

Vijeću Prirodno-matematičkog fakulteta

Predmet: Izvještaj Komisije za ocjenu podobnosti teme i kandidata za izradu magistarskog rada „*Namenski jezik za upravljanje ETL procesima u sistemima za skladištenje podataka*“ kandidata Nikole Trajkovića, specijaliste računarskih nauka.

Na osnovu člana 60 Statuta Univerziteta Crne Gore, a u vezi sa članom 24 Pravila studiranja na postdiplomskim studijama, na sjednici Vijeća PMF-a od 9. jula 2019. godine imenovani smo za članove komisije za ocjenu podobnosti teme i kandidata za izradu magistarskog rada „*Namenski jezik za upravljanje ETL procesima u sistemima za skladištenje podataka*“ kandidata Nikola Trajkovića, specijaliste računarskih nauka.

1) Podaci o kandidatu

Nikola Trajković je rođen 20. oktobra 1994. godine u Leskovcu. Nakon što je završio osnovnu školu, 2009. godine, preselio se u Podgoricu gdje je upisao Srednju ekonomsku školu „Mirko Vešović“, smjer ekonomski tehničar. Nosilac je diplome „Luča“. Po završetku srednje škole 2013. godine upisao je Prirodno-matematički fakultet, studijski program Računarske nauke. Zajedno sa kolegama napisao je rad o automatskoj kriptanalizi supstitucionog kriptosistema korišćenjem hill climbing algoritma sa heuristikom.

Prije upisivanja specijalističkih studija, 2016. godine, obavljao je dvomjesečnu praksu u firmi „Poslovna inteligencija“ u Podgorici. Posao je obuhvatao rad sa bazama podataka, konkretno ETL procesima u data warehouse sistemu. Specijalističke studije nastavio je odmah nakon završetka osnovnih studija i iste akademske godine specijalizirao u oblasti vještačke inteligencije na temu „*Paralelizacija konvolucionih neuronskih mreža na primeru klasifikacije slika*“.

U septembru 2017. godine nastavio je rad u firmi „Poslovna inteligencija“ i u isto vrijeme upisao magistarske studije na Prirodno-matematičkom fakultetu. U okviru svog radnog mjesta, kao BI konsultant, održava i nadograđuje data warehouse sisteme za brojne klijente, kako u zemlji tako i u inostranstvu.

2) Obrazloženje teme

a) Nučna oblast

Predložena tema pripada naučnoj oblasti **Namjenski jezici za domen – računarske nauke**.

b) Predmet rada

Data warehouse (DWH) podrazumijeva skup podataka koji su prikupljeni iz različitih izvora, kao što su operativne baze podataka, i sačuvani u posebno skladište podataka. Strukturu DWH sistema čine podaci i mehanizmi za manipulaciju nad podacima. U pomenute mehanizme spadaju procesi ekstrakcije, transformacije i punjenja podataka (ETL – extract, transform, load).

Posebna pažnja posvećena je ETL procesima i procedurama za punjenje i obradu podataka u različitim tipovima i tehnologijama izrade DWH sistema. Realizacija ETL procesa predstavlja glavni izazov prilikom kreiranja DWH sistema. ETL alati su razvijeni da podrže različite izvore podataka, izvrše transformaciju i skladište podatke u dati DWH sistem. Svaki ETL alat ima svoju specifičnu strukturu i funkcije koje su takođe vezane za tehnologiju na kojoj je alat izrađen. Orkestracija poziva ETL paketa koji vrše transformaciju i punjenje, kao i poziva procedura nad bazom podataka usko je vezana za tehnologiju za koju su ETL alati i DWH sistemi izrađeni. Ovdje se pored tehnologije govori i o jeziku koji je karakterističan za datu verziju baze u kojoj se podaci čuvaju. Različiti namjenski jezici za različite verzije DWH sistema, otežavaju proces organizacije punjenja baze podataka. Osim jezika za konkretnu bazu podataka, pozivi za pokretanje ETL procesa zavise od alata, kao i od servera na kojem se alat nalazi.

U ovom radu predlaže se jedno rješenje koje ima za cilj da prevaziđe ili ublaži prethodno opisane probleme koji su izazvani složenošću ETL procesa i heterogenošću pratećih tehnologija. Rješenje se zasniva na kreiranju namjenskog jezika za upravljanje koji će korisniku pružiti odgovarajuće apstrakcije i notacije koje su bliske njegovoj percepciji.

c) Naučni cilj rada

Glavni ciljevi rada su definisanje koncepata i kreiranje namjenskog jezika za upravljanje procesima čiji je zadatak da ublaži probleme izazvane složenošću ETL procesa u sistemima za skladištenje podataka koji se baziraju na različitim arhitekturama. Cilj je da se kompletna

poslovna logika centralizuje i definiše putem koncepata koji su nezavisni od konkretne platforme. Ovakav pristup bi olakšao održavanje sistema i omogućio prenosivost između različitih platformi. Pored toga, kreiranje univerzalnog jezika omogućiće i lakše kreiranje jedinstvenog okruženja za upravljanje procesima punjenja DWH sistema bez obzira na strukturu, jezik ili ETL alat.

Stvaranje namjenskog jezika korisniku će pružiti odgovarajuće apstrakcije i notacije koje su bliske njegovoj percepciji. Novi namjenski jezik će omogućiti transformaciju iz jednog u više konkretnih jezika vezanih za tehnologiju na kojoj se zasniva DWH sistem. Pored toga, biće izgrađena podrška pokretanja ETL procesa bez obzira na tehnologiju za koju je ETL alat vezan. Riječ je o prelasku na viši nivo apstrakcije. Ovakav pristup omogućava efikasnije organizovanje ETL procesa, lakšu manipulaciju nad procedurama, kao i bolje otkrivanje i postavljanje međuzavisnosti nad tabelama u bazi.

d) Naučne metode

Za testiranje hipoteza u radu će se, polazeći od problema, koristiti eksperimentalna metoda, metoda teorijske analize kao i deskriptivna naučno - istraživačka metoda.

- Eksperimentalna metoda podrazumijeva implementaciju jezika pomoću okruženja pod nazivom Xtext. U pozadini okruženja nalazi se programski jezik Java. Prevođenje namjenskog jezika podrazumijeva poznavanje jezika kao što su SQL, PLSQL, Windows i Linux komandi kao i komandi za pokretanje konkretnih ETL paketa i procesa. Eksperimenti se izvode na izdvojenom okruženju sa različitim postavkama i dobijeni rezultati se upoređuju sa pokretanjem originalnih upita i procesa na konkretnoj DWH platformi i u konkretnom ETL alatu respektivno.
- Metodom analize, kao postupkom naučnog istraživanja, doći će se raščlanjivanjem kompleksnih pojmova i procesa do njihovih jednostavnijih sastavnih djelova i elemenata od značaja za sisteme za skladištenje podataka.

- Deskriptivna metoda biće primijenjena u početnoj fazi naučnog istraživanja i opisivanje će biti povezano sa važnijim objašnjenjima osnovnih pojmova DWH sistema, struktura i tehnologije na kojima se DWH sistemi izgrađuju. Metodom deskripcije prikupiće se osnovni podaci o specifičnostima ETL procesa, zatim ETL alata i njihovih funkcionalnosti. Takođe biće opisani principi i načini kreiranja namjenskih jezika.

e) Aktuelnost problematike

Današnje poslovanje nezamislivo je bez postojanja sistema za podršku odlučivanju, tj. DWH sistema. Održavanje jednog takvog sistema podrazumijeva kreiranje okruženja pomoću koga je moguće kontrolisati procese punjenja sistema podacima. Dosadašnja rješenja su se oslanjala na konkretnu tehnologiju na kojoj su alati zasnovani. To podrazumijeva kreiranje aplikacije koja je zavisna od platforme na kojoj je izgrađen DWH sistem, konkretnog ETL alata, kao i jezika baze podataka. Različiti jezici za pojedinačne verzije DWH sistema, otežavaju proces organizacije punjenja baze. Osim jezika baze podataka, pozivi za pokretanje ETL procesa zavise od alata kao i od servera na kojem se alat nalazi. Ovakva rješenja su teška za održavanje, jer minimalna promjena u strukturi zahtjeva znatne promjene u kodu aplikacije za upravljanje ETL procesima.

3) Zaključak

Uvidom u podnesenu dokumentaciju, Komisija je utvrdila da predložena tema kandidata Nikole Trajkovića ima jasno definisane ciljeve i metode istraživanja i očekivane rezultate.

Predlažemo Vijeću Prirodno-matematičkog fakulteta da odobri izradu magistarskog rada kandidata Nikole Trajkovića pod nazivom „*Namenski jezik za upravljanje ETL procesima u sistemima za skladištenje podataka*“.

U Podgorici, 09. septembra 2019. godine

Komisija

Prof. dr Predrag Stanišić, član

Prof. dr Savo Tomović, član

Doc. dr Aleksandar Popović, mentor

