znanja, u cilju identifikacije, analize i procene rizika. Ukoliko bi se rudarska ontologija prilagodila potrebama određenog rudnika, mogla bi se kreirati dodatna mobilna aplikacija za monitoring čime bi se ceo proces upravljanja rizikom zaokružio. Aplikacija bi se oslanjala na ontološki model i pohranjene podatke, a osobama odgovornim za monitoring omogućila da na terenu kroz jednostavne ček liste sprovedu nadzor.

Dilema koja je bila prisutna tokom celog perioda istraživanja bila je da li se opredeliti za pristup jedne integralne ontologije, što je i usvojeno, ili bazirati rešenje na izdvajanje pojedinih klasa u posebne ontologije, tezauruse ili taksonomije koje bi se samostalno razvijale, a onda koristile kao domeni u centralnoj ontologiji. Ova dilema bi svakako mogla biti predmet daljih analiza i sa porastom broja instanci (recimo opreme) bi se taj deo izdvojio kao posebna ontologija. Imajući u vidu arhitekturu sistema i primenjene tehnologije, na samo korišćenje sistema to ne bi imalo nikakvog uticaja. U disertaciji je prevagnulo opredeljenje za izgradnju jedinstvene ontologije, sa ciljem uspostavljanja opšte baze znanja. Sa aspekta personalizovane ontologije za upotrebu u rudniku, možda bi bilo pogodnije rešenje sa nekoliko parcijalnih delova, gde bi se odgovarajuće službe fokusirale na svoj domen podataka.

7. LITERATURA

- Abarbanell, Jeffery S, William N Lanen, and Robert E Verrecchia. 1995. "Analysts' Forecasts as Proxies for Investor Beliefs in Empirical Research." *Journal of Accounting and Economics* 20 (1): 31–60.
- Abdelghany, Abdelghany Salah, Nagy Ramadan Darwish, and Hesham Ahmed Hefni. 2019. "An Agile Methodology for Ontology Development." *International Journal of Intelligent Engineering and Systems* 12 (2): 170–81.
- Alanen, Jarmo, Iiro Vidberg, Heikki Nikula, Nikolaos Papakonstantinou, Teppo Pirttioja, and Seppo Sierla. 2011. "Engineering Data Model for Machine Automation Systems."
- Allan, G. M., N. D. Allan, V. Kadiramanathan, and P. J. Fleming. 2007. "Risk Mining for Strategic Decision Making." In *Advances in Intelligent Web Mastering*, edited by Katarzyna M. Wegrzyn-Wolska and Piotr S. Szczepaniak, 21–28. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Andonovski, Jelena, Branislava Šandrih, and Olivera Kitanović. 2019. "Bilingual Lexical Extraction Based on Word Alignment for Improving Corpus Search." *The Electronic Library* 37 (2): 722–39. <https://doi.org/10.1108/EL-03-2019-0056>.
- Angelides, Demos, Yiannis Xenidis, Nick Bassiliades, Eva Loukogeorgaki, Alexandros Taflanidis, Dimitris Vrakas, Stella Arnaouti, and Georgios Meditskos. 2013. "The Development of a New Framework for Managing Risks in the European Industry: The IRIS RISK PARADIGM." *Risk* 1: 1.
- Assadi, H, and D Bourigault. 2000. "Analyses Syntaxique et Statistique Pour La Construction d'ontologies à Partir de Textes." *Charlet et Al*, 243–55.
- Atanacković, Nebojša B. 2018. "Procena Rizika Od Zaga\ djivanja Vodnih Resursa Pod Uticajem Napuštenih Rudarskih Radova Na Prostoru Srbije." PhD Thesis, Univerzitet u Beogradu-Rudarsko-geološki fakultet.
- Ayyub, Bilal M. 2014. *Risk Analysis in Engineering and Economics*. Crc Press.
- Badri, Adel, André Gbodossou, and Sylvie Nadeau. 2012. "Occupational Health and Safety Risks: Towards the Integration into Project Management." *Safety Science* 50 (2): 190–98.
- Beasley, Mark S, Richard Clune, and Dana R Hermanson. 2005. "Enterprise Risk Management: An Empirical Analysis of Factors Associated with the Extent of Implementation." *Journal of Accounting and Public Policy* 24 (6): 521–31.
- Beko, Lidija, Ivan Obradović, and Ranka Stanković. 2015. "Developing Students' Mining and Geology Vocabulary Through Flashcards and L1 in the CLIL Classroom." In *Proceedings of the Second International Conference on Teaching English for Specific Purposes and New Language Learning Technologies*. Niš, Serbia: Faculty of Electronic Engineering, University of Niš: Niš, Serbia.
- Beliga, Slobodan, Olivera Kitanović, Ranka Stanković, and Sanda Martinčić-Ipšić. 2017. "Keyword Extraction from Parallel Abstracts of Scientific Publications." In *Semanitic Keyword-Based Search on Structured Data Sources*, 44–55. Springer.
- Bernoulli, Daniel. 1896. *Specimen Theoriae Novae de Mensura Sortis*. Gregg.
- Blagojević, Danka, Stanković Ranka, Stejić Petar, and Nikolić Velizar. 2014. "Srbija u OneGeology Europe." *Zapisnici Srpskog Geološkog Društva Za 2013. Godinu*, 79–95.
- Bosque-Gil, Julia, Jorge Gracia, and Elena Montiel-Ponsoda. 2017. "Towards a Module for Lexicography in OntoLex." In *Proceedings of the LDK Workshops: OntoLex, TIAD and*

Challenges for Wordnets at 1st Language Data and Knowledge Conference (LDK 2017), Galway, Ireland, 1899:74–84.

- Bosque-Gil, Julia, Jorge Gracia, Elena Montiel-Ponsoda, and Guadalupe Aguado-de-Cea. 2016. "Modelling Multilingual Lexicographic Resources for the Web of Data: The K Dictionaries Case." In *Proc. of GLOBALEX'16 Workshop at LREC'16, Portoroz, Slovenia*, edited by Ilan Kernerman, Iztok Kosem, Simon Krek, and Lars Trap-Jensen, 65–72. European Language Resources Association. http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-GLOBALEX_Proceedings-v2.pdf.
- Bouaud, Jacques, Benoît Habert, Adeline Nazarenko, and Pierre Zweigenbaum. 2000. "Regroupements Issus de Dépendances Syntaxiques Sur Un Corpus de Spécialité: Catégorisation et Confrontation à Deux Conceptualisations Du Domaine." *Charlet, J., Zacklad, M., Kassel, G. et Bourigault, D., Éditeurs: Ingénierie Des Connaissances: Évolutions Récentes et Nouveaux Défis, Chapitre 17*: 275–90.
- Capponi, Nicolas, and Yannick Toussaint. 2000. *Interprétation de Classes de Termes Par Généralisation de Structures Prédicat-Arguments*. Eyrolles.
- Chen, Liang-Ching, Kuei-Hu Chang, and Hsiang-Yu Chung. 2020. "A Novel Statistic-Based Corpus Machine Processing Approach to Refine a Big Textual Data: An ESP Case of COVID-19 News Reports." *Applied Sciences* 10 (16). <https://doi.org/10.3390/app10165505>.
- Cox, Simon JD. 2004. "XXML? A Standards Conformant XML Language for Transfer of Exploration Data." *ASEG Extended Abstracts 2004* (1): 1–4.
- Crawley, Frank, Malcom Preston, and Brian Tyler. 2000. "HAZOP Guidelines: Guide to Best Practice (Guidelines to Best Practice for the Process and Chemical Industries)." *European Process Safety Centre/Institution of Chemical Engineers, Rugby*.
- Crawley, Frank, and Brian Tyler. 2015. *HAZOP: Guide to Best Practice*. Elsevier.
- Daille, Béatrice. 1994. "Study and Implementation of Combined Techniques for Automatic Extraction of Terminology." In *The Balancing Act: Combining Symbolic and Statistical Approaches to Language*.
- Dasgupta, Tirthankar, and Lipika Dey. 2016. "Enterprise Risk Analytics: Automatic Analysis of Risk Factors from Textual Feedbacks." In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 001735–40. <https://doi.org/10.1109/SMC.2016.7844488>.
- Dasgupta, Tirthankar, Lipika Dey, Prasenjit Dey, and Rupsa Saha. 2016. "A Framework for Mining Enterprise Risk and Risk Factors from News Documents." In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, 180–84. Osaka, Japan: The COLING 2016 Organizing Committee. <https://www.aclweb.org/anthology/C16-2038>.
- Dechow, Patricia M, and Ilia D Dichev. 2002. "The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors." *The Accounting Review* 77 (s-1): 35–59.
- Đorđević, Bojana. 2017. "Izrada Osnova Formalne Gramatike Srpskog Jezika Upotrebom Metagramatike." *Универзитет у Београду*.
- Dowd, Peter. 1994. "Risk Assessment in Reserve Estimation and Open-Pit Planning." *Transactions of the Institution of Mining and Metallurgy, Section A: Mining Technology* 103: 148–54.
- Elms, DG. 1992. *Risk Assessment. Engineering Safety. DI Blockley*. London, Mc Graw-Hill.
- Evert, Stefan. 2005. "The CQP Query Language Tutorial." *IMS Stuttgart. CWB Version 2*: b90.

- Fillmore, C. 1982. "Frame Semantics." *Linguistics in the Morning Calm*, 111–37.
- Friburger, Nathalie. 2002. "Reconnaissance Automatique Des Noms Propres: Application à La Classification Automatique de Textes Journalistiques." PhD Thesis, Tours.
- Galatescu, Alexandra, Adriana Alexandru, Corneliu Zaharia, and Stefan Kovacs. 2010. "Ontology-Based Modeling and Inference for Occupational Risk Prevention." In *SEMAPRO 2010: The Fourth International Conference on Advances in Semantic Processing*, 205–12.
- Gašević, Dragan, Dragan Đurić, and Vladan Devedžić. 2006. *Model Driven Architecture and Ontology Development*. Springer Science & Business Media.
- Godoy, M., and R Dimitrakopoulos. 2004. "Managing Risk and Waste Mining in Long-Term Production Scheduling of Open-Pit Mines." In .
- Graham, R. L. 1996. *Dictionary of Mining, Mineral & Related Terms*. US Bureau of Mines.
- Gross, Maurice. 1987. "The Use of Finite Automata in the Lexical Representation of Natural Language." In *LITP Spring School on Theoretical Computer Science*, 34–50. Springer.
- Gruber, Thomas R. 1995. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing?" *International Journal of Human-Computer Studies* 43 (5–6): 907–28.
- Hearst, Marti A. 1992. "Automatic Acquisition of Hyponyms from Large Text Corpora." In *Coling 1992 Volume 2: The 15th International Conference on Computational Linguistics*.
- Henley, Ernest J, and Hiromitsu Kumamoto. 1996. "Probabilistic Risk Assessment and Management for Engineers and Scientists." *IEEE Press (2nd Edition)*.
- Horridge, Matthew, Holger Knublauch, Alan Rector, Robert Stevens, and Chris Wroe. 2004. "A Practical Guide to Building OWL Ontologies Using the Protégé-OWL Plugin and CO-ODE Tools Edition 1.0." *University of Manchester*.
- Ignjatović, Dragan, Dinko Knežević, Božo Kolonja, Nikola Lilić, and Ranka Stanković. 2007. *Upravljanje Kvalitetom Uglja*. Univerzitet u Beogradu, Rudarsko-geološki fakultet.
- Jian Pei, Jiawei Han, B. Mortazavi-Asl, H. Pinto, Qiming Chen, U. Dayal, and Mei-Chun Hsu. 2001. "PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth." In *Proceedings 17th International Conference on Data Engineering*, 215–24. <https://doi.org/10.1109/ICDE.2001.914830>.
- Justeson, John S., and Slava M. Katz. 1995. "Technical Terminology: Some Linguistic Properties and an Algorithm for Identification in Text." *Natural Language Engineering* 1 (1): 9–27. <https://doi.org/10.1017/S1351324900000048>.
- Karović, S, and N Komazec. 2010. "Upravljanje Rizikom Na Sistemskim Osnovama." *Vojno Delo* 62: 228–40.
- Kirin, Snežana, Wei Li, Miodrag Brzaković, Igor Miljanović, and Aleksandar Sedmak. 2020. "Rules of Risk Management-Case Study of Open Pit Mine." *Procedia Structural Integrity* 28: 764–69.
- Kirin, Snežana, Milorad Pantelic, Igor Miljanović, and Violeta Talovic. 2013. "RISK MANAGEMENT OF MACHINES IN OPEN PIT MINES." *STRUCTURAL INTEGRITY AND LIFE-INTEGRITET I VEK KONSTRUKCIJA* 13 (3): 197–202.
- Kirin, Snežana, Petar Stanojević, Igor Miljanović, Aleksandar Sedmak, Tijana Perić, and Petar Ilić. 2015. "Influence of the Human Factor on Risks in an Open-Pit Mine." *Structural Integrity and Life* 15 (2): 117–28.

- Kitanović, Olivera. 2016. "Short Term Scientific Mission Report, IC1302 KEYSTONE COST Action."
- Kitanović, Olivera, Ranka Stanković, Aleksandra Tomašević, Mihailo Škorić, Ivan Babić, and Ljiljana Kolonja. 2021. "A Data Driven Approach for Raw Material Terminology." *Applied Sciences* 11 (7): 2892.
- Knublauch, Holger, Ray W Ferguson, Natalya F Noy, and Mark A Musen. 2004. "The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications." In *International Semantic Web Conference*, 229–43. Springer.
- Koeva, Svetla, Cvetana Krstev, and Duško Vitas. 2008. "Morpho-Semantic Relations in Wordnet—a Case Study for Two Slavic Languages." In *Proceedings of Global WordNet Conference 2008*, 239–53. University of Szeged, Department of Informatics.
- Kogan, Shimon, Dimitry Levin, Bryan R Routledge, Jacob S Sagi, and Noah A Smith. 2009. "Predicting Risk from Financial Reports with Regression." In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 272–80.
- Kolonja, Ljiljana. 2016. "Sistem Poslovne Inteligencije Za Upravljanje Zaštitom Na Radu u Rudarskoj Industrijii." *Универзитет у Београду*.
- Kolonja, Ljiljana, Ranka Stanković, Ivan Obradović, Olivera Kitanović, and Aleksandar Cvjetić. 2016. "Development of Terminological Resources for Expert Knowledge: A Case Study in Mining." *Knowledge Management Research & Practice* 14 (4): 445–56. <https://doi.org/10.1057/kmrp.2015.10>.
- Krstev, Cvetana. 2008. *Processing of Serbian - Automata, Texts and Electronic Dictionaries*. Faculty of Philology of the University of Belgrade, Belgrade, Serbia.
- Krstev, Cvetana, Ranka Stanković, Ivan Obradović, and Biljana Lazić. 2015. "Terminology Acquisition and Description Using Lexical Resources and Local Grammars." In *Proceedings of the 11th International Conference on Terminology and Artificial Intelligence*, edited by Thierry Poibeau and Pamela Faber, 1495:81–89. CEUR Workshop Proceedings. CEUR-WS.org. http://ceur-ws.org/Vol-1495/paper_13.pdf.
- Krstev, Cvetana, Ranka Stanković, and Duško Vitas. 2018. "Knowledge and Rule-Based Diacritic Restoration in Serbian." In *Proceedings of the Third International Conference Computational Linguistics in Bulgaria (CLIB 2018)*, edited by Svetla Koeva, 41–51. Sofia, Bulgaria: Institute for Bulgarian Language "Prof. Lyubomir Andreychin", Bulgarian Academy of Sciences.
- Krstev, Cvetana, Duško Vitas, and R Stanković. 2015. "A Lexical Approach to Acronyms and Their Definitions." In *Proceedings of the 7th Language & Technology Conference*, 219–23.
- Krstev, Cvetana, Anđjelka Zečević, Duško Vitas, and Tita Kyriacopoulou. 2013. "NERosetta for the Named Entity Multi-Lingual Space." In *Language and Technology Conference*, 327–40. Springer.
- Lazić, Predrag. 2020. *Terminološki Rečnik Iz Pripreme Mineralnih Sirovina [The Terminological Dictionary of the Mineral Processing]*. draft version.
- Lazić, Svetislav. 1976. *Englesko-Hrvatsko-Srpski Naftni Rječnik [English Croatian-Serbian Petroleum Dictionary]*. Poslovno udruženje Nafta: Zagreb, Hrvatska.
- Leidner, Jochen L, and Frank Schilder. 2010. "Hunting for the Black Swan: Risk Mining from Text." In *Proceedings of the ACL 2010 System Demonstrations*, 54–59.

- Lilić, Nikola, Srećko Ćurčić, and Milan Radosavljević. 2009. "Risk Assessment and Managing Technical Systems in Case of Mining Industry." *Strojniški Vestnik-Journal of Mechanical Engineering* 55 (2): 119–13.
- Lilić, Nikola, Ivan Obradović, and Aleksandar Cvjetić. 2010. "An Intelligent Hybrid System for Surface Coal Mine Safety Analysis." *Engineering Applications of Artificial Intelligence* 23 (4): 453–62.
- Lilic, NIKOLA, IVAN Obradovic, and RANKA Stankovic. 2002. "A Knowledge-Based Approach to Mine Ventilation Planning in Yugoslav Mining Practice." *Mineral Resources Engineering* 11 (4): 361–82.
- Lu, Hsin-Min, Nina WanHsin Huang, Zhu Zhang, and Tsai-Jyh Chen. 2009. "Identifying Firm-Specific Risk Statements in News Articles." In *Intelligence and Security Informatics*, edited by Hsinchun Chen, Christopher C. Yang, Michael Chau, and Shu-Hsing Li, 42–53. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Lykourantzou, Ioanna, Katerina Papadaki, Apostolis Kalliakmanis, Younes Djaghloul, Thibaud Latour, Ioannis Charalabis, and Epaminondas Kapetanios. 2011. "Ontology-Based Operational Risk Management." In *2011 IEEE 13th Conference on Commerce and Enterprise Computing*, 153–60. IEEE.
- Makki, Jawad, Anne-Marie Alquier, and Violaine Prince. 2008a. "An NLP-Based Ontology Population for a Risk Management Generic Structure." In *Proceedings of the 5th International Conference on Soft Computing as Transdisciplinary Science and Technology*, 350–55. CSTST '08. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/1456223.1456296>.
- . 2008b. "Semi Automatic Ontology Instantiation in the Domain of Risk Management." In *International Conference on Intelligent Information Processing*, 254–65. Springer.
- . 2009. "Ontology Population via NLP Techniques in Risk Management." *International Journal of Humanities and Social Science (IJHSS)* 3 (3): 212–17.
- McCrae, John, Guadalupe Aguado-de-Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asunción Gómez-Pérez, Jorge Gracia, et al. 2012. "Interchanging Lexical Resources on the Semantic Web." *Language Resources and Evaluation* 46 (4): 701–19. <https://doi.org/10.1007/s10579-012-9182-3>.
- McCrae, John, Dennis Spohr, and Philipp Cimiano. 2011. "Linking Lexical Resources and Ontologies on the Semantic Web with Lemon." In *The Semantic Web: Research and Applications: 8th Extended Semantic Web Conference, ESWC 2011, Heraklion, Crete, Greece, May 29-June 2, 2011, Proceedings, Part I*, 245–59. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21034-1_17.
- Meditkos, Georgios, Nick Bassiliades, Dimitris Vrakas, and Ioannis Vlahavas. 2012. "IRISPortal: A Semantic Portal for Industrial Risk Cases Management." In *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics*, 1–8.
- Mel'cuk, Igor. 1995. "Phrasemes in Language and Phraseology in Linguistics." *Idioms: Structural and Psychological Perspectives*, 167–232.
- Miljanović, Igor, Aleksandar Milutinović, and Rade Tokalić. 2016. "Risk Management in Coal Mining." In *Proceedings, 2nd Virtual International Conference on Science, Technology and Management in Energy*, 89–96.
- Milutinović, Aleksandar, Igor Miljanović, and Biljana Panić. 2013. "Qualification of Open Pit Mines Based on External Risks Assessment." In *Proceedings of the XI Balkan Conference on Operations Research BALCOR, 7-11.09. 2013.*, 705–9.

- Mladenović, Miljana, Ranka Stanković, and Cvetana Krstev. 2017. "A WordNet Ontology in Improving Searches of Digital Dialect Dictionary." In *European Conference on Advances in Databases and Information Systems*, 373–83. Springer.
- Nešić, Gojko. 1970. *Rudarski Rečnik: Srpsko-Hrvatski: English: Français: Deutsch: Russkij [Mining Dictionary: Serbo-Croatian: English: French: German: Russian]*. Rudarski institut [Mining institute], Beograd, Srbija.
- Nišić, Dragana. 2019. "Razvoj Modela Za Procenu Rizika Na Deponijama Industrijskog Otpada Mineralnog Porekla." Univerzitet u Beogradu, Rudarsko-geološki fakultet.
- Nišić, Dragana, Dinko Knežević, and Nikola Lilić. 2018. "Assessment of Risks Associated with the Operation of the Tailings Storage Facility Veliki Krivelj, Bor (Serbia)." *Archives of Mining Sciences* 63 (1).
- Nišić, Dragana, Dinko Knežević, Nevena Sijerković, Uroš Pantelić, and Mirjana Banković. 2016. "Comparative Risk Assessment of Exploitation of Old and New Thermal Power Plant's Ash Disposal Sites in Kostolac, According to Hydrological Scenario." *Tehnika* 71 (5): 677–84.
- Nota, Giancarlo, Rossella Aiello, and Maria Pia Di Gregorio. 2010. "Ontology Based Risk Management." In *Decision Theory and Choices: A Complexity Approach*, 235–51. Springer.
- Noy, Natalya F, Deborah L McGuinness, and others. 2001. *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford knowledge systems laboratory technical report KSL-01-05 and
- Obradović, Ivan, Ranka Stanković, Jelena Prodanović, and Olivera Kitanović. 2013. "A TEL Platform Blending Academic and Entrepreneurial Knowledge." In *Proceedings of the Fourth International Conference on E-Learning (ELearning-2013)*, 65–70. Belgrade, Serbia: Belgrade Metropolitan University Belgrade.
- Obradović, Ivan, Aleksandra Tomašević, Ranka Stanković, and Biljana Lazić. 2017. "Uvođenje domenskih i semantičkih markera za oblast rudarstva u srpske elektronske rečnike." In *Srpski jezik i njegovi resursi: teorija, opis i primene*, edited by Rajna Dragičević and Aleksandar Milanović, 46:147–58. Naučni sastanak slavista u Vukove dane 3. Beograd: Međunarodni slavistički centar, Filološki fakultet, Univerzitet u Beogradu. http://doi.fil.bg.ac.rs/volume.php?pt=eb_ser&issue=msc-2017-46-3&i=10.
- O'Connor, Martin J, Ravi D Shankar, Csongor Nyulas, Samson W Tu, and Amar K Das. 2008. "Developing a Web-Based Application Using OWL and SWRL." In *AAAI Spring Symposium: AI Meets Business Rules and Process Management*, 93–98.
- Olofsson, Tobias. 2020. "Imagined Futures in Mineral Exploration." *Journal of Cultural Economy* 13 (3): 265–77.
- Pajić, Vesna, Staša Vujičić Stanković, Ranka Stanković, and Miloš Pajić. 2018. "Semi-Automatic Extraction of Multiword Terms from Domain-Specific Corpora." *The Electronic Library*.
- Perić, M. 2007. *Englesko-Hrvatski Enciklopedijski Rječnik Istraživanja i Proizvodnje Nafta i Plina [English-Croatian Encyclopedic Dictionary of Petroleum Exploration & Production]*. INA Industrija nafte d.d. Sektor korporativnih komunikacija, Zagreb, Hrvatska.
- Petrović, Dejan V, Miloš Tanasijević, Vitomir Milić, Nikola Lilić, Saša Stojadinović, and Igor Svrkota. 2014. "Risk Assessment Model of Mining Equipment Failure Based on Fuzzy Logic." *Expert Systems with Applications* 41 (18): 8157–64.
- Polanyi, Michael. 1964. *Personal Knowledge: Towards a Post-Critical Philosophy*. Harper Torch Books.

- Polguère, Alain. 2014. "From Writing Dictionaries to Weaving Lexical Networks." *International Journal of Lexicography* 27 (4): 396–418.
- Radojičić, Marija, Ivan Obradović, Ranka Stanković, Miloš Utvić, and Sebastijan Kaplar. 2018. "A Mathematical Learning Environment Based on Serbian Language Resources." In *Proceedings of the 7th International Scientific Conference Technics and Informatics in Education*, 248–54. Faculty of Technical Sciences: Čačak, Serbia.
- Radosavljević, Slobodan. 2009. "Model Procene Rizika Procesa Zaštite Na Radu u Pogonu Suva Separacija, 'KOLUBARA PRERADA' Vreoci." *Универзитет у Београду*.
- Rausand, Marvin, and Arnljot Hoyland. 2003. *System Reliability Theory: Models, Statistical Methods, and Applications*. Vol. 396. John Wiley & Sons.
- Rendu, Jean-Michel. 2017. *Risk Management in Evaluating Mineral Deposits*. Society for Mining, Metallurgy & Exploration (SME).
- Roche, Christophe. 2012. "Ontoterminology: How to Unify Terminology and Ontology into a Single Paradigm." In *LREC 2012, Eighth International Conference on Language Resources and Evaluation*, p-2626. European Language Resources Association.
- Rojabanu, M. S., and K. Alagarsamy. 2012. "A Model For The Proactive Risk Management Based On The Text Mining Classification." *International Journal of Engineering Research and Technology* 1.
- Roller, Stephen, Douwe Kiela, and Maximilian Nickel. 2018. "Hearst Patterns Revisited: Automatic Hypernym Detection from Large Text Corpora." *ArXiv Preprint ArXiv:1806.03191*.
- Rubin, Victoria L, Noriko Kando, and Elizabeth D Liddy. 2004. "Certainty Categorization Model." In *AAAI Spring Symposium: Exploring Attitude and Affect in Text: Theories and Applications*, Stanford, CA.
- Ružin, Tatjana Ž. 2015. "Glagoli Uzrokovanja u Engleskom i Srpskom Jeziku." PhD Thesis, Univerzitet u Beogradu-Filološki fakultet.
- Salton, Gerard. 1989. "Automatic Text Processing: The Transformation, Analysis, and Retrieval Of." *Reading: Addison-Wesley* 169.
- Salton, Gerard, and Christopher Buckley. 1988. "Term-Weighting Approaches in Automatic Text Retrieval." *Information Processing & Management* 24 (5): 513–23.
- Šandrih, Branislava, Cvetana Krstev, and Ranka Stanković. 2020. "Two Approaches to Compilation of Bilingual Multi-Word Terminology Lists from Lexical Resources." *Natural Language Engineering* 26 (4 July): 455–79. <https://doi.org/10.1017/S1351324919000615>.
- Schmid, Helmut. 1999. "Improvements in Part-of-Speech Tagging with an Application to German." In *Natural Language Processing Using Very Large Corpora*, edited by Susan Armstrong, Kenneth Church, Pierre Isabelle, Sandra Manzi, Evelyne Tzoukermann, and David Yarowsky, 13–25. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-2390-9_2.
- . 2013. "Probabilistic Part-of-Speech Tagging Using Decision Trees." In *New Methods in Language Processing*, edited by Somers H. L. Jones D. B., 154–64. Routledge: London, United Kingdom.
- Sen, Marcus, and Tim Duffy. 2005. "GeoSciML: Development of a Generic Geoscience Markup Language." *Computers & Geosciences* 31 (9): 1095–1103.

- Shu-hsien, Liao. 2003. "Knowledge Management Technologies and Applications—Literature Review from 1995 to 2002." *Expert Systems with Applications* 25 (2): 155–64.
- Škorić, Mihailo, and Mauro Dragoni. 2019. "Medical Domain Document Classification via Extraction of Taxonomy Concepts from MeSH Ontology." *Infotheca*.
- Slywotzky, Adrian J, and John Drzik. 2005. "Countering the Biggest Risk of All." *Harvard Business Review* 83 (4): 78–88.
- Stanković, Ranka, Cvetana Krstev, Biljana Lazić, and Mihailo Škorić. 2018. "Electronic Dictionaries - from File System to Lemon Based Lexical Database." In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation - W23 6th Workshop on Linked Data in Linguistics: Towards Linguistic Data Science (LDL-2018), LREC 2018*, edited by John P. McCrae, Christian Chiarcos, Thierry Declerck, Jorge Gracia, and Bettina Klimek, 18–23. Paris, France: European Language Resources Association (ELRA).
- Stanković, Ranka, Cvetana Krstev, Biljana Lazić, and Dalibor Vorkapić. 2015. "A Bilingual Digital Library for Academic and Entrepreneurial Knowledge Management." In *Proceeding of 10th International Forum on Knowledge Asset Dynamics-IFKAD*, 1764–77. Bari, Italy.
- Stanković, Ranka, Cvetana Krstev, Ivan Obradović, Biljana Lazić, and Aleksandra Trtovac. 2016. "Rule-Based Automatic Multi-Word Term Extraction and Lemmatization." In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, edited by N. Calzolari and et al, 507–14. Portorož, Slovenia.
- Stanković, Ranka, Cvetana Krstev, Duško Vitas, Nikola Vulović, and Olivera Kitanović. 2016. "Keyword-Based Search on Bilingual Digital Libraries." In *Semantic Keyword-Based Search on Structured Data Sources*, edited by A. Cali, D. Gorgan, and M. Ugarte, 112–23. Cluj-Napoca, Romania: Springer.
- Stanković, Ranka, Ivan Obradović, Olivera Kitanović, and Ljiljana Kolonja. 2012. "Towards a Mining Equipment Ontology." In *Proceedings of the 12th International Conference 'Research and Development in Mechanical Industry' (RaDMI 2012)*, 108–18. Vrnjačka Banja, Serbia: SaTCIP (Scientific and Technical Center for Intellectual Property) Ltd., Serbia.
- Stanković, Ranka, Ivan Obradović, Cvetana Krstev, and Duško Vitas. 2011. "Production of Morphological Dictionaries of Multi-Word Units Using a Multipurpose Tool." In *Proceedings of the Computational Linguistics-Applications Conference, October 2011, Jachranka, Poland*, 77–84.
- Stanković, Ranka, Ivan Obradović, and Miloš Utvić. 2014. "Developing Termbases for Expert Terminology under the TBX Standard." In *Natural Language Processing for Serbian-Resources and Applications. In Proceedings of the Conference "35th Anniversary of Computational Linguistics in Serbia"*, edited by Gordana Pavlović Lažetić, Cvetana Krstev, and Duško Vitas, 12–26. Belgrade, Serbia: University of Belgrade, Faculty of Mathematics.
- Stanković, Ranka, Branislava Šandrih, Cvetana Krstev, Miloš Utvić, and Mihailo Škoric. 2020. "Machine Learning and Deep Neural Network-Based Lemmatization and Morphosyntactic Tagging for Serbian." In *Proceedings of The 12th Language Resources and Evaluation Conference*, 3954–62. Marseille, France: European Language Resources Association.
- Stanković, Ranka, Branislava Šandrih, Rada Stijović, Cvetana Krstev, Duško Vitas, and Aleksandra Marković. 2019. "SASA Dictionary as the Gold Standard for Good Dictionary Examples for Serbian." In *Electronic Lexicography in the 21st Century: Smart*

- Steffen, OKH. 1997. "Planning of Open Pit Mines on a Risk Basis." *Journal of the Southern African Institute of Mining and Metallurgy* 97 (2): 47–56.
- Stellato, Armando, Manuel Fiorelli, Andrea Turbati, Tiziano Lorenzetti, Willem van Gemert, Denis Dechandon, Christine Laaboudi-Spoiden, et al. 2020. "VocBench 3 - A Collaborative Semantic Web Editor for Ontologies, Thesauri and Lexicons." *Semantic Web* 11 (5): 855–81. <https://doi.org/10.3233/SW-200370>.
- Stenetorp, Pontus, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. "Brat: A Web-Based Tool for NLP-Assisted Text Annotation." In *Proceedings of the Demonstrations Session at EACL 2012*. Avignon, France: Association for Computational Linguistics.
- Tarle, Snežana Pejčić, Marijana Petrović, and Nataša Bojković. 2009. *Upravljanje Rizikom Prema Modelu ISO 31000 u Pružanju Poštanskih Usluga*. Beograd.
- Tay, Kai Meng, and Chee Peng Lim. 2008. "On the Use of Fuzzy Inference Techniques in Assessment Models: Part II: Industrial Applications." *Fuzzy Optimization and Decision Making* 7 (3): 283–302.
- Todirascu, Amalia, Laurent Romary, and Dalila Bekhouche. 2002. "Combining Syntax and Ontologies for Information Extraction." In *Terminology and Knowledge Engineering-TKE'02*.
- Tomašević, Aleksandra. 2018. "Развој Модела За Управљање Рударском Пројектном Документацијом." *Универзитет у Београду*.
- Tomašević, Aleksandra, Ranka Stanković, Miloš Utvić, Ivan Obradović, and Božo Kolonja. 2018. "Managing Mining Project Documentation Using Human Language Technology." *The Electronic Library* 36 (6): 993–1009. <https://doi.org/10.1108/EL-11-2017-0239>.
- Tovar, Mireya, Gerardo Flores, José A Reyes-Ortiz, and Meliza Contreras. 2018. "Validation of Semantic Relation of Synonymy in Domain Ontologies Using Lexico-Syntactic Patterns and Acronyms." In *Mexican Conference on Pattern Recognition*, 199–208. Springer.
- Tripathy, Debi Prasad, and Charan Kumar Ala. 2018. "Identification of Safety Hazards in Indian Underground Coal Mines." *Journal of Sustainable Mining* 17 (4): 175–83. <https://doi.org/10.1016/j.jsm.2018.07.005>.
- Tserng, H. Ping, Samuel Y. L. Yin, R. J. Dzen, B. Wou, M. D. Tsai, and W. Y. Chen. 2009. "A Study of Ontology-Based Risk Management Framework of Construction Projects through Project Life Cycle." *Automation in Construction* 18 (7): 994–1008. <https://doi.org/10.1016/j.autcon.2009.05.005>.
- Tufano, Peter. 1996. "Who Manages Risk? An Empirical Examination of Risk Management Practices in the Gold Mining Industry." *The Journal of Finance* 51 (4): 1097–1137.
- Utvić, Miloš. 2014. "The Construction of Reference Corpus of Contemporary Serbian [Izgradnja Referentnog Korpusa Savremenog Srpskog Jezika]." Doctoral dissertation, University of Belgrade.
- Vitas, Duško, and Cvetana Krstev. 1989. "On Representing Morpho-Graphemic Information in a Serbo-Croatian Dictionary." In *Međunarodni Seminar Po Mašnomu Perevodu "IVM i Perevod 89", Tbilisi 1989, Eds. I.I. Ubin, E.E. Lovckiy*. Vsesoznij centr perevodov naučno-tehničkoj literaturi i dokumentacij; Gruzinskij naučno-issledovateljskij institut naučno-tehničkoj informaciji, Moskva.

- . 2012. "Processing of Corpora of Serbian Using Electronic Dictionaries." *Prace Filologiczne* 63: 279–92.
- Vujanić, Milica, Darinka Gortan-Premk, Milorad Dešić, Rajna Dragičević, Miroslav Nikolić, Ljiljana Nogo, Vasa Pavković, et al. 2011. *Rečnik Srpskoga Jezika [Dictionary of the Serbian Language]*. Edited by Miroslav Nikolić. Matica srpska: Novi Sad, Serija.
- Whittle, D. 1997. "The Relationship between Economic Design Objectives and Reserve Estimates, in Resource to Reserve Inputs Seminar." *AusIMM, Melbourne Branch*, 27–29.
- Wynne, Brian. 1988. "Unruly Technology: Practical Rules, Impractical Discourses and Public Understanding." *Social Studies of Science* 18 (1): 147–67.
- Yadav, Usha, G. S. Narula, Neelam Duhan, and Vishal Jain. 2016. "Ontology Engineering and Development Aspects: A Survey." *International Journal of Education and Management Engineering* 6: 9–19.
- Yang, Bo, and Xinfei Wang. 2019. "Construction of Logistics Financial Risk Ontology Model Based on Risk Correlation." In . <https://doi.org/10.2991/smонт-19.2019.14>.
- Ykhlef, Mourad, and Danah Algawiaz. 2014. "A New Strategic Risk Reduction for Risk Management." *International Journal of Computational Intelligence Systems* 7 (6): 1054–63.
- Zhao, W, and JK Liu. 2008. "OWL/SWRL Representation Methodology for EXPRESS-Driven Product Information Model: Part I. Implementation Methodology." *Computers in Industry* 59 (6): 580–89.
- Расулић, Катарина, and Д Кликовац. 2014. *Језик и Сазнање. Хрестоматија Из Когнитивне Лингвистике*. Београд: Филолошки факултет.

Biografija

Olivera M. Kitanović rođena je 13. septembra 1974. godine u Surdulici, Srbija. Završila je IX Beogradsku gimnaziju „Mihajlo Petrović Alas“, prirodno-matematički smer, u Beogradu 1993. godine. Iste godine upisala je Fakultet organizacionih nauka Univerziteta u Beogradu, smer Informacioni sistemi. Studije je završila sa prosečnom ocenom 8,05, a diplomski rad pod nazivom "Realizacija mrežnog analizatora pod operativnim sistemom Linux." je odbranila 2003. godine, sa ocenom 10.

Od 2002. godine zaposlena je na Rudarsko-geološkom fakultetu, najpre na radnom mestu administratora računarskog softvera i mrežnih komunikacionih resursa, a od 2018. godine prelazi na poziciju projektanta informacionih sistema i programa.

Autor je i koautor u preko 20 naučnih radova objavljenih u domaćim i inostranim časopisima, od kojih su 3 sa SCI liste. Naučno zvanje istraživač – saradnik stiče decembra 2019. godine. Uporedo sa naučno-istraživačkim radom, učestvuje u izradi i implementaciji više značajnih domaćih i međunarodnih naučno-istraživačkih projekata vezanih za primenu novih tehnologija u rudarstvu, kao i u izradi nekoliko tehničkih rešenja sa primenom u oblasti rudarstva. Učestvovala je na preko 10 projekata vezanih za projektovanje, razvoj i implementaciju informacionih sistema u oblasti rudarstva i geologije, stičući neophodna iskustva u radu prema najvišim standardima u okviru rudarske industrije.

Član je Društva za jezičke resurse i tehnologije. U okviru COST akcije „Semantic keyword-based search on structured data sources (KEYSTONE)“ je 2016. godine, tokom kratkog studijskog boravka u Rijeci, Hrvatska, za potrebe disertacije radila na ekstrakciji i poravnavanju ključnih reči i ključnih fraza iz jednojezičnih (srpski) i višejezičnih tekstualnih izvora (srpsko-engleski).

Изјава о ауторству

Име и презиме аутора ОЛИВЕРА КИТАНОВИЋ

Број индекса Р712/11

Изјављујем

да је докторска дисертација под насловом

ОНТОЛОШКИ МОДЕЛ УПРАВЉАЊА РИЗИКОМ У РУДАРСТВУ

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 12.04.2021.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора ОЛИВЕРА КИТАНОВИЋ

Број индекса Р712/11

Студијски програм РУДАРСКО ИНЖЕЊЕРСТВО

Наслов рада:

ОНТОЛОШКИ МОДЕЛ УПРАВЉАЊА РИЗИКОМ У РУДАРСТВУ

Ментор Др РАНКА СТАНКОВИЋ, ванредни професор

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањена у **Дигиталном репозиторијуму Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке,

Потпис аутора

У Београду, 12.04.2021.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку "Светозар Марковић" да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

ОНТОЛОШКИ МОДЕЛ УПРАВЉАЊА РИЗИКОМ У РУДАРСТВУ

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)
5. Ауторство – без прерада (CC BY-ND)
6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 12.04.2021.

1. **Ауторство.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.