

ISTAKNUTE ZNAČAJKE SQL SERVERA 2019

FEATURED CHARACTERISTICS OF SQL SERVER 2019

Martina Petrovečki, Željko Kovačević

Tehničko veleučilište u Zagrebu, Vrbik 8, Zagreb, Hrvatska

Sažetak

Tvrtka Microsoft je 1989. godine predstavila prvu verziju SQL Servera 1.0., čiji se razvoj nastavlja sve do danas. SQL Server koriste mnogobrojne tvrtke za razvoj velikih i malih sustava čiji je glavni cilj pohrana i obrada podataka te u konačnici analiza podataka. Kroz godine razvoja, tvrtka Microsoft je svakim novim izdanjem unaprijedila SQL Server. Sukladno tome, u SQL Server verziji 2017 pojavljuje se mogućnost korištenja na platformi Linux, što administratorima olakšava rad na različitim razvojnim platformama. Izlaskom SQL Server verzije 2019 unaprijeđena je verzija za Linux čime je sve sličnija verziji za Windows platformu. Kroz ovaj rad opisane su neke od najvažnijih značajki SQL Servera 2019 koje omogućuju jednostavnije upravljanje bazama podataka. Drugim riječima, isti zadaci sada se mogu izvršiti u puno kraćem vremenu te korištenjem šireg spektra mogućnosti. Pristup različitim vanjskim izvorima podataka, obrada velikih podataka te uvođenje UTF-8 kodnog jezika čija upotreba može smanjiti veličinu baze podataka, dodatne su značajke ove verzije SQL Servera.

Ključne riječi: *SQL Server, UTF-8, podatkovni klasteri, uvijek kriptirani podaci, SQL Server Linux*

Abstract

In 1989, Microsoft introduced the first version of SQL Server 1.0. which continued to develop up to nowadays. A great number of companies specialized for the development of large and small systems use SQL Server with its main purpose of storing, processing and eventually data analysis. Throughout the years of its development, Microsoft has upgraded the SQL Server with each version.

Therefore, SQL Server 2017 version presents the possibility of its usage on the Linux platform which enables work on a variety of developing platforms to administrators. SQL Server 2019 is an advanced version for use on Linux, with features similar to the version created for the Windows platform. This work describes the main characteristics of SQL Server 2019 which eases the control over the database. In other words, the same tasks can be completed in a shorter time period, providing a wider range of possibilities. Access to different polybase, processing a large number of data, as well as introduction of UTF-8 code which can reduce the size of the database, are additional features of this SQL Server version.

Keywords: *SQL Server, UTF-8, data clusters, always encrypted, SQL Server Linux*

1. Uvod

1. Introduction

Microsoft SQL Server sustav je za upravljanje relacijskim bazama podataka koji je razvila tvrtka Microsoft. Radi se o softverskom proizvodu kojem je primarna funkcija pohranjivanje i dohvaćanje podataka prema zahtjevima drugih softverskih aplikacija koje se mogu pokretati na jednom ili više računala preko mreže. Microsoft je izdao preko deset različitih verzija Microsoft SQL Servera namijenjenih različitim korisnicima pojedinačnih aplikacija do velikih aplikacija s velikim brojem korisnika.

Posljednje izdana verzija je Microsoft SQL Server 2019 koja donosi inovativne značajke u području sigurnosti i usklađenosti te naprednije mogućnosti podatkovne analitike s podrškom za obradu velikih podataka. SQL Server 2019 može raditi s petabajtom podataka, pri čemu je znatno poboljšana obrada podataka te izvođenje upita

iz različitih vanjskih izvora podataka, kao što su SQL Server, Oracle, Teradata te MongoDB. Izlaskom Microsoft SQL Servera 2017 omogućeno je njegovo korištenje na različitim platformama kao što su Windows, Linux te Docker kontejneri pa je ova značajka unaprijeđena i na novoj verziji SQL Servera 2019.

Kada je riječ o performansama i sigurnosti, SQL Server je jedan od vodećih RDBMS sustava te pripada kategoriji najmanje ranjivih baza podataka. Sve dobre značajke replicirane su na noviju inačicu SQL Servera 2019 čije će značajke biti opisane u ovom članku.

2. Značajke SQL Server 2019

2.1. SQL Server 2019 features

SQL Server 2019 nadogradnja je svih prethodnih inačica SQL Servera te novom inačicom nastavlja evoluciju razvoja novih mogućnosti obrade podataka, bolje podrške za aplikacije kojima je temelj obrada podataka. Novitete u novijoj inačici SQL Servera 2019 možemo opisati kroz pristup podacima od bilo kuda, mogućnost odabira jezika ili platforme te unaprijeđenje SQL Servera za Linux platformu koji je uveden već u verziji SQL Server 2017.

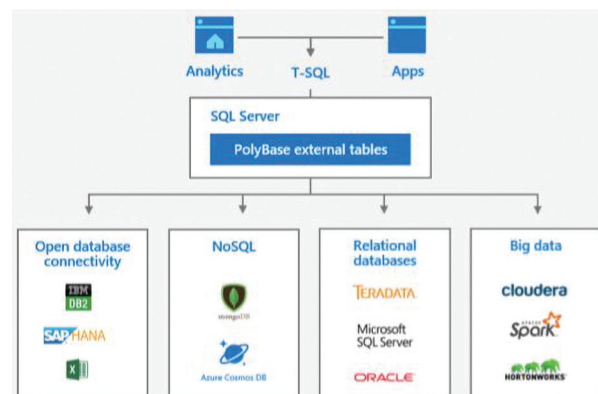
Mogućnost upravljanja podacima od bilo kuda uključuje znatno bolju integraciju te obradu velikih količina podataka, dok izbor operacijskog sustava olakšava rad ukoliko se koriste različita radna okruženja operacijskih sustava. Bilo da se radi o Windows ili Linux operacijskom sustavu, tvrtka Microsoft odlučila je proširiti mogućnost implementacije SQL Servera i time olakšati korištenje SQL Servera na raznim drugim platformama. Linux inačica SQL servera 2017 uključuje sve standardne funkcionalnosti za upravljanje SQL Serverom koje dolaze i s Windows verzijom što je dodatno nadograđeno unutar novije verzije SQL Server 2019. Nadograđene su performanse i sigurnost SQL Servera te dodatne značajke koje su usklađene s uredbom o zaštiti privatnih podataka, tzv. GDPR (*engl. General Data Protection Regulation*) odredbom. U odnosu na SQL Server 2017 dodatno je poboljšana verzija za Linux pri čemu je unaprijeđena podrška za replikaciju baze podataka te izvođenje distribuiranih transakcija.

3. Pristup vanjskim izvorima podataka

3.1. Polybase

Pristup vanjskim izvorima podataka omogućen je prvi puta u verziji SQL Servera 2016 gdje se SQL Server instanci omogućuje izvršavanje Transact-SQL upita koji učitavaju podatke iz vanjskih izvora podataka kao što su Hadoop (*engl. Hadoop Distributed File System - HDFS*) te Azure Blob Storage definirajući vanjsku tablicu koja predstavlja podatke iz nekog od spomenutih vanjskih servisa. To znači da korisnici mogu izvršiti Transact-SQL upite koji se odnose na vanjsku tablicu kao da je normalna tablica SQL poslužitelja, a kada se upit izvrši, podaci iz vanjske tablice se dohvaćaju i prikazuju korisniku.

U odnosu na SQL Server 2016 razvojem SQL Servera 2019 proširena je mogućnost pristupa podacima i drugim izvorima podataka što uključuje sam SQL Server, Oracle, Teradata, MongoDB ili bilo koji drugi izvor podataka koji koristi mogućnost ODBC (*engl. Open Database Connectivity*) upravljački program. [1]



Slika 1 Vanjski izvori podataka [2]

Figure 1 Polybase

Nakon što se kreiraju tablice unutar SQL Servera, za upravljanje podacima može se koristiti aktivni direktorij (*engl. Active Directory*) što znatno olakšava pristup i upravljanjem podacima, pravima korisnika te grupama korisnika. SQL Server 2019 omogućuje čitanje podatka iz samog izvora podataka koristeći instancu SQL Servera na podatkovnim čvorovima za filtriranje i agregiranje podataka. Za povećanje performansi SQL Servera mogu se kreirati različite grupe SQL instanci koje se mogu dodati u skupinu vanjskih izvora (*engl. Polybase*) koje će biti pod glavnom skupinom.

Na taj način upiti se predaju glavnom čvoru koji zatim predaje radno opterećenje na računalne instance skupine vanjskih izvora, što omogućuje paralelnu obradu podataka s vanjskih izvora podataka.

Dodavanjem novih podataka te novih izvora, skupina vanjskih izvora može se povećati dodavanjem novih instanci te bez obzira na radno opterećenje i virtualizaciju podataka mogu se dodavati i uklanjati instance unutar nekoliko trenutaka. S poboljšanom verzijom vanjskih izvora podataka mogu se osigurati različiti izvori podataka za izradu izvješća te analizu unutar SQL Servera.

4. Podatkovni klasteri za SQL Server

4. SQL Server data clusters

Veliki podatkovni klasteri (*engl. cluster*) koriste poboljšanu verziju vanjskih izvora podataka kako bi se unaprijedila virtualizacija podataka između SQL Servera te drugih izvora podataka, čime bi se omogućila brža i sigurnija dvosmjerna integracija s velikim sustavima kao što su Hadoop i Apache Spark. SQL Server 2019 nudi tri različite funkcionalnosti:

- Virtualizacija podataka – kombiniranje podataka iz mnogih izvora podataka bez njihovog repliciranja pri čemu se smanjuje predmemoriranje podataka te time povećava učinkovitost.

- Upravljeni SQL Server, Spark i centralizirano središte pohranjenih strukturiranih i nestrukturiranih podataka (*engl. data lake*) – mogućnost pohrane velike količine strukturiranih i nestrukturiranih podataka kojima je moguće pristupiti preko SQL Servera ili Sparka.

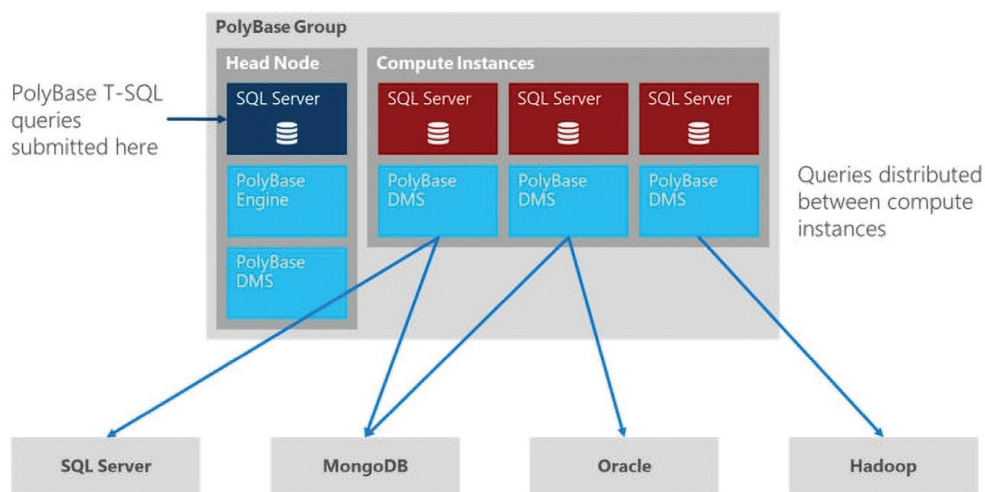
- AI platforma – jednostavno unošenje podataka iz različitih izvora unutar modela što nudi mogućnost pripremanja podataka, a zatim pohrana te obrada modela unutar jednog sustava.

Veliki podatkovni klasteri za SQL Server poprimaju oblik računalnog klastera za obradu velikih podataka gdje se podaci mogu paralelno obrađivati u različitim instancama SQL Servera, bilo da se obrađuju s računala ili nekih drugih spremnika unutar klastera, a izrađeni su za Linux platformu, Docker kontejnere u suradnji sa velikim spremnicima podataka kao što su Kubernetes, HDFS i Spark. [4]

5. Poboljšanje SQL Server baze podataka

5. SQL Server database improvement

Razvoj SQL Servera 2019 nije ograničen samo na korištenje vanjskih izvora podataka i rad s velikim klasterima podataka, već su poboljšane i same performanse servera, dostupnost na drugim platformama, sigurnost i usklađenost, podrška za UTF-8 kodne znakove te znatno poboljšanje za Linux podršku.



Slika 2 Prikaz grupe vanjskih izvora podataka [3]

Figure 2 Display of the polybase group

U ovoj inačici SQL Servera posebno je poboljšana podrška za uređaje koji koriste fizičku memoriju (PMEM). Svaka datoteka SQL Servera koja se nalazi na fizičkoj memoriji uređaja, SQL Server pristupa izravno uređaju zaobilazeći memoriju operacijskog sustava. Proces pokretanja samog SQL Server poslužitelja znatno je brži za baze podataka koje koriste tablice u memoriji za hibridnu transakcijsku i analitičku obradu podataka. Mogućnost ponovnog korištenja DLL operacija kao i podrška za stvaranje online indeksa dodatno je poboljšanje nove verzije SQL Servera 2019.

SQL Server 2019 razvija podršku za više dostupnih scenarija te platformi pri čemu uključuje mogućnost konfiguracije SQL Servera koji se izvodi u spremnicima. Povećan je broj sinkronih replika s tri na pet sinkronih replika i to na jednu primarnu i četiri sekundarne replike s automatskim prebacivanjem u slučaju kvara na jednoj od replika.

5.1. Sigurnost i usklađenost

5.1. *Security and compatibility*

SQL Server 2019 uvodi nekoliko poboljšanja koja olakšavaju administratorima baza podataka i službenicima za zaštitu podataka osiguranje osjetljivih podataka te poštivanje politike i zakona o zaštiti podataka.

5.1.1. Pretraživanje i klasifikacija podataka u SQL Serveru 2019

5.1.1. *SQL Server 2019 data search and classification*

Ova mogućnost dodana je u SQL Serveru Management Studio (SSMS) verziji 17.5 što omogućuje pretraživanje i klasifikaciju stupaca u bazi podataka koja sadrži osjetljive podatke. Stupci se mogu klasificirati prema vrsti informacija koje sadrže imena, adrese, lozinke te po stupnju osjetljivosti podataka u stupcu – javne (*engl. public*), opće (*engl. general*) i povjerljive (*engl. confidential*). Izvještaje je vrlo jednostavno generirati prema klasifikaciji podataka koja je pripremljena kako bi bile zadovoljene zakonske i regulatorne zahtjeve poput zaštite osobnih podataka.

SSMS uključuje čarobnjaka za pretraživanje i klasifikaciju podataka u SQL bazi podataka koji će pokušati identificirati stupce u bazi podataka koji mogu sadržavati osjetljive informacije. SQL pretraživanje podataka i klasifikacija koriste temeljni mehanizam *SQL Server Extended Properties* te je tako kompatibilan sa sustavom SQL Server 2008 te svim novijim inačicama.

5.1.2. Uvijek kriptirani podaci

5.1.2. *Always encrypted data*

Opcija koja je uvedena u verziji SQL Server 2016 enkriptijska je tehnologija koja štiti povjerljive i osjetljive podatke od različitih zlonamjernih programa, neovlaštenih korisnika SQL Servera, uključujući administratore baze podataka te druge neovlaštene pristupe. Uvođenjem ove mogućnosti povjerljivi podaci nisu prikazani kao čisti tekst pa ako administrator baze podataka nema pristup ključevima kojima bi mogao dešifrirati podatke podaci neće biti vidljivi ili sam upit neće biti izvršen, ovisno o svojstvima baze podataka. [5] U SQL Serveru 2019, opcija uvijek kriptiranih podataka postiže iste pogodnosti šifriranja podataka kao i u prethodnim verzijama, samo što se šifriranje vrši na strani klijenta te nikada ne dopušta da se podaci ili odgovarajući kriptografski ključevi prikažu u čistom tekstu unutar SQL Servera. Kao rezultat toga, funkcionalnost kriptiranih stupaca baze podataka strogo je ograničena. Jedina operacija koju *SQL Server Engine* može izvesti na kriptiranim podacima je usporedba jednakosti, dok sve ostale operacije, kao što su kriptografske operacije tipa podudaranje uzoraka, nisu podržane unutar baze podataka i moraju se izvršavati u aplikacijskom sloju.

5.1.3. Procjena ranjivosti

5.1.3. *Vulnerability assessment*

SQL Server 2019 omogućuje procjenu ranjivosti baze podataka pri čemu prati usklađenost instanci SQL Servera i instanci Azure SQL baze podataka. Za lokalne instance SQL Servera procjenama ranjivosti može se pristupiti preko SSMS alata, dok je za instance Azure SQL baze podataka ona dostupna preko Azure portala.

Procjena ranjivosti radi na principu skeniranja na način da uspoređuje konfiguraciju baze podataka sa sigurnosnom praksom koju održava Microsoft te na kraju skeniranja alat generira izvješće o stanju sigurnosti i pojedinosti o pronađenim ranjivostima te koracima sanacije. Na kraju skeniranja potrebno je odabrati primjenu generiranih preporuka na bazu podataka jer se ne primjenjuju automatikom. Procjena ranjivosti pruža jednostavan način proaktivnog praćenja i poboljšanja sigurnosti baze podataka te bolju usklađenost sa GDPR propisima.

5.1.4. Upravljanje certifikatima

5.1.4. *Certificate management*

SSL/TLS certifikati se primjenjuju za siguran pristup SQL Serveru. U ranijim verzijama SQL Servera bilo je potrebno uložiti puno veće napore za održavanje infrastrukture certifikata unutar SQL Servera i to često kroz razvoj različitih SQL skripti ili samostalnim izvršavanjem naredbi. Razvojem SQL Servera, upravljanje certifikatima integrirano je unutar *SQL Server Configuration Managera*, što je pojednostavilo upravljanje zadacima kao što su pregled i provjera valjanosti certifikata unutar instance SQL Servera te izrada izvješća o certifikatima koji su pred istekom.

5.1.5. Podrška za UTF-8

5.1.5. *UTF-8 support*

Izlaskom SQL Server 2019 inačice omogućeno je i korištenje UTF-8 kodnog jezika koji pripada široko korištenom kodnom jeziku. Podrška za UTF-8 kodni jezik od velikog je značaja tvrtkama koje širesvoje poslovanje na globalnu razinu te razvijaju višejezične aplikacije koje koriste SQL bazu podataka za unos i obradu podataka. UTF-8 je za sada omogućen u postojećim podatkovnim tipovima kao što su CHAR i VARCHAR, dok se podaci tipa STRING automatski pretvaraju u UTF-8 kodni jezik kod promjene tipa kodnog jezika, tako primjerice promjena iz LATIN1_GENERAL_100_CI_AS_SC vrši promjenu u LATIN1_GENERAL_100_CI_AS_SC_UTF8. Korištenjem kodnog jezika UTF-8 mogu se postići značajne uštede u pohrani podataka, ovisno o skupu podataka koji se koriste.

Promjenom postojećeg tipa podataka unutar stupca koji koristi NCHAR (10) te kodni tip UTF-16 pretvorbom u UTF-8 može smanjiti 50% zauzeća memorije iz razloga što NCHAR (10) zahtjeva 22 bajta za pohranu, dok CHAR (10) zahtjeva 12 bajtova za isti *Unicode* niz. [6]

5.1.6. SQL Server na Linux platformi

5.1.6. *SQL Server on Linux platform*

Microsoft je razvojem SQL Servera 2017 predstavio značajan korak u razvoju SQL servera koji daje mogućnost instalacije na različitim platformama. SQL Server 2017 razvijen je za korištenje u Windows i Linux operacijskom sustavu što uvelike olakšava rad administratorima koji koriste različite razvojne platforme. [7] Ovu važnu značajku Microsoft je unaprijedio razvojem SQL Servera 2019 te time približio razvoj SQL Servera verziji koja se koristi na Windows operativnom sustavu. Značajke koje su unaprijeđene za Linux platformu su:

- Implementacija replikacija
- Lakša integracija u aktivni direktorij
- Distribuirane transakcije

6. Dostupni alati unutar SQL Servera 2019

6. *Tools available within SQL Server 2019*

SQL Server 2019 dolazi s novim poboljšanjima te novim besplatnim alatima kao što je Microsoft Ignite 2018 te Azure Data Studio 7. Azure Data Studio 7 platforma je podržana na operacijskim sustavima Windows, Linux i MacOS kao desktop aplikacija koja služi za administriranje te izvršavanje upita unutar SQL Server instanci na bilo kojoj platformi.

Azure Data Studio omogućuje izvođenje upita, provođenje administrativnih zadataka te stvara prilagođene nadzorne ploče za praćenje stanja instanci SQL Servera, a temelji se na istoj platformi kao i Visual Studio Code što uključuje Git integraciju.

SQL Server Management Studio besplatan je alat dostupan na Internetu koji nudi sveobuhvatne alate za dizajn i upravljanje bazama podataka, analizama, integracijom te izradu izvješća. SSMS je izrađen u alatu Visual Studio, što omogućuje jednostavnije proširenje različitih značajki. SSMS alatom omogućena je jednostavna administracija baze podataka, konfiguracija te održavanje baze podataka te isto tako jednostavno upravljanje SQL Server Agentom. [5]

Visual Studio Code besplatan je alat koji se može koristiti na više platformi, a s ekstenzijom mysql omogućuje izvršavanje Transact-SQL upita na instancama SQL Servera.

Osim grafičkih alata, upravljanje SQL Serverom može se vršiti i preko naredbenog retka (*engl. shell tools*). Primjeri naredbi i njihov opis u navedeni su u nastavku:

- sqlcmd – koristi se za izvršavanje upita iz naredbenog retka
- bcp – koristi se za uvoz i izvoz podataka unutar SQL Server instance
- sqlpackage – služi za skriptiranje baze podataka kako bi se pojednostavila migracija baze podataka

7. Zaključak

7. Conclusion

Razvoj novih inačica SQL Servera te razvoj novih tehnologija čine ovaj alat još primamljivijim alatom za korištenje unutar velikih kompanija. Programeri, analitičari i stručnjaci za obradu podataka te administratori baze podataka uvelike mogu povećati produktivnost sa svim novim značajkama koje nudi SQL Server. Rad s velikim podatkovnim klasterima unutar SQL Servera 2019 znatno olakšavaju rad s velikim količinama podataka te rad sa visokovrijednim relacijskim bazama podataka na jednoj vrlo stabilnoj platformi. S obzirom da novije verzije SQL servera nude mogućnost korištenja servera na različitim platformama, mnogim kompanijama upravo ovaj element daje veću fleksibilnost odabira alata za dizajniranje i upravljanje bazom podataka.

8. REFERENCE

8. REFERENCES

- [1] S. Minukhin, V. Fedko i Y. Gnusov, »Enhancing the performance of distributed big data processing systems using Hadoop and Polybase,« Eastern-European Journal of Enterprise Technologies, pp. 16-28, 7 27 2018.
- [2] T. Ford, »Polybase Expansion, Big Clusters Are Key Features in the New SQL Server 2019,« 26 09 2018. [Mrežno]. Available: <https://www.itprotoday.com/sql-server/polybase-expansion-big-clusters-are-key-features-new-sql-server-2019>.
- [3] »Microsoft SQL Server 2019,« 09 2018. [Mrežno]. Available: <https://info.microsoft.com/rs/157-GQE-382/images/EN-US-CNTNT-white-paper-DBMod-Microsoft-SQL-Server-2019-Technical-white-paper.pdf>. [Pokušaj pristupa 8 1 2019].
- [4] M. Shmel i V. Mužny, »Hands on Data Science with SQL Server 2017,« Birmingham, Packt Publishing Ltd., 2018, pp. 450-452.
- [5] D. Sarka, M. Radivojević i W. Durkin, »SQL Server 2017 Developer's Guide: A professional guide to designing and developing enterprise database applications,« Birmingham, Packt Publishing Ltd., 2018, pp. 11; 18-19.
- [6] 3 2 2019. [Mrežno]. Available: <https://cloudblogs.microsoft.com/sqlserver/2018/12/18/introducing-utf-8-support-in-sql-server-2019-preview/>.
- [7] B. Ward, »Pro SQL Server on Linux: Including Container - Based Deployment with Docker and Kubernetes,« North Richland Hills, Apress Media LLC, 2018, pp. 1-6.

AUTORI · AUTHORS**Martina Petrovečki**

Vanjska je suradnica na Tehničkom veleučilištu u Zagrebu gdje sudjeluje u nastavi iz objektno orijentiranih programskih jezika. Završila je specijalistički studij na

Tehničkom veleučilištu u Zagrebu 2016. godine, a izbor u zvanje asistenta stječe 2018. godine. U slobodno vrijeme bavi se istraživanjem iz područja baza podataka.

Korespondencija

mpetrove1@tvz.hr

**Željko Kovačević**

Predavač je na Tehničkom veleučilištu u Zagrebu gdje sudjeluje u nastavi iz kolegija orijentiranih prema učenju programskih jezika i baza podataka. Diplomirao je na

Tehničkom veleučilištu u Zagrebu 2010.g, a zvanje predavača stječe 2016.g. Autor je tri knjige i desetak stručnih i znanstvenih radova iz područja programskih jezika i baza podataka, a 2015.g. dobiva posebno priznanje MVP (Most Valuable Professional) tvrtke Embarcadero za rad u RAD Studio alatima C++ Builder i Delphi. Trenutno pohađa doktorski studij Računarstva i Informatike u Mariboru, gdje radi na istraživanju iz područja domenski-specifičnih programskih jezika.

Korespondencija

zkovacev1@tvz.hr