

**KONCEPT JEZERA
PODATAKA I
MOGUĆNOST PRIMENE
U JAVNOM SEKTORU U
REPUBLICI SRBIJI**

decembar 2023.

UREDNIKA

Ana Milovanović, NALED

AUTORKA

Marija Suzić, NALED

SARADNICA

Milica Anđelković Đoković, NALED

OBAVEŠTENJE O AUTORSKOM PRAVU

© 2023 NALED

Makedonska 30/VII, 11000 Beograd, Srbija

www.naled.rs

Ovaj dokument pripremio je stručni tim NALED-a u sklopu projekta *Javne nabavke i dobro upravljanje za veću konkurentnost* koji sprovodi NALED, a uz podršku Švedske agencije za međunarodni razvoj i saradnju (Sida). Korišćenje, kopiranje i distribucija sadržaja ovog dokumenta dozvoljena je isključivo u neprofitne svrhe i uz odgovarajuće naznačenje imena, odnosno autorskih prava NALED-a. Učinjeni su svi napori kako bi se osigurala pouzdanost, tačnost i ažurnost informacija iznetih u ovom dokumentu.

Sadržaj:

Sadržaj:.....	3
Kratak pregled	4
Uvod.....	8
Definisanje koncepta jezera podataka	9
Baze podataka	9
Magacin podataka.....	11
Jezera podataka.....	12
Ključne razlike između baza, magacina i jezera podataka	14
Regulatorni okvir za uspostavljanje jezera podataka	16
Modeli upravljanja i korišćenja jezera podataka.....	18
Modeli upravljanja jezerom podataka	18
Modeli korišćenja jezera podataka	19
Mogućnosti koje proističu iz upotrebe jezera podataka	20
Analiza prakse primene koncepta jezera podataka.....	23
Naučene lekcije na osnovu primera iz prakse	24
Zainteresovane strane i potencijalne mogućnosti primene koncepta jezera podataka.....	25
Jezera podataka u sektoru poljoprivrede	26
Analiza spremnosti i preporuke za implementaciju koncepta jezera podataka u Republici Srbiji	28
Primer upotrebe 1 - Jezero podataka u sektoru zdravstva.....	31
Primer upotrebe 2 – Jezero podataka u Kragujevcu	36
Zaključak.....	40
Rečnik.....	42

Kratak pregled

Skladišta podataka, kao centralizovani sistemi za upravljanje velikim količinama podataka, ključna su za informisano donošenje odluka u međunarodnim organizacijama, preduzećima i državnim institucijama. Postoje različiti tipovi skladišta podataka, uključujući baze podataka, magacine podataka i jezera podataka, a izbor odgovarajuće opcije zavisi od specifičnih potreba i ciljeva organizacije (privatnog preduzeća ili organizacije u javnom sektoru).

Baze podataka su ključne u svakodnevnom radu organizacija, omogućavajući efikasno upravljanje transakcijama poput zakazivanja termina. Baze podataka primarno skladište strukturisane podatke i predstavljaju dobro rešenje za skladištenje i brz pristup podacima u različitim oblastima, poput e-trgovine, bankarstva, sporta i javnog sektora. Iako su baze podataka široko korišćene za skladištenje i upravljanje podacima, one imaju određena tehnološka ograničenja, poput kompleksnosti upravljanja velikim podacima, slabljenja performansi sa rastom količine skladištenih podataka i ažuriranja podataka. Iz navedenih razloga, magacini ili jezera podataka mogu biti bolja opcija za skladištenje velikih podataka.

Magacini podataka su specijalizovani za skladištenje uglavnom strukturisanih podataka koristeći procese za ekstrakciju, transformaciju i učitavanje podataka iz raznih izvora, s izazovom vremenske diskrepancije usled pokretanja procesa u određenom vremenskom trenutku. Magacini podataka su idealni za čuvanje obimnih strukturisanih istorijskih podataka namenjenih detaljnim analizama.

Jezera podataka omogućavaju skladištenje strukturisanih, polustrukturisanih i nestrukturisanih podataka u njihovom izvornom formatu, obezbeđujući vremenske uštede i veću fleksibilnost organizacijama. Ona su ključni deo digitalne transformacije i omogućavaju organizacijama da ostvare konkurentne prednosti u eri informacija. Jezera podataka pružaju troškovno efikasno rešenje za skladištenje podataka, zahvaljujući mogućnosti korišćenja postojeće infrastrukture i odvajanja skladištenja od obrade podataka, a dodatno pružaju ekonomsku efikasnost omogućavajući slojevito skladištenje podataka po klasama važnosti. Izazovi u upravljanju jezerima podataka, poput očuvanja kvaliteta i sprečavanja stvaranja "močvare podataka", zahtevaju pažljivo planiranje i efikasno upravljanje podacima. Pri odabiru između baza podataka, magacina i jezera podataka, ključna razmatranja uključuju tip podataka, potrebu za unapred definisanim shemama za olakšavanje analiza i trenutno mesto skladištenja podataka.

Ukoliko se određena organizacija javnog ili privatnog sektora opredeli da unapredi poslovanje i upravljanje podacima izgradnjom jezera podataka, bitno je da zna da donošenje posebnih propisa nije neophodno za gradnju i upravljanje jezerom podataka, ali da mora poštovati propise o zaštiti ličnih podataka. U slučaju javnog sektora, dodatna pažnja mora biti usmerena i na odgovarajuće propise koji u okviru sektora regulišu pitanje upravljanja podacima. U zavisnosti od delokruga javnog sektora u kome se država opredeli za gradnju jezera podataka, može biti potrebno donošenje novih ili izmena postojećih propisa koji se tiču upravljanja specifičnim tipom podataka, a koji će biti skladišteni u jezeru podataka. Takođe, bitno je da rukovaoci i obrađivači podataka imenuju lice koje će biti zaduženo za zaštitu podataka o ličnosti. Svaki vlasnik jezera podataka predstavlja rukovaoca podacima, a može biti i obrađivač podataka.

Koristi koje države imaju od centralizovanog skladištenja raznovrsnih podataka ogledaju se u praćenju porekla i upotrebe podataka, što je ključno za koordinaciju rada različitih organizacija u javnom sektoru i usklađivanje sa strateškim ciljevima. Efikasna akumulacija i konsolidacija podataka iz različitih izvora olakšavaju identifikaciju neusklađenosti i grešaka u podacima, a time se unapređuje tačnost analiza i doprinosi integritetu i pouzdanosti sistema koji generišu i koriste podatke. Podaci skladišteni u jezeru

podataka se mogu dopuniti metapodacima čime se omogućava dublji uvid u podatke i čine jezera podataka superiornijim u odnosu na tradicionalna skladišta podataka. Fleksibilnost jezera podataka omogućava organizacijama lako prilagođavanje i reorganizuju podataka za različite svrhe i korisnike.

Uspešna izgradnja jezera podataka zahteva jasno definisanje problema koji se rešava i plan upotrebe skladištenih podataka kako bi se izbeglo pretvaranje jezera u „močvaru podataka“ i osigurala vrednost ulaganja u odnosu na troškove održavanja. Formiranje multidisciplinarnog tima stručnjaka i uspostavljanje kulture saradnje, uz obuku zaposlenih za korišćenje i analizu podataka, ključni su preduslovi za uspešnu izgradnju i upravljanje jezerom podataka.

Zemlje Evropske unije usvajaju koncept jezera podataka sa ciljem unapređenja poslovanja i pružanja boljih usluga građanima. Neke od zemalja koje su implementirale koncept jezera podataka u javnom sektoru su: Švajcarska, Finska, Francuska, Italija i Estonija. Analizom primera implementacije jezera podataka u navedenim zemljama može se zaključiti da modernizacija sistema za prikupljanje i obradu podataka postaje ključna za osiguranje sigurnog skladištenja podataka, smanjenje operativnih troškova, povećanje efikasnosti i unapređenje dostupnosti, deljenja i analize podataka.

Lanac vrednosti podataka ilustruje proces stvaranja nove ekonomske i društvene vrednosti kroz otvaranje i korišćenje podataka, a sa ciljem unapređenja javnih usluga, stvaranja radnih mesta i podsticanja saradnje između javnog i privatnog sektora. U ovom modelu učestvuju različiti akteri: snabdevači, sakupljači podataka i programeri. Javni sektor igra ključnu ulogu u fazi kreiranja i validacije podataka, kao i na kraju, kao korisnik rezultata analiza. Privatni sektor je važan za proces sakupljanja podataka, njihovu analizu i razvoj usluga zasnovanih na tim analizama. Koristi od podataka nisu ograničene samo na ekonomski rast i efikasnost; one se takođe, odnose na društvene koristi kao što su transparentnost, odgovornost i inovacije koje doprinose opštem dobru. Otvaranjem i korišćenjem podataka, kako privatni tako i javni sektor mogu zajedno raditi na stvaranju održivog i inkluzivnog ekonomskog razvoja koji odgovara potrebama i izazovima savremenog društva.

Izgradnja jezera podataka u javnom sektoru bi trebalo da omogući visokim državnim funkcionerima, predstavnicima lokalnih vlasti, i celim sektorima poput zdravstva i obrazovanja, da vrše složene analize i donose strateške odluke. Osim toga, kontrolori ispravnosti i transparentnosti mogu koristiti jezera podataka za nadzor nad korišćenjem budžetskih sredstava i etičku upotrebu podataka, dok IT stručnjaci igraju direktnu ulogu u izgradnji i održavanju jezera podataka. Građani i privreda ostvaruju indirektnu korist kroz poboljšane usluge zahvaljujući agilnosti javnog sektora koju jezera podataka treba da omoguće, dok privatni sektor i akademska zajednica kroz saradnju sa javnim sektorom, koriste podatke za inovacije i unapređenje usluga.

Uvođenje jezera podataka u javni sektor Srbije bi trebalo da omogući jednostavan pristup i transformaciju istorijskih podataka za analitičke svrhe, nakon što se podaci uvezu u jezero iz starih sistema. Za postizanje maksimizacije vrednosti podataka javnog sektora u Srbiji, a pre implementacije koncepta jezera podataka, biće potrebno sprovesti sistematizaciju i katalogizaciju proizvedenih podataka. Trenutno u Srbiji nedostatak standarda za podatke i neusklađenost informacionih sistema u različitim delovima javnog sektora otežava njihovu međusobnu povezanost i analizu. Izazovi implementacije jezera podataka uključuju raznovrsnost formata podataka i nedostatak razumevanja značaja i mogućnosti elektronskih podataka i analitike. Definisane jasne uloge svake institucije uključene u upotrebu jezera podataka kroz politike upravljanja podacima, ključno je za razvoj jezera podataka u Srbiji. Nedostatak stručnjaka iz IKT oblasti i niske plate u javnom sektoru su još jedan izazov koji je potrebno prevazići pre početka implementacije koncepta jezera podataka. Postojanje Državnog centra za upravljanje i čuvanje podataka u Kragujevcu olakšava tehnički deo implementacije jezera podataka, ali su ipak potrebna su dodatna ulaganja u hardverska rešenja za skladište podataka.

Međutim, na osnovu dva predloga upotrebe jezera podataka u javnom sektoru u Srbiji može se zaključiti da su brojni benefiti koje bi celokupno društvo i država ostvarili od uvođenja jezera podataka.

Uvod

Podaci postaju sve vredniji resurs za donošenje odluka, kako u međunarodnim organizacijama, tako i u preduzećima različitih veličina i državnim institucijama. Stoga, skladišta podataka, koja su centralizovani sistemi za čuvanje i upravljanje velikim količinama podataka, igraju ključnu ulogu u savremenom dobu. Ona omogućavaju jednostavan pristup, obradu i analizu podataka za informisano donošenje odluka. Postoje različiti tipovi skladišta podataka, uključujući baze podataka, magacine podataka i jezera podataka, a izbor odgovarajuće opcije zavisi od specifičnih potreba i ciljeva organizacije.

Kada su u pitanju podaci proizvedeni i generisani u javnom sektoru Srbije, uprkos razvoju brojnih elektronskih usluga javne uprave, pitanje integracije i povezivanja podataka nije rešeno, što ograničava njihovu punu vrednost. Podaci se često skladište u različita skladišta bez efektivnog povezivanja, što utiče na efikasnost njihove analize i korišćenja u procesu donošenja odluka. U cilju poboljšanja ove situacije, jezera podataka se javljaju kao obećavajuće tehnološko rešenje. Ona omogućavaju centralizovano skladištenje i bolje upravljanje raznovrsnim podacima, uključujući i njihovo efikasno povezivanje. Prednosti upotrebe jezera podataka uključuju ne samo integraciju i lakši pristup podacima, već i mogućnost detaljnijih analiza, što može značajno doprineti informisanijem donošenju odluka. Međutim, važno je biti svestan i potencijalnih izazova, kao što su potreba za naprednom infrastrukturom, obukom zaposlenih i adekvatnim upravljanjem podacima.

Analiza Koncept jezera podataka i mogućnosti za primenu u Republici Srbiji ima za cilj da predstavi jezera podataka kao tehnološko rešenje za skladištenje i upravljanje podacima, sa svim prednostima i nedostacima koje ona nude. Analiza je podeljena u sedam glavnih poglavlja. Prvo poglavlje posvećeno je predstavljanju koncepta skladišta podataka, u kome su predstavljena tri tipa skladišta podataka: baze podataka, magacini podataka i jezera podataka. U drugom poglavlju je predstavljen regulatorni okvir za uvođenje jezera podataka. Naredno, treće poglavlje, posvećeno je modelima upravljanja i korišćenja jezera podataka. Četvrto poglavlje je koncentrisano na predstavljanje mogućnosti koje proizilaze iz koncepta jezera podataka za korisnike. Potom je analizirana praksa primene koncepta jezera podataka u javnom sektoru u zemljama Evropske unije i predstavljene lekcije naučene iz ove prakse. Šesto poglavlje je posvećeno predstavljanju zainteresovanih strana u slučaju primene jezera podataka u javnom sektoru. U sedmom poglavlju je analizirana pozicija Srbije u pogledu mogućnosti implementacije koncepta jezera podataka, uz predstavljanje dva detaljna primera mogućnosti implementacije jezera podataka u javnom sektoru – u sektoru zdravstva i u gradu Kragujevcu.

Definisanje koncepta jezera podataka

Skladišta podataka se koriste za čuvanje i upravljanje velikim količinama podataka iz različitih izvora, omogućavajući lak pristup, obradu i analizu podataka za informisano donošenje odluka. U eri naprednih informacionih tehnologija, podaci postaju ključni resurs za donošenje odluka na svim nivoima - od međunarodnih organizacija, velikih i malih preduzeća do državnih institucija. Države ne samo da su veliki proizvođači podataka, već su i odgovorne za njihovo upravljanje, skladištenje i primenu u različite svrhe. Skladišta podataka predstavljaju centralizovane sisteme za čuvanje i upravljanje velikim količinama podataka, koji se koriste kako u javnom tako i u privatnom sektoru. Osnovna svrha skladišta podataka jeste da obezbede jednostavan, lak i brz pristup podacima. U skladištima se mogu naći podaci prikupljeni iz različitih izvora, kao što su različite aplikacije, poput aplikacija za prijavu na određeni servis, bankarskih aplikacija, aplikacija za korišćenje javnih usluga, aplikacija na pametnim uređajima poput onih za brojanje načinjenih koraka, podaci prikupljeni posetama različitim sajtovima, servisima za slušanje muzike ili gledanje video sadržaja i slično. Osim čuvanja, ona nude i mogućnost obrade i analize podataka, što je važno za izradu izveštaja i donošenje odluka na osnovu podataka.

Različite vrste skladišta podataka nude različite prednosti i ograničenja, pri čemu izbor zavisi od potreba i ciljeva organizacije. Postoje različite vrste skladišta podataka, ali većina ih se može grupisati u tri tipa: baze podataka, magacini podataka i jezera podataka. Odabir odgovarajuće opcije za skladištenje podataka zavisi od specifičnih potreba i ciljeva organizacije (privatnog preduzeća ili organizacije u javnom sektoru). Svaka od navedenih opcija ima svoje prednosti i ograničenja, koja će biti predstavljena u nastavku, uz poseban osvrt na jezera podataka.

U Srbiji, iako su razvijene mnoge elektronske usluge javne uprave, usled nedostatka integracije i povezivanja, podataka, ne ostvaruje se njihova puna vrednost. Država i lokalne samouprave su prethodnih godina aktivno radile na razvoju elektronske uprave kroz razvoj brojnih aplikacija i servisa za pružanje elektronskih usluga javne uprave. Međutim, podaci se skladište u različita skladišta i ne povezuju se, te ne pružaju svoju maksimalnu vrednost. Analize se često vrše po potrebi određenih projekata ili zahteva visokih državnih funkcionera, ali ne postaju sastavni deo sistema izveštavanja i odlučivanja. Primena Uredbe o načinu rada Portala otvorenih podataka¹ je delimična, iako Uredba propisuje da dužnost državnih organa da objavljuju podatke u mašinski čitljivom formatu. Da bi se situacija promenila, potrebno je dodatno edukovati zaposlene o značaju upravljanja podacima, otvorenim podacima, kao i koristima u poslovanju koje im mogu doneti bolje upravljanje i korišćenje podatka.

Baze podataka

Baze podataka su ključne u svakodnevnom radu organizacija, omogućavajući efikasno upravljanje transakcijama poput zakazivanja termina ili obračuna komunalnih usluga u slučaju javnog sektora. Baza podataka predstavlja skup podataka, mada se često pod ovim terminom obuhvataju i sistemi upravljanja bazama podataka. One najčešće podržavaju obradu OLTP (Online Transaction Processing) transakcija, odnosno transakcija koje se javljaju u organizaciji tokom svakodnevnog rada. U slučaju javnog sektora, OLTP transakcije se recimo, aktiviraju prilikom zakazivanja online termina za izdavanje ličnih dokumenata, zakazivanja termina kod izabranog doktora, prilikom obračuna komunalnih usluga i slično.

¹ Sl. glasnik RS, br. 104/2018

Sistemi za upravljanje bazama podataka (DBMS) su ključni alati za skladištenje i pristup podacima u bazama, koje se mogu razvrstati u dve kategorije: relacione i ne-relacione baze podataka. Sistemi za upravljanje bazama podataka, poznati kao DBMS (Database Management Systems), koriste se za skladištenje podataka u baze podataka. Osim skladištenja, ovi sistemi omogućavaju korisnicima da pristupe i rade sa skladištenim podacima, kao i da interaguju sa različitim aplikacijama. DBMS se posebno koristi kada se obrađuju podaci dobijeni iz svakodnevnih transakcija u organizacijama, poznatih kao OLTP transakcije. Poslednjih decenija je razvijeno više različitih tipova baza podataka, koje se razlikuju u određenim specifičnim karakteristikama. Različiti tipovi baza podataka se mogu svrstati u dve grupe:

- Relacione baze podataka – u njima se skladište podaci u tabele koje imaju fiksni broj redova i kolona,
- Ne-relacione baze podataka – poznate su i kao SQL baze podataka, a skladište podatke kroz različite modele, uparene vrednosne ključeve i/ili table sa promenljivim kolonama, čvorovima i granicama.

Baze podataka primarno skladište strukturisane podatke, čija obrada je jednostavnija od obrade polustrukturisanih podataka čije skladištenje omogućavaju neke baze podataka. U baze podataka se najčešće mogu skladištiti samo strukturisani, odnosno podaci koji imaju jasnu strukturu i najčešće su organizovani u tabelama. Primer strukturisanih podataka bi bili podaci iz telefonskog imenika u Excel tabeli, gde se prema kolonama unose ime, prezime, kontakt telefon, dok svaki red predstavlja jednog korisnika. Pored strukturisanih, neke baze podataka nude mogućnost skladištenja polu-strukturisanih podataka, odnosno podataka koji nisu striktno organizovani kao strukturisani podaci. To mogu biti podaci elektronske pošte, podaci za logovanje na određene aplikacije i slično. Dakle, glavna razlika između navedenih tipova podataka je u stepenu organizovanosti podataka, a u skladu sa tim i u jednostavnosti obrade. Strukturisane podatke je lakše i jednostavnije obrađivati, jer ne zahtevaju prethodno usaglašavanje, odnosno strukturisanje.

Za obezbeđenje tačnosti, konzistentnosti i smanjenje mogućnosti greške, ključno je definisati sheme podataka pre skladištenja podataka u baze podataka. Da bi podaci mogli biti skladišteni u baze podataka, potrebno je unapred odrediti kolone i tipove podataka koji će biti skladišteni, odnosno potrebno je izraditi takozvanu shemu podataka. Na taj način se osiguravaju tačnost i konzistentnost i umanjuje mogućnost greške.

Karakteristike baze podataka :

- Podacima mogu pristupiti samo korisnici sa odobrenim pristupom,
- Integritet podataka je obezbeđen ACID (Atomicity, Consistency, Isolation, Durability) transakcijama, koje osiguravaju stabilnost i doslednost u sistemima za upravljanje bazama podataka
- Jednostavna interakcija sa podacima,
- Optimizacija performansi upita je obezbeđena uz pomoć indeksa, odnosno indeksi se koriste kako bi se ubrzala pretraga u tabelama baze podataka,
- Postoje opcije pretrage teksta,
- Optimizovane su za mobilne uređaje,
- Raspodelom resursa i komponenti sistema na način koji najbolje podržava specifične vrste radnog opterećenja optimizuju se efikasnost i performanse sistema,
- Opcije „hostovanja“ mogu biti na matičnoj lokaciji ili na privatnom, javnom, hibridnom ili multi „cloud“-u.

Baze podataka su dobro rešenje za skladištenje i brz pristup podacima u različitim oblastima, poput e-trgovine, bankarstva, sporta i javnog sektora. Uzimajući u obzir navedene karakteristike baza podataka, može se zaključiti da su one odgovarajuća opcija za skladištenje podataka kojima je potrebno brzo pristupiti. Dobri primeri korišćenja baza podataka jesu za skladištenje proizvoda u online prodavnicama, finansijskih dosijea u bankarskom sektoru, sportskih rezultata i statistika, ocena i rezultata studenata, informacija sa aplikacija sa mobilnih telefona. U slučaju javnog sektora, kao što je već navedeno, baze podataka su dobra opcija kada je potrebno skladištiti podatke online rezervacije termina za zamenu ličnih dokumenata, medicinskih kartona pacijenata, poreske administracije, podatke o imovini, koji se koriste trenutno i u operativne svrhe i slično.

Dobar primer upotrebe baza podataka u javnom sektoru Srbije je Centralni registar obaveznog socijalnog osiguranja, koji je uspostavljen sa jasno definisanom shemom i jedinstvenim kodovima za efikasno upravljanje podacima socijalnog osiguranja. Baze podataka su u velikoj meri u upotrebi u javnom sektoru u Srbiji. Dobar primer baze podataka koja je povezana sa drugim bazama podataka u javnom sektoru jeste baza podataka Centralnog registra obaveznog socijalnog osiguranja. U pitanju je baza podataka koja je uspostavljena na osnovu postojećih baza podataka koje vode organizacije obaveznog socijalnog osiguranja i Poreska uprava². U Zakonu o Centralnom registru obaveznog socijalnog osiguranja su tačno navedeni podaci, odnosno kolone koje ova baza mora da sadrži. Uspostavljen je i jedinstveni kodeks šifara, kako bi baza bila uređena.

Iako su baze podataka široko korišćene za skladištenje i upravljanje podacima, one imaju određena tehnološka ograničenja, poput kompleksnosti upravljanja velikim podacima, slabljenja performansi sa rastom količine skladištenih podataka i ažuriranja podataka. Upravljanje bazama podataka zahteva detaljno planiranje i dobro razvijene tehnike upravljanja, što sa rastom količine podataka postaje sve kompleksnije. Rast količine podataka donosi i slabljenje performansi, naročito ukoliko je potrebno sprovesti više upita u istom trenutku ili podržati simultani rad više korisnika. Dodatno, savremeno poslovanje zahteva redovno ažuriranje podataka, što je izazovno sa aspekta očuvanja tačnosti i konzistentnosti podataka, kao i neprekidne dostupnosti sistema korisnicima. Iz navedenih razloga, magacini ili jezera podataka mogu biti bolja opcija za skladištenje velikih podataka.

Magacin podataka

Magacini podataka su specijalizovani za skladištenje uglavnom strukturisanih podataka koristeći procese za ekstrakciju, transformaciju i učitavanje podataka iz raznih izvora, s izazovom vremenske diskrepancije usled pokretanja procesa u određenom vremenskom trenutku. Magacin podataka je tip skladišta podataka u koji se mogu smestiti uglavnom strukturisani podaci, mada postoje i neki magacini za polu-strukturisane podatke. Slično bazama podataka, magacini zahtevaju postojanje unapred definisanih relacionih shema, dok su ključni takozvani ETL procesi („extract, transform, and load“), koji obuhvataju ekstrakciju podataka iz izvora podataka (aplikacija, baza podataka, platformi, društvenih mreža, itd.), potom njihovu transformaciju u odgovarajući format i nakon toga podaci se skladište u magacin. Ovaj proces zahteva vremensku odrednicu u smislu da se pokreće u određeno vreme, možda svakog sata ili jednom dnevno, što dovodi do postojanja vremenske diskrepancije između podataka koji se nalaze u magacinu i poslednje unetih podataka u sistem.

Magacini podataka su idealni za čuvanje obimnih strukturisanih istorijskih podataka namenjenih detaljnim analizama. Magacini podataka su pogodni za korišćenje u slučajevima kada je potrebno skladištiti veliku količinu strukturisanih istorijskih podataka, ali nije potrebno nad njima sprovesti brze i hitne analize, već detaljne i dubinske analize koje zahtevaju vreme. Dakle, magacini podataka nisu

² Sl. glasnik RS, br. 95/2018 i 91/2019

pogodni za skladištenje podataka iz aplikacija koje zahtevaju pravovremenu obradu podataka u transakcione svrhe, što znači da bi bilo poželjno da gradnju magacina podataka prati i izrada baze podataka u koju će se skladištiti podaci potrebni za obavljanje dnevnih transakcija, dok će se magacini koristiti za složene dubinske analize poslovanja.

Naučnici podataka i poslovni analitičari uz pomoć BI alata mogu jednostavno povezati i analizirati podatke skladištene u magacinu. Analiza podataka smeštenih u magacin je jednostavan proces koji mogu obavljati poslovni analitičari i naučnici podataka. Podaci iz različitih magacina podataka se mogu koristiti za poslovne analize, ali se prethodno moraju povezati uz pomoć „Business Intelligence“ alata (BI alati). U pitanju su alati koje koriste poslovni analitičari i naučnici podataka za istraživanje, ispitivanje i analizu podataka na osnovu kojih generišu izveštaje za sve zainteresovane strane.

Karakteristike magacina podataka:

- Omogućavaju integrisanost podataka iz više baza podataka kako bi se stekao jedinstven uvid u podatke,
- Nema volatlnosti podataka, odnosno jednom kada su skladišteni, podaci se ne mogu brisati, menjati, odnosno čuva se svaki prethodni oblik skladištenog podatka,
- Jednostavan pristup podacima i izradi izveštaja koji se postiže konsolidacijom podataka iz više različitih izvora,
- Podržavaju kompleksne upite, agregacije i izveštaje,
- Optimizovani su za radna opterećenja sa velikim brojem upita,
- Kvalitet skladištenih podataka je od velike važnosti, pa zahtevaju definisanje mera za obezbeđenje tačnosti i pouzdanosti podataka
- Sigurnost podataka je obezbeđena kroz kontrolu pristupa zasnovanu na ulogama i enkripciju podataka.

Magacini podataka nalaze široku primenu kako u privatnom sektoru, tako i u javnom sektoru. Magacini podataka su pogodni za korišćenje kako u privatnom, tako i u javnom sektoru. Kada je u pitanju privatni sektor, mogu se koristiti za skladištenje i analizu podataka kupaca u slučaju e-trgovine, upravljanje finansijskim rizikom, praćenje kamatnih stopa, potom za skladištenje i analizu podataka iz proizvodnje i lanaca snabdevanja. U javnom sektoru, magacini podataka su pogodni za praćenje istorijskih evidencija imovine, prikupljanje, analizu i izveštavanje o poreskim transakcijama, kao i u bezbednosne svrhe.

Magacini podataka su manje korišćen tip skladišta podataka u javnom sektoru Srbije. Mali broj institucija je prepoznao prednosti magacina podataka u odnosu na baze i da njihovom upotrebom može unaprediti svoje poslovanje i analize koje obavljaju, što posledično utiče i na kvalitet javnog posla u njihovoj nadležnosti. Jedna od institucija koja koristi magacine podataka je Uprava za sprečavanje pranja novca koja je organ u sastavu Ministarstva finansija. Takođe, preporuka Svetske banke je da se u sistem ePravosuđe³ uvedu magacini podataka kako bi se pratili podaci o broju i tipu obrađenih slučajeva, a na osnovu njih da se izrađuju analize o efikasnosti sudova.

Jezera podataka

Jezera podataka mogu biti od velike koristi u digitalnoj transformaciji organizacija u privatnom i javnom sektoru, nudeći centralizovano skladištenje i naprednu analizu podataka koji potiču iz raznovrsnih izvora. Jezera podataka su moćno sredstvo za sakupljanje, skladištenje i analizu velike količine podataka iz različitih izvora. Ova centralizovana skladišta podataka omogućavaju

³ https://vk.sud.rs/sites/default/files/attachments/Kratak%20pregled%20nalaza%20sa%20preporukama_0.pdf

organizacijama da efikasnije upravljaju informacijama, olakšavaju pristup podacima i omogućavaju dublju analizu i razumevanje informacija o poslovanju. Uvođenjem jezera podataka u poslovanje, organizacije dobijaju bolji uvid u poslovanje, koje od tog trenutka može da počiva na velikom broju odluka donetih na osnovu robusnih analiza velike količine podataka. Razvoj inovativnih rešenja primenom tehnologija poput veštačke inteligencije i analize podataka obezbediće organizaciji konkurentsku prednost i pružanje kvalitetnijih usluga korisnicima, a za sve to osnov će biti jezera podataka.⁴ Ova rešenja se primenjuju kako u privatnom sektoru, tako i u javnom sektoru. U cilju boljeg razumevanja primene koncepta jezera podataka i koristi koje jezera podataka donose, biće analizirana njihova primena u sektoru zdravstva i jedinicama lokalne samouprave. Jezera podataka su ključni deo digitalne transformacije i omogućavaju organizacijama da ostvare konkurentске prednosti u eri informacija.

Jezera podataka omogućavaju skladištenje strukturisanih, polustrukturisanih i nestrukturisanih podataka u njihovom izvornom formatu, obezbeđujući vremenske uštede i veću fleksibilnost organizacijama. Ova karakteristika izdvaja jezero podataka od prethodna dva tipa skladišta podataka, koji dozvoljavaju skladištenje samo strukturisanih i polustrukturisanih podataka. Svi podaci se skladište u izvornom, odnosno sirovom formatu, bez prethodne potrebe za transformacijom u određeni format ili modeliranjem, što omogućava uštede u vremenu organizaciji koja se opredeli za upotrebu jezera podataka. Takođe, sa porastom količine podataka koja se skladišti, jezera podataka mogu da rastu, što znači da u početku nije potrebno definisati veličinu jezera, već ono raste sa potrebama organizacije. Veličina i rast jezera podataka su ograničeni isključivo:

- infrastrukturom na kojoj se gradi jezero podataka,
- raspoloživim finansijskim resursima organizacije,
- kapacitetima menadžmenta i
- karakteristikama koje mogu uticati na kvalitet podataka i performanse jezera podataka.

Sa rastom veličine jezera podataka, upravljanje jezerom podataka postaje sve kompleksnije, a održavanje konzistentnosti, efikasnosti i pristupačnosti podacima skladištenim u jezeru je sve teže.

Jezera podataka pružaju troškovno efikasno rešenje za skladištenje podataka, zahvaljujući mogućnosti korišćenja postojeće infrastrukture i odvajanja skladištenja od obrade podataka. Prednost jezera podataka je troškovna efikasnost u odnosu na druge opcije za skladištenje podataka, koja potiče od činjenice da se jezera podataka mogu graditi i na postojećoj informacionoj infrastrukturi kojom organizacija već raspolaže.⁵ Jezera podataka se mogu implementirati koristeći softver otvorenog koda, koji često dolazi bez licencnih troškova. Pored navedenog, dodatni izvor troškovne efikasnosti jezera podataka leži u tome da je skladište podataka odvojeno od obrade podataka.⁶ To znači da se podaci mogu čuvati u sirovom obliku bez potrebe za skupim postupcima transformacije podataka pre skladištenja.

Jezera podataka pružaju ekonomsku efikasnost organizacijama omogućavajući slojevito skladištenje podataka po klasama važnosti, što znači da podaci koji se retko koriste mogu biti smešteni u delovima jezera sa lošijim performansama u smislu dužeg vremena odgovora na upite, dok bi se češće korišćeni podaci čuvali u delu jezera sa boljim performansama.⁷ Na taj način se obezbeđuju ekonomske koristi,

⁴ <https://www.matillion.com/uploads/pdf/essential-guide-to-data-lakes-designing-data-lakes-to-optimize-analytics.pdf>

⁵ Data lakes, Božić V., Mart 2023

⁶ <https://www.matillion.com/uploads/pdf/essential-guide-to-data-lakes-designing-data-lakes-to-optimize-analytics.pdf>

⁷ <https://www.matillion.com/uploads/pdf/essential-guide-to-data-lakes-designing-data-lakes-to-optimize-analytics.pdf>

jer se resursi usmeravaju tamo gde su najpotrebniji. Dakle, skladištenje podataka po slojevima omogućava prilagođavanje skladištenja po učestalosti pristupa.

Izazovi u upravljanju jezerima podataka, poput očuvanja kvaliteta i sprečavanja stvaranja "močvare podataka", zahtevaju pažljivo planiranje i efikasno upravljanje podacima. Pored navedenih prednosti, jezera podataka donose sa sobom i određene izazove, koji mogu negativno uticati na korisnike. Jedan od izazova je obezbeđivanje kvaliteta podataka koji se skladište u jezeru, kao i način upotrebe tih podataka. Ovaj izazov potiče od mogućnosti skladištenja sirovih podataka u različitim formatima, što može dovesti do nekonzistentnosti podataka i mogućih grešaka. Takođe, sa rastom količine i raznovrsnosti podataka skladištenih u jezeru podataka raste mogućnost da jezero preraste u tzv. „močvaru podataka“. Kako bi se predupredio ovaj izazov, potrebno je pre izgradnje jezera podataka izraditi plan upravljanja skladištenim podacima, sistem skladištenja metapodataka i katalog podataka.

Strukturisani podaci su odmah dostupni poslovnim analitičarima za vršenje analiza i praćenje toka poslovanja, dok im je za nestrukturisane podatke potrebna podrška programera, naučnika i inženjera podataka. Ukoliko su u jezeru skladišteni strukturisani podaci, onda ih bez ikakve transformacije ili modeliranja, mogu koristiti poslovni analitičari podataka. Ukoliko je pak, potrebno transformisati podatke pre korišćenja, onda će poslovnim analitičarima biti potrebna pomoć programera, naučnika podataka i inženjera podataka, koji će podatke prevesti u format koji je potreban analitičarima za dalje sprovođenje analiza, dati predlog u koje svrhe i na koji način se podaci mogu koristiti, kao i sagledati da li postoje neki neočekivani obrasci u podacima koji daju nove uvide u poslovanje.

Karakteristike jezera podataka :

- Skalabilnost jezera podataka je omogućena kroz mogućnost rasta jezera sa obimom podataka koji se u njih skladišti
- Fleksibilnost u obradi podataka je obezbeđena kroz mogućnost skladištenja različitih podataka u sirovom formatu bez prethodno sprovedenih transformacija,
- Troškovna efikasnost se postiže kroz mogućnost gradnje jezera podataka na postojećoj infrastrukturi organizacije,
- Ubrzanje procesa donošenja odluka kroz mogućnost analize podataka u realnom vremenu i brzu identifikaciju i praćenje trendova i pravila u podacima,
- Lak i brz pristup podacima omogućava agilnost organizacija, što ubrzava sprovođenje analiza, inovacija i eksperimenata,
- Kvalitet i konzistentnost podataka se omogućavaju uspostavljanjem sistema upravljanja podacima.

Jezera podataka još nisu našla svoju primenu u javnom sektoru Srbije. Cilj ove analize je upravo da predloži načine na koji se jezera podataka mogu koristiti u javnom sektoru u Srbiji i sektore koji bi ostvarili značajne koristi od implementacije ovog koncepta.

Ključne razlike između baza, magacina i jezera podataka

Baze, magacini i jezera podataka povezuje to što su u pitanju tehnološka rešenja koja se koriste za skladištenje podataka. Sledeća tabela može dati odgovor na pitanje razlika između navedenih tehnologija.

Tabela 1

Karakteristika	Baza podataka	Magacin podataka	Jezero podataka
Tip zadatka	Operacioni i transakcioni	Analitički	Analitički
Tip podataka	Strukturisani i polustrukturisani	Strukturisani i polustrukturisani	Strukturisani, polustrukturisani i nestrukturisani
Schema fleksibilnosti	Rigidna i fleksibilna shema u zavisnosti od tipa baze podataka	Predefinisana shema ili fiksna shema za unos podataka	Ne traži se definisana shema za unos podataka
Schema osvežavanja podataka	U realnom vremenu	Ne mora biti ažurirana na osnovu ETL procesa	Ne mora biti ažurirana na osnovu ETL procesa
Korisnici	Programeri aplikacija	Naučnici podataka i poslovni analitičari	Poslovni analitičari, programeri aplikacija, naučnici podataka
Prednosti	Brzi upiti za skladištenje i ažuriranje podataka	Fiksna shema koja čini rad sa podacima lakim za poslovne analitičare	Lako skladištenje podataka koje pojednostavljuje unos sirovih podataka; Shema koja se primenjuje kasnije kako bi olakšala rukovanje podacima za poslovne analitičare; Razdvojenost skladištenja i obrade podataka
Mane	Može biti ograničenih analitičkih mogućnosti	Definisanje i razvoj šeme podataka mogu biti izazovni, a skaliranje sistema može zahtevati dodatno skaliranje skladišta, jer su ova dva procesa međusobno usko povezana	Zahteva organizaciju i pripremu podataka za korišćenje

*Izvor: <https://www.mongodb.com/databases/data-lake-vs-data-warehouse-vs-database>

Pri odabiru između baza podataka, magacina i jezera podataka, ključna razmatranja uključuju tip podataka, potrebu za unapred definisanim shemama za olakšavanje analiza i trenutno mesto skladištenja podataka. Baze podataka su neophodne prilikom razvoja i rada sa interaktivnim aplikacijama. Ukoliko su pak, organizaciji potrebne detaljnije analize podataka, ona može izgraditi magacin ili jezero podataka, a kako bi se utvrdilo koja od dve opcije je bolja, potrebno je odgovoriti na sledeća pitanja⁸:

1. *Kakvi su podaci – strukturisani, polustrukturisani ili nestrukturisani?* – ukoliko organizacija planira da skladišti samo prva dva tipa podatka, magacin podataka je zadovoljavajuće rešenje,

⁸ <https://www.mongodb.com/databases/data-lake-vs-data-warehouse-vs-database>

dok je jezero podataka bolje rešenje u slučaju potrebe skladištenja i nestrukturisanih podataka.

2. *Koliko će unapred definisane fiksne sheme olakšati analize podataka?* – Magacini podataka zahtevaju unapred definisane sheme koje olakšavaju i pojednostavljaju analize podataka, dok su u jezeru podataka skladišteni sirovi podaci, koji se mogu tek kasnije, kada se javi potreba za njihovim korišćenjem, transformisati.
3. *Gde su podaci trenutno skladišteni?* – Za premeštanje podataka sa trenutnog mesta čuvanja podataka u magacin, potrebno je razviti ETL procese, dok neki tipovi jezera podataka ne zahtevaju nužno prenos podataka, već mogu direktno da pristupe trenutnom mestu na kome se nalaze podaci.

Regulatorni okvir za uspostavljanje jezera podataka

Donošenje posebnih propisa nije neophodno za gradnju i upravljanje jezerom podataka, ali se moraju poštovati propisi o zaštiti ličnih podataka i može se javiti potreba za izmenom sektorskih zakona. Donošenje posebnih propisa koji bi preciznije uredili oblast skladištenja podataka nije potrebno za gradnju ili upravljanje jezerom podataka, ali postoji nekoliko regulatornih zahteva koji moraju biti ispunjeni u pogledu izgradnje i održavanja jezera podataka, naročito ukoliko će jezero podataka sadržati podatke o ličnosti. U zavisnosti od delokruga javnog sektora u kome se država opredeli za gradnju jezera podataka, može biti potrebno donošenje novih ili izmena postojećih propisa koji se tiču upravljanja specifičnim tipom podataka, a koji će biti skladišteni u jezeru podataka. Na nivou Evropske unije, privatnost podataka regulisana je Opštom uredbom o zaštiti podataka (General Data Protection Regulation – GDPR). Odredbe navedene Uredbe moraju se poštovati i prilikom skladištenja i upravljanja podacima u jezeru podataka. U Srbiji ovu oblast uređuje Zakon o zaštiti podataka o ličnosti. Ovaj Zakon daje pravni okvir zaštiti podataka i njegove odredbe moraju biti razmotrene i uvažene prilikom gradnje jezera podataka.

Prilikom skladištenja podataka u jezeru podataka, neophodno je poštovati šest ključnih načela obrade podataka o ličnosti propisanih Zakonom o zaštiti podataka o ličnosti, kao i da se obrada vrši u svrhe propisane Zakonom. Imajući u vidu da podaci koji se skladište u jezero podataka predstavljaju zbirku podataka i da bi trebalo da budu sistematizovani ili strukturisani, na jezera podataka se primenjuju odredbe Zakona o zaštiti podataka o ličnosti. A u skladu sa ovim Zakonom, potrebno je voditi računa o sledećih šest načela obrade podataka o ličnosti⁹:

1. *Zakovitost, poštenje i transparentnost* – što znači da je obaveza obrađivača podataka o ličnosti da podaci budu obrađeni na pošten i transparentan način,
2. *Ograničenje u odnosu na svrhu obrade* – prikupljanje podataka o ličnosti treba da bude u konkretno određene, izričite, opravdane i zakonite svrhe i podaci se ne smeju dalje obrađivati na način koji nije u skladu sa tim svrhama,
3. *Minimizacija podataka* – podrazumeva da podaci budu primereni, bitni i ograničeni na ono što je neophodno u odnosu na svrhu obrade,
4. *Tačnost podataka* – neophodno preduzeti sve mere kojima se obezbeđuje da se netačni podaci o ličnosti bez odlaganja izbrišu ili isprave,
5. *Ograničenje čuvanja podataka* – podaci se moraju čuvati u obliku koji omogućava identifikaciju lica samo u roku koji je neophodan za ostvarivanje svrhe obrade,
6. *Integritet i poverljivost podataka* – podaci se moraju obrađivati na način koji obezbeđuje odgovarajuću zaštitu podataka o ličnosti, što obuhvata i zaštitu od neovlašćene ili nezakonite

⁹ https://www.paragraf.rs/propisi/zakon_o_zastiti_podataka_o_licnosti.html

obrade, slučajnog gubitka, uništenja ili oštećenja. Ovo se čini uz pomoć odgovarajućih tehničkih, organizacionih i kadrovskih mera.

Kako bi obrada podataka bila zakonita, potrebno je da svrha obrade može da se podvede pod jedan od sledećih interesa: zaštita životno važnih interesa, pristanak, legitimni interes, izvršenje ugovora, poštovanje pravnih obaveza, obavljanje poslova u pravnom interesu.

Rukovaoci i obrađivači podataka moraju imenovati lice zaduženo za zaštitu podataka o ličnosti. Imajući u vidu da su jezera podataka zapravo vid skladištenja podataka i da se podaci iz jezera koriste za dalje obrade, svaki vlasnik jezera podataka predstavlja rukovaoca podacima, a može biti i obrađivač. Svaki rukovalac i obrađivač su dužni da, u skladu sa Zakonom o zaštiti podataka o ličnosti imenuju lice zaduženo za zaštitu podataka o ličnosti u sledećim slučajevima¹⁰:

- Obrada se vrši od strane organa vlasti¹¹, osim ako se radi o obradi koju vrši sud u svrhu obavljanja njegovih sudskih ovlašćenja,
- Kada se osnovne aktivnosti rukovaoca ili obrađivača sastoje u radnjama obrade koje po svojoj prirodi, obimu, odnosno svrhama zahtevaju redovan i sistematski nadzor velikog broja lica na koja se podaci odnose,
- Ukoliko se osnovne aktivnosti rukovaoca ili obrađivača sastoje u obradi posebnih vrsta podataka o ličnosti (podaci o ličnosti čijom se obradom otkriva rasno ili etničko poreklo, političko mišljenje, versko ili filozofsko uverenje ili članstvo u sindikatu, kao i obrada genetskih podataka, biometrijskih podataka u cilju jedinstvene identifikacije lica, podataka o zdravstvenom stanju ili podataka o seksualnom životu ili seksualnoj orijentaciji fizičkog lica) ili u obradi podataka koji se odnose na krivične presude, kažnjiva dela i mere bezbednosti, uz dodatni uslov da se obrada vrši u velikom obimu.

Zakon o informacionoj bezbednosti propisuje mere zaštite od bezbednosnih rizika i načela za planiranje i primenu mera zaštite informaciono-komunikacionih sistema, što mora biti uzeto u obzir prilikom gradnje jezera podataka. Osim Zakonom o zaštiti podataka o ličnosti, gradnja jezera podataka je uređena i Zakonom o informacionoj bezbednosti kojim se utvrđuju mere zaštite od bezbednosnih rizika u informaciono-komunikacionim sistemima i određuju se nadležni organi za sprovođenje mera zaštite, koordinaciju između članica i praćenje pravilne primene propisanih mera.¹² Zakon o informacionoj bezbednosti takođe, propisuje načela kojima se treba rukovoditi prilikom planiranja i primene mera zaštite IKT sistema:

1. *Načelo upravljanja rizikom* propisuje da izbor i nivo primene mera treba da se zasniva na proceni rizika,
2. *Načelom sveobuhvatne zaštite* je propisano da se mere primenjuju na svim organizacionim, fizičkim i tehničko-tehnološkim nivoima i tokom celog životnog ciklusa IKT sistema,
3. *Načelo stručnosti i dobre prakse* kaže da se mere primenjuju u skladu sa stručnim i naučnim saznanjima i iskustvima u oblasti informacione bezbednosti,
4. *Načelo svesti i osposobljenosti* propisuje da sva lica koja svojim postupcima efektivno ili potencijalno utiču na informacionu bezbednost treba da budu svesna rizika i poseduju odgovarajuća znanja i veštine.

¹⁰ https://www.paragraf.rs/propisi/zakon_o_zastiti_podataka_o_licnosti.html

¹¹ „Organ vlasti” je državni organ, organ teritorijalne autonomije i jedinice lokalne samouprave, javno preduzeće, ustanova i druga javna služba, organizacija i drugo pravno ili fizičko lice koje vrši javna ovlašćenja.

¹² <https://www.paragraf.rs/dnevne-vesti/150923/150923-vest16.html>

U pogledu bezbednosti podataka, jezera podataka bi trebalo da budu opremljena robusnim sigurnosnim merama kako bi se odbranila od neautorizovanog pristupa podacima, hakerskih napada i krađe podataka i curenja podataka.

Modeli upravljanja i korišćenja jezera podataka

Modeli upravljanja jezerom podataka

Uspešna izgradnja jezera podataka zahteva jasno definisanje problema koji se rešava i plan upotrebe skladištenih podataka kako bi se izbeglo pretvaranje jezera u „močvaru podataka“ i osigurala vrednost ulaganja u odnosu na troškove održavanja. Proces gradnje jezera podataka je identičan u tehničkom smislu bilo u javnom ili u privatnom sektoru, a jezero podataka će doneti benefite organizaciji ukoliko se pridržava sledećih preporuka za gradnju:

- Unapred prepoznati problem koji se očekuje da bude rešen gradnjom jezera podataka;
- Precizno razraditi modele upotrebe i stvarne potrebe, odnosno unapred definisati svrhu gradnje;
- Testirati funkcionalnost i osnovne postavke na realnim problemima;
- Odabrati lokaciju jezera između opcija lokalnog skladištenja (na samoj organizaciji, "in-premises") i skladištenja u oblaku ("cloud"),
- Definirati okvir upravljanja podacima koji obuhvata jasne procedure i politike upravljanja podacima, alate za manipulaciju i obradu podataka i mere bezbednosti,
- Postaviti pravila za označavanje i klasifikaciju podataka za lakšu pretragu i upravljanje.

Dodatno, u slučaju javnog sektora, trebalo bi odrediti instituciju koja je odgovorna za upravljanje jezerom podataka, koja bi pratila i kontrolisala kvalitet podataka, ali i opominjala druge institucije koje u tom jezeru skladište podatke da ih redovno ažuriraju. Takođe, u kontekstu upravljanja podacima u javnom sektoru u Srbiji, uspostavljanje kataloga podataka je od naročite važnosti, imajući u vidu da trenutno nije poznato kojim sve podacima javni sektor raspolaže.

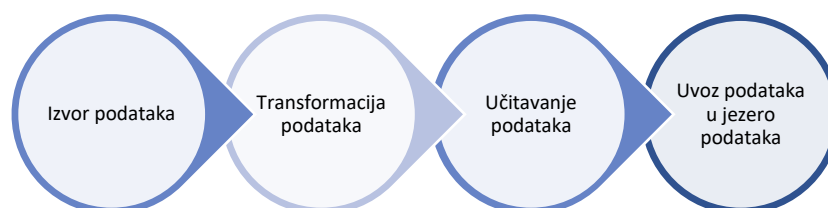
Formiranje multidisciplinarnog tima stručnjaka i uspostavljanje kulture saradnje, uz obuku zaposlenih za korišćenje i analizu podataka, ključni su preduslovi za uspešnu izgradnju i upravljanje jezerom podataka. U pogledu ljudskih resursa, formiranje tima stručnjaka za izgradnju i upravljanje jezerom podataka, koji će činiti stručnjaci za podatke, arhitekturu podataka, upravljanje podacima, analizu podataka i poznavaoци poslovnih procesa, bitno je kako bi se tehničke potrebe uskladile sa postojećim poslovanjem¹³. Ovakav tim će osigurati da se jezero podataka pravilno planira i izgradi na način koji će zadovoljiti sve potrebe korisnika podataka. Dodatni preduslov za izgradnju i upotrebu jezera podataka je kultura saradnje koja bi trebalo da podstakne različite timove u okviru organizacije da sarađuju kako bi svojim specifičnim znanjima i iskustvom doprineli rešenju zadatog problema i donosili zajedničke odluke na osnovu analize raspoloživih podataka. Takođe, obuka zaposlenih za korišćenje podataka i izvođenje analiza je važan korak pre izgradnje jezera podataka.

Izgradnja uspešnog jezera podataka zahteva pažljivo planiranje, investiranje u pravu infrastrukturu i zaposlene, kao i kulturu saradnje koja pridaje vrednost odlukama donetim na osnovu analize podataka¹⁴.

¹³ <https://www.thoughtworks.com/insights/blog/curse-data-lake-monster>

¹⁴ <https://www.matillion.com/uploads/pdf/essential-guide-to-data-lakes-designing-data-lakes-to-optimize-analytics.pdf>

Proces gradnje jezera podataka se sastoji od sledećih koraka:



*Izvor: <https://www.matillion.com/uploads/pdf/essential-guide-to-data-lakes-designing-data-lakes-to-optimize-analytics.pdf>

Efikasno upravljanje jezerom podataka zahteva imenovanje odgovornog pojedinca, razvoj centralizovanog kataloga podataka, jasno definisane uloge i odgovornosti, redovne revizije, kao i detaljan plan integracije i upravljanja raznovrsnim podacima. Ključ ostvarivanja dodate vrednosti od koncepta jezera podataka krije se u odgovarajućem upravljanju. Analize su pokazale da je primenom sledećih praksi upravljanja jezerom podataka ostvarena maksimalna vrednost^{15,16}:

1. Imenovanje pojedinca (data officer-a) koji će biti zadužen za upravljanje jezerom podataka,
2. Izrada i korišćenje centralizovanog kataloga podataka,
3. Izrada matrice uloga i odgovornosti,
4. Vršenje redovnih revizija skladištenih podataka.

Za efikasno korišćenje jezera podataka, organizacija mora da ima plan o tome kako će integrisati različite vrste podataka iz različitih izvora. Ovaj plan treba da identifikuje gde će organizacija prikupljati podatke i kakve podatke će prikupljati. Da bi se različiti podaci usklađivali u okviru jezera podataka, potrebno je koristiti razne alate i strategije koje su prilagođene potrebama organizacije. Na primer, u slučaju Poreske uprave i podataka o poreskim obveznicima, za dobro upravljanje jezerom, potrebno je da bude izrađen plan prikupljanja i povezivanja podataka kao i katalog podataka, a sve u cilju dobijanja celovite slike izvora poreskih prihoda i toka njihove naplate. Pristup ovim podacima omogućio bi Poreskoj upravi da unapredi predviđanje toka i obima naplate poreza, otkrije izuzetke koji se razlikuju od pravila i proveri da nisu u pitanju poreski prekršaji, ali i da unapredi usluge koje pruža korisnicima, kako građanima tako i privredi.

Modeli korišćenja jezera podataka

Jezera podataka se mogu koristiti kao proširena skladišta podataka, u napredne analitičke svrhe, za internet stvari ili za razvoj pametnih gradova. Upotrebna vrednost jezera podataka zavisi od spremnosti organizacije, bilo u javnom ili privatnom sektoru da investira u prave alate i razvoj snažne strategije za integraciju podataka. Koji su to pravi alati zavisi upravo od svrhe u koju organizacija planira da koristi jezero podataka. Neke od mogućnosti za primenu koncepta jezera podataka su¹⁷:

1. *Prošireno skladište podataka* – Jezero podataka može se koristiti kao produženo skladište podataka za informacije koje se retko menjaju ili koriste. Jezero podataka treba koristiti kao

¹⁵ <https://www.matillion.com/uploads/pdf/essential-guide-to-data-lakes-designing-data-lakes-to-optimize-analytics.pdf>

¹⁶ Data lakes, Božić V., Mart 2023

¹⁷ Data Lakes, Božić V., Mart 2023

prošireno skladište samo ukoliko je troškovno efikasnije skladištiti podatke u njemu nego u magacinu. Na primer, organizacija može skladištiti arhivske podatke ili nestrukturisane podatke koji se ne koriste često. U sektoru bezbednosti, jezero podataka može biti korisno za skladištenje podataka iz različitih izvora, poput podataka Ministarstva unutrašnjih poslova, carinskih podataka, podataka o saobraćaju, internet bezbednosti i drugih, a sve u cilju unapređenja različitih segmenata bezbednosti. Naravno, pristup podacima može da se ograniči za različite korisnike.

2. *Napredna analitika* – Jezero podataka može biti odgovarajući izbor u slučaju potrebe za brzim pristupom netransformisanim podacima. To je posebno korisno za naučnike podataka koji razvijaju modele mašinskog učenja. Sektor zdravstva je odličan primer upotrebe jezera podataka u napredne analitičke svrhe, jer se u tom sektoru generišu veliki podaci iz različitih izvora koje je potrebno ne samo skladištiti, već i obraditi i analizirati, a sve u svrhe donošenja odluka. Primer upotrebe jezera podataka u zdravstvu će kasnije biti detaljnije analiziran.
3. *Podrška za internet stvari* – Jezera podataka su ključna komponenta za skladištenje, obradu i analizu velikih količina podataka dobijenih internetom stvari (IoT), jer omogućavaju organizacijama vredne uvide iz raznovrsnih IoT uređaja i senzora. Kroz jezera podataka, organizacije mogu efikasno upravljati i koristiti IoT podatke za informisane odluke i optimizaciju svojih poslovnih procesa.
4. *Pametni gradovi* – Jezera podataka se nalaze u osnovi razvoja pametnih gradova, jer omogućavaju da se podaci iz različitih baza podataka sliju u jezero i da se koriste kasnije za upravljanje pametnim uslugama. Na primer, podaci o saobraćaju, energetici, vodovodu i drugim aspektima grada mogu se integrisati u jezero podataka kako bi se olakšalo pametno upravljanje gradom¹⁸.

Mogućnosti koje proističu iz upotrebe jezera podataka

Jezera podataka nude rešenje za centralizovanje i efikasnije upravljanje velikim količinama podatka koje države proizvode, a koji su često raspoređeni po različitim skladišta podataka. Podaci postaju sve vredniji resurs u svetu. Velika količina podataka nalazi se u vlasništvu države, a da same nisu ni svesne kako količine podataka koju proizvode tako ni količine podataka kojom raspolažu. Države su jedan od najvećih proizvođača podataka, ako ne i najveći, ocene su stručnjaka¹⁹. Svakoga dana različite organizacije u javnom sektoru proizvode, sakupe, analiziraju i distribuiraju velike količine podataka, od onih o vremenskoj prognozi, preko ekonomskih indikatora do podataka o zdravstvenom stanju građana. Ni najrazvijenije države sveta, poput Sjedinjenih Američkih Država, nisu sigurne kog obima su podaci koje proizvode.²⁰ Procena za SAD je da se može govoriti o nekoliko petabajta podataka (jedan petabajt je jednak milion gigabajta)²¹. Međutim, ti petabajti nisu skladišteni na jednom mestu, već svaka agencija federalne vlasti raspolaže svojim skladištima podataka.

¹⁸

https://cris.vtt.fi/ws/portalfiles/portal/73428990/IET_Smart_Cities_2022_Paiho_Opportunities_of_collected_city_data_for_smart_cities.pdf

¹⁹ <https://fedtechmagazine.com/article/2019/01/data-lakes-what-they-are-and-how-they-can-benefit-feds-perfcon>

²⁰ <https://fedtechmagazine.com/article/2019/01/data-lakes-what-they-are-and-how-they-can-benefit-feds-perfcon>

²¹ <https://fedtechmagazine.com/article/2019/01/data-lakes-what-they-are-and-how-they-can-benefit-feds-perfcon>

Koristi koje države imaju od centralizovanog skladištenja raznovrsnih podataka ogledaju se u praćenju porekla i upotrebe podataka, što je ključno za koordinaciju rada različitih organizacija u javnom sektoru i usklađivanje sa strateškim ciljevima. Jezera podataka omogućavaju državama da čuvaju raznolike vrste podataka, od dokumenata i fotografija do video i audio zapisa. Jedna od ključnih prednosti ovakvog sistema je transparentnost koja se ogleda u mogućnosti praćenja ne samo sadržaja podataka, već i njihovog porekla, kao i informacija o tome ko ih koristi i u koje svrhe. Centralizovano skladištenje podataka, kakvo nude jezera podataka, ima značajnu ulogu u koordinaciji i analizi rada različitih organizacija unutar javnog sektora, a koja je značajna sa aspekta pružanja pregleda i usklađivanja aktivnosti različitih organizacija sa strateškim ciljevima države i povezivanjem različitih sektorskih strategija kako bi se lakše koordinisao zajednički rad više sektora. Na taj način, jezera podataka predstavljaju ključni alat za praćenje i unapređenje efikasnosti i efektivnosti u javnom sektoru.

Efikasna akumulacija i konsolidacija podataka iz različitih izvora u jezeru olakšavaju identifikaciju neusklađenosti i grešaka u podacima, čime se unapređuje tačnost analiza i doprinosi integritetu i pouzdanosti sistema koji generišu i koriste podatke. Jezera podataka olakšavaju akumulaciju podataka o određenim indikatorima iz raznih izvora. Na primer, u sektoru zdravstva, podaci o učestalosti oboljevanja od određene bolesti mogu dolaziti iz različitih bolnica i klinika, kao i iz nacionalnih zdravstvenih istraživanja, te bi konsolidacija podataka omogućila jednostavnije poređenje podataka koji se odnose na učestalost oboljevanja, čime bi se lakše identifikovale anomalije ili greške u podacima, kao i pravilnosti u pojavama. Jednom kada se greške otkriju, postaje lakše adresirati ih i ispraviti, što vodi ka poboljšanjima u samim sistemima koji generišu podatke. Pružajući ujedinen pogled na podatke iz više sektora, jezera ne samo da povećavaju tačnost analize, već doprinose i celokupnom integritetu i pouzdanosti sistema koji generiše i koristi podatke.

Podaci skladišteni u jezeru podataka se mogu dopuniti metapodacima čime se omogućava dublji uvid u podatke i čine jezera podataka superiornijim u odnosu na tradicionalna skladišta. Jezera podataka omogućavaju dopunu sačuvanih podataka metapodacima, koji uključuju informacije poput porekla podataka, vremena kreiranja i druge relevantne oznake. Kroz jezera podataka, podaci se mogu detaljnije analizirati, pružajući tako bolji i precizniji uvid. Ova funkcionalnost je posebno korisna za poreske uprave u cilju efikasnije detekcije i prevencije prevara, utaje poreza ili zloupotrebe javnih sredstava od strane javnih funkcionera. Na primer, kada Poreska uprava pretražuje podatke o poreskim obveznicima, može lako identifikovati neobične obrasce ili neusklađenosti u finansijskim izveštajima, upotrebom metapodataka kao što su vremenski pečati ili geografski podaci, što joj pruža snažan alat u borbi protiv finansijskih prevara. Sposobnost dubinskog uvida u podatke čini jezera podataka superiornijim rešenjem za čuvanje podataka u odnosu na tradicionalne metode skladištenja podataka.

Fleksibilnost jezera podataka omogućava organizacijama lako prilagođavanje i reorganizaciju podataka za različite svrhe i korisnike. Fleksibilnost jezera podataka je posebno korisna kada organizacije treba da ponude setove podataka drugim korisnicima, kao u slučaju Zavoda za statistiku. Tradicionalno, kada se setovi podataka dele, oni su često strukturirani na specifičan način, prilagođen njihovoj prvobitnoj nameni. Međutim, uz pomoć jezera podataka, struktura setova podataka može biti otvorena i fleksibilnija. To znači da se podaci mogu lako prilagoditi i reorganizovati za nove svrhe i potrebe korisnika. Umesto da se ograničavaju na fiksnu strukturu koja je možda usko usmerena na određenu upotrebu, različiti korisnici mogu efikasno pristupiti i koristiti širok spektar podataka u različitim formatima, čime je omogućeno lakše kombinovanje i analiza podataka iz različitih izvora, što može dovesti do boljeg razumevanja složenih pitanja i informisanijeg donošenja odluka.

Kroz jezera podataka, korisnici koji skladište podatke mogu pratiti ko i u koje svrhe koristi podatke, omogućujući transparentnost i odgovornost. Svaki korisnik jezera podataka koji skladišti podatke ima

mogućnost da prati koje organizacije koriste te podatke i u koje svrhe, odnosno za koje analize. Ovo je značajno za osiguranje transparentnosti i odgovornosti u upotrebi podataka. Naravno, važan preduslov za ovakvu razmenu i analizu podataka je da korisnici koji skladište podatke daju odgovarajuće dozvole drugim korisnicima za pristup i obradu. Na taj način je omogućeno kontrolisano i sigurno deljenje podataka, što je od ključnog značaja u javnom sektoru, gde se često rukuje osetljivim informacijama. To u praksi znači da, na primer, jedna opštinska administracija može skladištiti podatke o lokalnim projektima i istovremeno koristiti podatke Ministarstva finansija o budžetu. Tako se obezbeđuje bolje razumevanje i koordinacija između različitih nivoa upravljanja, omogućavajući efikasnije planiranje i realizaciju projekata.

Razdvajanjem skladištenja i obrade podataka u jezerima, korisnicima se pruža mogućnost optimizacije troškova i jačanja tržišne pozicije. Skladištenje i obrada podataka su razdvojeni u jezerima, što vodi optimizaciji troškova za korisnike. Njima stoji na raspolaganju mogućnost odabira naprednih alata za obradu podataka dok istovremeno koriste manje sofisticirane, a time i jeftinije opcije za skladištenje podataka, ili obrnuto. To znači da organizacija može platiti samo za resurse koji su joj od važnosti za poslovanje, dok za druge može iskoristiti već postojeće resurse. U kontekstu tržišne konkurencije, jezera podataka olakšavaju ulazak na tržište smanjujući potrebu za velikim investicijama u novu tehnologiju. Preduzeća mogu koristiti postojeće tehnologije kao osnovu za izgradnju novih sistema, čime smanjuju troškove i jačaju svoju tržišnu poziciju. Preciznije, jezera podataka nude ekonomičan i fleksibilan pristup upravljanju podacima koji omogućava organizacijama da efikasnije koriste svoje resurse.

Lanac vrednosti podataka ilustruje proces stvaranja nove ekonomske i društvene vrednosti kroz otvaranje i korišćenje podataka, a sa ciljem unapređenja javnih usluga, stvaranja radnih mesta i podsticanja saradnje između javnog i privatnog sektora. Koncept lanca vrednosti podataka se odnosi na proces stvaranja vrednosti kroz otvaranje i korišćenje podataka²². U ovom modelu učestvuju različiti akteri: snabdevači, sakupljači podataka i programeri. Snabdevači su najčešće organizacije javnog sektora, koje su odgovorne za proizvodnju i validaciju podataka. Sakupljači podataka su odgovorni za proces agregiranja podataka, dok su programeri odgovorni za analizu podataka, kao i razvoj usluga i proizvoda. U Evropskoj uniji, model lanca vrednosti podataka ima cilj da pokaže kako otvoreni podaci mogu doprineti stvaranju novih radnih mesta i poboljšanju opšteg blagostanja, što se postiže ponovnom upotrebom i dopunom podataka kroz saradnju javnog i privatnog sektora. Javni sektor igra ključnu ulogu u fazi kreiranja i validacije podataka, kao i na kraju, kao korisnik rezultata analiza. Privatni sektor je važan za proces sakupljanja podataka, njihovu analizu i razvoj usluga zasnovanih na tim analizama. Krajnji rezultat ovog procesa bi trebalo da budu unapređene javne usluge i rast broja novoootvorenih radnih mesta, zahvaljujući brojnim komercijalnim uslugama koje nastaju na temelju otvorenih podataka. Lanac vrednosti podataka predstavlja način na koji podaci postaju izvor nove ekonomske i društvene vrednosti.

Otvaranje i korišćenje podataka podstiče produktivnost i smanjuje troškove, nudeći ekonomske i društvene prednosti poput inovacija, transparentnosti i unapređenja usluga, doprinoseći održivom razvoju. Motivacija za otvaranje, obradu i korišćenje podataka u svrhu donošenja odluka i razvoja usluga koje stvaraju dodatnu vrednost proizlazi iz kontinuirane potrebe za rastom produktivnosti i smanjenjem troškova, izazova koji su prisutni kako u privatnom, tako i u javnom sektoru. Na primer, kroz otvorene podatke, kompanije mogu identifikovati nove tržišne mogućnosti ili unaprediti svoje

22

https://cris.vtt.fi/ws/portalfiles/portal/73428990/IET_Smart_Cities_2022_Paiho_Opportunities_of_collected_city_data_for_smart_cities.pdf

operacije, dok javni sektor može poboljšati javne usluge i politike, čineći ih preciznijim i prilagođenijim potrebama građana. Dakle, koristi od otvorenih podataka nisu ograničene samo na ekonomski rast i efikasnost; one se takođe, odnose na društvene koristi kao što su transparentnost, odgovornost i inovacije koje doprinose opštem dobru. Otvaranjem i korišćenjem podataka, kako privatni tako i javni sektor mogu zajedno raditi na stvaranju održivog i inkluzivnog ekonomskog razvoja koji odgovara potrebama i izazovima savremenog društva. Cilj ovakvog pristupa je da se maksimiziraju koristi od podataka, uključujući poboljšanje korisničkog iskustva kroz personalizaciju, smanjenje operativnih troškova, optimizaciju vremena potrebnog za analizu podataka, te osnaživanje zaposlenih da efikasnije analiziraju trendove i vrše poslovna istraživanja. Prelazak na poslovanje zasnovano na podacima je izazovan ne samo za javni sektor, već i za privatni, jer zahteva promenu dugogodišnjih praksi i adaptaciju na nove tehnologije i metode rada.

Analiza prakse primene koncepta jezera podataka

Zemlje Evropske unije usvajaju koncept jezera podataka sa ciljem unapređenja poslovanja i pružanja boljih usluga građanima. Ove zemlje su shvatile da rastuća količina podataka koja se generiše kako u pojedinačnim delovima javnog sektora, tako i u javnom sektoru u celini treba da bude skladištena, a potom i analizirana u cilju unapređenja poslovanja javnog sektora i pružanja boljih usluga građanima. U skladu sa tim, mnoge od njih su se opredelile za primenu koncepta jezera podataka. Neke su primenile ovaj koncept samo u jednom segmentu, dok su druge izgradile jezero podataka koje predstavlja centralno skladište za sve delove javnog sektora.

Švajcarska je potrebu za skladištenjem i obradom sve veće količine podataka rešila zamenom starog sistema prikupljanja i obrade podataka, novim, odnosno izgradnjom jezera podataka. Stari sistem se pokazao kao nebezbedan i sa brojnim nedostacima, među kojima su bile i operacione mogućnosti sistema. Kako bi prevazišla ove izazove, modernizovala postojeće državne sisteme, spojila podatke kojima raspolažu različiti delovi javnog sektora i povezala sve zainteresovane strane, švajcarska Vlada se odlučila za izgradnju jezera podataka. Uvođenje ovog koncepta u javni sektor omogućilo je veću bezbednost podataka i lakše preuzimanje podataka. Takođe, pokazalo se kao troškovno efikasnije rešenje od skladištenja podataka u tradicionalnim skladištima. Još jedna uočena prednost je pojednostavljenje poslovne analitike i pretrage podataka²³.

U Finskoj više univerzitetskih klinika koristi jezera podataka u svrhe zdravstvenih istraživanja i ispitivanja. Jedna od najstarijih biobanaka (Auria biobank)²⁴ koristi jezera podataka u svom poslovanju. U pitanju je univerzitetska klinika koja u svom sastavu ima istraživački centar, ali se bavi i dijagnostikom i lečenjem pacijenata. Oni prikupljaju podatke iz kliničkih istraživanja i ispitivanja, kombinuju ih sa podacima koje dobijaju kroz zdravstvene tretmane pacijenata i povezuju ih sa uzorcima iz biobanke. Kako bi se sakupili nestrukturisani podaci, razvijeni su različiti algoritmi. Takođe, koriste modele veštačke inteligencije kako bi predvideli ishode lečenja. Još jedna univerzitetska klinika istražuje podatke prikupljene testiranjem krvi, tkiva, cerebrospinalne tečnosti i drugih zdravstvenih analiza izvršenih nad pacijentima, a sve u cilju otkrivanja uzroka i načina lečenja pacijenata koji boluju od vaskularnih, metaboličkih, muskulaturnih, neuroloških ili onkoloških bolesti²⁵. Svi podaci se skladište u jezero podataka koje je prilagođeno za prikupljanje i analizu ovakvih podataka. Treća klinika je Helsinška univerzitetska bolnica koja skladišti i obrađuje podatke od više 100.000 pacijenata, a imaju

²³ <https://www.consulteer.com/bringing-digitization-to-the-public-sector/>

²⁴ <https://www.auria.fi/tietopalvelu/en/>

²⁵ <https://www.tuni.fi/en/news/profiling-areas-focus-health-data-science-studies-how-machine-learning-and-ai-could-help-treat>

više od milion uzoraka različitih bolesti²⁶. Navedenim podacima upravljaju pomoću dobro organizovanog jezera podataka koje funkcioniše u okviru „cloud“-a i omogućava analizu ogromnog seta podataka. Ove analize se koriste za istraživanja, u patološke i brojne druge svrhe.

Ministarstvo finansija Francuske od 2019. godine radi na izgradnji moderne infrastrukture za upravljanje podacima, koja uključuje izgradnju demo verzije jezera podataka. Počelo se sa proširivanjem online usluga za građane i privredu. Neki od tih servisa su otvaranje jedinstvenog upravnog mesta, jedinstveni nalog za sve usluge javne uprave, izrada usluge sigurnosnih poruka za profesionalne korisnike, uvođenje mogućnosti online plaćanja. Takođe, papirni platni listići za državne službenike su ukinuti i uveden je državni portal za penzijsko osiguranje zaposlenih. Kako bi se upravljalo svim prikupljenim podacima, velika sredstva su uložena u modernu infrastrukturu upravljanja podacima, kako po pitanju internog korišćenja prikupljenih podataka, tako i za eksterne korisnike. Tada se počelo i sa izgradnjom demo verzije jezera podataka.²⁷

U cilju unapređenja nacionalne arhitekture i infrastrukture za upravljanje podacima, Italija je razvila Nacionalnu Digitalnu Platformu za Podatke²⁸, koja obuhvata i razvoj jezera podataka. Cilj izgradnje Digitalne Platforme je lakši pristup, deljenje i analiza velikih količina podataka kojima raspolaže javni sektor. Na ovaj način je država pokazala da ima razumevanja za potrebe razvoja modernih tehnologija poput mašinskog učenja i veštačke inteligencije.

Estonija koristi jezera podataka kako bi unapredila statistiku zemlje. Statistički zavod Estonije koristi skladišta podataka u kojima su podaci sistematizovani i opisani, odnosno svi podaci su pre skladištenja transformisani kako bi se prilagodili skladištu. Dalji razvoj skladištenja podataka treba da olakša ponovnu upotrebu podataka, omogući kombinovanje podataka i poveže podatke iz različitih skladišta. Ova unapređenja treba da dovedu i do brže obrade podataka, kako bi se izradila potpuno nova baza znanja, ali na postojećim informacijama. Kako bi se realizovala ova unapređenja, planirano je da se izgradi jezero podataka, jer ono jedino može zadovoljiti sve kriterijume.²⁹

Naučene lekcije na osnovu primera iz prakse

Modernizacija sistema za prikupljanje i obradu podataka postaje ključna za skladištenje podataka, smanjenje operativnih troškova, povećanje efikasnosti i unapređenje dostupnosti, deljenja i analize podataka. Sigurno i pouzdano skladištenje podataka postaje sve teže obezbediti u starim i nesigurnim sistemima za prikupljanje i obradu podataka u javnom sektoru evropskih zemalja. Stoga ih je potrebno modernizovati i pružiti veći stepen zaštite podataka kako bi se izbegli potencijalni bezbednosni problemi. Pored unapređenja bezbednosti, modernizacija sistema za prikupljanje i analizu podataka pomaže u smanjenju operativnih troškova i povećanju efikasnosti. Modernizacija infrastrukture za upravljanje podacima, uključujući izgradnju jezera podataka, ključna je za poboljšanje dostupnosti, deljenja i analize podataka javnog sektora.

Integracija podataka iz različitih delova javnog sektora olakšava upotrebu veštačke inteligencije za analizu podataka, što može značajno poboljšati donošenje odluka, preciznost prognoza, optimizaciju resursa i kvalitet usluga pruženih građanima. Integracija podataka iz različitih delova javnog sektora omogućava bolje razumevanje i analizu celokupnih operacija, što olakšava donošenje odluka na osnovu šireg konteksta. Upotreba tehnologija poput veštačke inteligencije i analize podataka može

²⁶ <https://healthcapitalhelsinki.fi/finland-best-health-data-in-the-world/>

²⁷ Annual Report, Key Performance Indicators, 2019

²⁸ <https://www.oecd-ilibrary.org/sites/9cada708-en/index.html?itemId=/content/component/9cada708-en>

²⁹ Development Plan of Statistics Estonia 2018 – 2022

transformisati način na koji se podaci koriste za donošenje odluka. To može rezultirati u preciznijim prognozama, optimizaciji resursa i unapređenju usluga. Pored toga, korišćenje jezera podataka može poboljšati usluge koje se pružaju građanima i obezbediti veću transparentnost u radu javnog sektora.

Uvođenje novih tehnologija zahteva obuku osoblja i zapošljavanje stručnjaka za upravljanje podacima i analitikom. Ovaj korak je neophodan kako bi se iskoristile prednosti modernizacije. Građani moraju biti obavješteni o tome kako se njihovi podaci prikupljaju i koriste u okviru jezera podataka. Transparentnost i poverenje su ključni za uspešnu implementaciju.

Razvoj i unapređenje jezera podataka je dugoročan proces. Stalno praćenje tehnoloških trendova i potreba korisnika je ključno za uspeh.

Zainteresovane strane i potencijalne mogućnosti primene koncepta jezera podataka

Jezera podataka se mogu koristiti kako u privatnom, tako i u javnom sektoru, ali imajući u vidu da je cilj ove analize ispitivanje mogućnosti i koristi koje javni sektor može imati od uvođenja jezera podataka u svoje poslovanje, zainteresovane strane će se posmatrati kroz prizmu upravo ovog cilja.

Visoki državni funkcioneri i predstavnici lokalnih vlasti mogu koristiti jezera podataka za složene analize i donošenje strateških odluka. U slučaju javnog sektora zainteresovane strane mogu biti državni funkcioneri na najvišem nivou, poput ministara, direktora organa poput Poreske uprave ili Uprave carina, javnih agencija, državni sekretari i pomoćnici ministara, ali i gradonačelnici. Njima su potrebne složene analize zasnovane na velikim podacima koje bi im pomogle da donesu odluke o strateškim pravcima delovanja sektora i razvoju novih usluga koje unapređuju živote građana.

Zainteresovane strane mogu biti čitavi sektori. Jezera podataka se najčešće grade sektorski: u zdravstvu, prosveti, saobraćaju, ili pak, na nivou jedinice lokalne samouprave. Stoga, među zainteresovanim stranama mogu biti delokruzi javnog sektora ili jedinice lokalne samouprave. Oni su odgovorni za koordinaciju procesa izgradnje jezera podataka, regulaciju i kasnije upravljanje jezerom podataka, kako bi donosioci odluka imali pravovremene informacije.

Kontrolori ispravnosti i transparentnosti takođe, ključni su korisnici jezera podataka u javnom sektoru, koristeći ih za nadzor nad korišćenjem budžetskih sredstava, etičnu i odgovornu upotrebu podataka, razvoj različitih inicijativa. Pored donosilaca odluka, kontrolori ispravnosti i transparentnosti rada javnog sektora bi mogli da budu korisnici podataka iz jezera podataka. Državna revizorska agencija bi mogla lakše da kontroliše poslovanje korisnika budžetskih sredstava. Takođe, nevladine organizacije bi trebalo da kontrolišu i prate da li se jezera podataka izgrađena u javnom sektoru koriste na etičan i odgovoran način, odnosno da li se poštuje Zakon o zaštiti podataka o ličnosti prilikom prikupljanja i obrade podataka. Osim toga, nevladine organizacije mogu koristiti podatke iz jezera podataka za kreiranje svojih ekonomskih, socijalnih ili ekoloških inicijativa.

Za izgradnju i optimalno funkcionisanje jezera podataka neophodni su stručnjaci iz IT sektora, koji predstavljaju direktne zainteresovane strane. Da bi jezero podataka bilo izgrađeno i stvaralo očekivanu dodatnu vrednost, potrebno je uposliti stručnjake iz oblasti informacionih tehnologija, poput inženjera podataka, naučnika i analitičara podataka, koji predstavljaju još jednu zainteresovanu stranu koja je direktno uključena u proces gradnje jezera podataka. Zadatak inženjera je tehničke prirode, odnosno oni treba da rade na dizajnu jezera, da obezbede njegovu funkcionalnost i integrisanost sa postojećom informacionom infrastrukturom. Sa druge strane, naučnici i analitičari podataka treba da pruže uvid u podatke i izvrše analize koje su potrebne donosiocima odluka za unapređenje poslovanja i usluga koje pruža javni sektor. Sigurnost podataka skladištenih u jezero ne sme nikako biti skrajnuta i

zaboravljena, stoga je potrebno da u sektoru u kome je izgrađeno jezero podataka bude imenovano lice u čijoj nadležnosti će biti zaštita podataka skladištenih u jezeru.

Građani i privreda, kao indirektno zainteresovane strane, mogu imati koristi od unapređenih usluga javnog sektora zahvaljujući jezerima podataka, a IT i telekomunikacione firme, mogu biti i direktne zainteresovane strane kroz pružanje podrške izgradnji jezera. Ne smeju se zaboraviti građani i privreda, kao indirektno zainteresovane strane. Usluge koje javni sektor pruža građanima i privredi mogu biti unapređene upravo uvođenjem tehnološkog rešenja kao što je jezero podataka. Raste agilnost javnog sektora, od čega će i građani i privreda imati koristi. Takođe, privatni sektor može biti zainteresovana strana kroz usluge koje pruža javnom sektoru u procesu gradnje jezera podataka. Tu se javljaju IT firme koje se bave prodajom softvera ili hardvera za izgradnju jezera podataka, kao i konsultantske firme koje mogu pružiti podršku javnom sektoru kroz svoje znanje i ekspertizu. Telekomunikacione kompanije su bitne zbog obezbeđivanja potrebne infrastrukture. Međutim, osim kao podrška izgradnji jezera podataka, privatni sektor može biti zainteresovan za kupovinu podataka. Agencije za istraživanje tržišta, banke, privatni osiguravajući fondovi, agencije koje se bave prodajom nekretnina i građevinske firme su samo neke od kompanija u privatnom sektoru koje bi mogle biti zainteresovane za kupovinu podataka od države. Kao što je ranije napomenuto, sigurnost i zaštita ličnih podataka moraju biti na prvom mestu, te pre bilo kakve transakcije, jasno moraju biti propisani okviri za njeno sprovođenje, kako ne bi došlo do zloupotrebe podataka.

Akadska i istraživačka zajednica, u saradnji sa javnim sektorom, može značajno doprineti unapređenju javnih usluga kroz analize i inovacije. Kada se govori o zainteresovanim stranama, akademska i istraživačka zajednica nikako ne smeju biti izostavljene. U saradnji sa javnim sektorom, ova zajednica može svojim znanjem doprineti unapređenju javnih usluga kroz analize, ali i brojne nove ideje i predloge. Osim javnog, akademskoj i istraživačkoj zajednici partner i korisnik istraživanja može biti privatni sektor ili nevladine organizacije. Podaci iz jezera podataka mogu, na primer, biti korišćeni kako bi privatni sektor ponudio usluge građanima ili javnom sektoru, a od kojih će svi imati koristi. Takođe, ovo partnerstvo nevladinom sektoru može doneti nove ideje, ali i alate za borbu za bolje socijalne, ekonomske ili ekološke politike.

U nastavku će kroz primer primene jezera podataka u poljoprivredi biti prikazana uloga zainteresovanih strana, a potom i dodatno razrađena kroz analizu spremnosti Srbije za primenu koncepta jezera podataka i dva predloga za primenu u praksi.

Jezeru podataka u sektoru poljoprivrede

U sektoru poljoprivrede jezera podataka predstavljaju efikasno rešenje za skladištenje, obradu i analizu podataka, važnih ne samo za državu u kontekstu održivog razvoja i klimatskih promena, već i za poljoprivrednike, istraživačke institute, preduzeća i finansijske institucije. Poljoprivreda je sektor u koji je digitalizacija u velikoj meri prodrila i pravi značajne promene, naročito u pogledu obima podataka koji se generišu u poljoprivredi i odluka koje se na osnovu njih donose. Kako podaci ne bi postojali radi podataka, njih je potrebno negde skladištiti, obraditi i iskoristiti u procesu donošenja odluka. Podaci iz sektora poljoprivrede su značajni za državu u pogledu donošenja odluka u segmentu održivog razvoja i klimatskih promena. Pored toga, podaci su značajni za poljoprivrednike, naučne institute koji se bave ispitivanjima u ovoj i srodnim oblastima, preduzeća koja se bave preradom poljoprivrednih proizvoda, ali i finansijske institucije koje plasiraju svoje proizvode na ovom tržištu i brojne druge. Svi oni donose odluke o svom poslovanju na osnovu podataka, a da bi odluke na osnovu podataka mogle biti donete, potrebno je sakupiti i obraditi podatke. I ovaj sektor, kao i većinu drugih sektora, karakteriše veliki broj različitih informacionih sistema koji generišu podatke, kako u Srbiji tako

i u drugim zemljama. Kako bi svi raspoloživi podaci mogli da se iskoriste i kapitalizuju, jezera podataka mogu biti dobro rešenje³⁰.

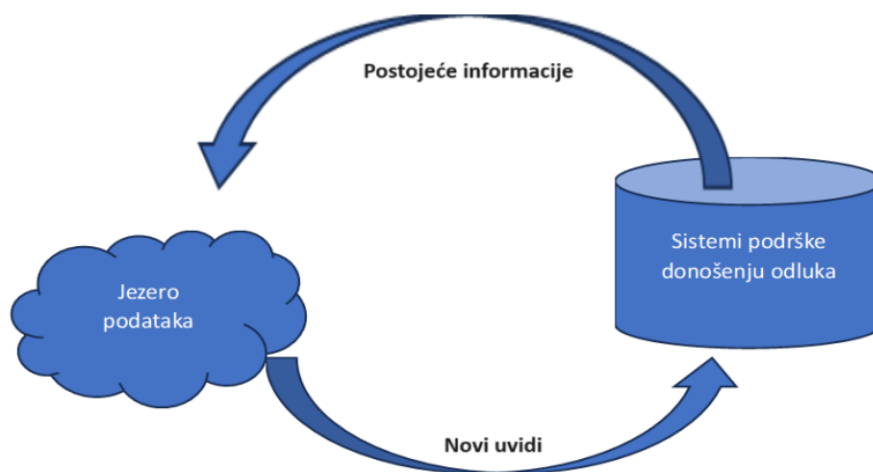
Savremene tehnologije omogućavaju brže i efikasnije analize različitih setova podataka u poljoprivredi, koji treba da identifikuju korelacione veze i pomognu u proceni uticaja novih praksi na okolinu. Za poljoprivredu je ključno istražiti različite setove podataka, poput onih koji se tiču vrsta biljaka i/ili zemljišta, koji potiču iz različitih izvora, kako bi se pronašla korelaciona veza i kako bi se procenio uticaj novih poljoprivrednih praksi na životnu sredinu, na primer³¹. Ključ jeste da se istraže podaci kako bi se pronašlo najbolje rešenje za postojeće probleme poljoprivrednika. U te svrhe je potrebno istražiti i ispitati što je moguće više opcija, a savremene tehnologije nude brži i lakši način da se podaci od eksperimentalnih istraživanja i ispitivanja brže istraže i analiziraju i da se na osnovu njih donesu bolje odluke.

Jezero podataka nudi rešenje koje omogućava integraciju sistema i pristup podacima za širok spektar korisnika, omogućavajući da se podaci kapitalizuju na najbolji mogući način. Poljoprivredne sisteme odlikuje kompleksnost. Uticaj različitih poljoprivrednih praksi na životnu sredinu, brojni akteri poljoprivrednih procesa, mogući brojni uticaji poljoprivrednih praksi na istraživanje i razvoj, finansije, različiti nivoi zrelosti donosilaca odluka, (ne)potpunosť podataka, različiti proizvođači i korisnici podataka su sve zainteresovane strane i elementi koji treba da budu uzeti u razmatranje kada je u pitanju odlučivanje o različitim rešenjima za izazove sa kojima se poljoprivreda suočava. Stoga, koncept jezera podataka može da bude dobro rešenje, uzimajući već navedene karakteristike ovog koncepta. Jezero podataka će doneti prednosti poljoprivredi u vidu mogućnosti interakcije između različitih sistema koji generišu podatke, pristup podacima će biti omogućen različitim korisnicima, od kreatora politika, do naučnika i samih poljoprivrednika. Na ovaj način podaci će moći da se iskoriste u svrhe kapitalizacije na najbolji način od strane svih onih kojima su zapravo potrebni.

³⁰ Cédrine Madera, Anne Laurent, Thérèse Libourel Rouge, André Miralles. How can the data lake concept influence information system design for agriculture?. 11th European conference dedicated to the future use of ICT in the agri-food sector, bioresource and biomass sector (EFITA 2017), Jul 2017, Montpellier, France. pp.181-182

³¹ Cédrine Madera, Anne Laurent, Thérèse Libourel Rouge, André Miralles. How can the data lake concept influence information system design for agriculture? 11th European conference dedicated to the future use of ICT in the agri-food sector, bioresource and biomass sector (EFITA 2017), Jul 2017, Montpellier, France. pp.181-182

Ilustracija 1



**Izvor: Cédrine Madera, Anne Laurent, Thérèse Libourel Rouge, André Miralles. How can the data lake concept influence information system design for agriculture?*

Analiza spremnosti i preporuke za implementaciju koncepta jezera podataka u Republici Srbiji

Strategijom razvoja informacionog društva i informacionih tehnologija u Republici Srbiji za period od 2021. do 2026³² definisan opšti cilj zemlje u ovom domenu, a to je razvijeno informaciono društvo i elektronska uprava u službi građana i privrede i unapređena informaciona bezbednost građana, javne uprave i privrede. U istom dokumentu je analizirano stanje Srbije u trenutku pisanja Strategije u posmatranoj oblasti i predstavljeni su napori i planovi javnog sektora u oblasti digitalizacije. Gotovo da nema sektora u kome nije moguće i ne planira se uvođenje elektronskih usluga. To znači da će količina podataka koje država proizvodi rasti sa svakim novim servisom čije uvođenje se planira. Međutim, ukoliko se ovi servisi ne uvežu i podaci ne ukrste, potencijal za naprednu analitiku i donošenje odluka na osnovu velikih podataka, biće propušten.

Integracija podataka u zajedničko jezero podataka bi mogla znatno ubrzati analize velikih podataka i dugoročno doneti značajne prednosti u kreiranju javnih politika, praćenju strateških dokumenata i procesu donošenja odluka. U javnom sektoru Srbije već se u velikoj meri koriste baze podataka, a neke institucije su uvele ili razmatraju uvođenje magacina podataka. Međutim, sa razvojem novih elektronskih usluga i korišćenjem baza i magacina podataka za skladištenje podataka, analize velikih podataka bi mogle da zahtevaju više vremena nego u slučaju organizovanja svih sistema tako da se njihovi podaci skladište u jedno zajedničko jezero podataka. Iz trenutne perspektive razvoja i upravljanja podacima u javnom sektoru Srbije, ovo može biti veliki zalogaj. Međutim, nesumnjivo bi u dugom roku doneo značajne benefite kreatorima javnih politika, olakšao praćenje realizacije strateških dokumenata i donošenje odluka.

Uvođenje jezera podataka u javni sektor Srbije bi trebalo da omogući jednostavan pristup i transformaciju istorijskih podataka za analitičke svrhe, nakon što se podaci uvezu u jezero iz starih sistema. Uvođenje jezera podataka u javni sektor u Srbiji zahtevalo bi migraciju podataka iz postojećih skladišta i, u nekim slučajevima, već zastarelih sistema, u noviji sistem, odnosno noviju arhitekturu. Pri

³² Sl. glasnik RS, br. 86/2023-5

tome, mora se uzeti u obzir zakonom propisana obaveza da se neke baze podataka, odnosno dokumenta, moraju čuvati decenijama ili trajno. Ovo nije prepreka za izgradnju jezera podataka, jer se svi ovi podaci mogu skladištiti tu i čuvati koliko god da je potrebno. Takođe, podaci mogu da budu skladišteni u formatu koji potiče iz prethodnog sistema, odnosno podaci iz baza ili magacina se mogu migrirati u jezero. Jezera podataka su upravo iz tog razloga odličan izbor, jer omogućavaju na jednostavan način da se pristupi starim istorijskim podacima, ali i da se oni uz pomoć transformacije prevedu u neki drugi format i koriste u analitičke svrhe.

Za maksimizaciju vrednosti podataka javnog sektora u Srbiji, a pre implementacije koncepta jezera podataka, biće potrebno sprovesti sistematizaciju i katalogizaciju proizvedenih podataka. Upravljanje podacima predstavlja i predstavljaće veliki izazov, imajući u vidu sve veći priliv različitih podataka iz različitih izvora, a u budućnosti će ih biti još više. U razvijenim zemljama akcentat je na skladištenju i analizama podataka i određivanju optimalnog načina za sprovođenje ova dva procesa. Srbija, kao zemlja koja ima za cilj da se visoko pozicionira na digitalnom nebu, takođe treba da razmišlja u ovom smeru. Kako skladištiti velike količine podataka i kako ih iskoristiti na najbolji mogući način. Jezera podataka mogu biti rešenje za prvi od ovih izazova. Međutim, da bi se koncept jezera podataka implementirao u Srbiji, potrebno je da se pre svega svi podaci sistematizuju, odnosno da se sagleda ko su proizvođači podataka i da se popiše koji se sve podaci generišu. Odnosno, trebalo bi sačiniti katalog podataka koje generiše javni sektor. Kako bi imale jasnu sliku o podacima koje proizvode, ali i o podacima koji su im potrebni za poslovanje, a proizvode ih druge institucije, potrebno je da postoji katalog podataka. Trenutno država Srbija nema informaciju o svim podacima koje proizvodi i koji joj mogu biti na raspolaganju za donošenje odluka. Uzimajući u obzir plan da se dodatno razvijaju digitalne usluge, podaci će se gomilati, a jasan uvid niko neće imati. Stoga, katalog podataka je neophodan zarad dobrog upravljanja podacima i dobijanja maksimalne vrednosti od njih.

Trenutno u Srbiji nedostatak standarda za podatke i neusklađenost informacionih sistema u različitim delovima javnog sektora otežava njihovu međusobnu povezanost i analizu. Pored nedostatka kataloga podataka, u Srbiji se kao izazov izdvaja činjenica da nisu propisani standardi u pogledu podataka, već da svaki proizvođač podataka može da proizvodi podatke u formatu u kome to njemu odgovara. Proces digitalizacije je otpočeo simultanim uvođenjem informacionih sistema u različite delove javnog sektora, bez plana i ideje o njihovom povezivanju i uvezivanju kako bi se podaci mogli međusobno povezati zarad vršenja analiza. Stoga, različite institucije danas proizvode različite podatke koje je trenutno nemoguće ukrstiti i analizirati. Na primer, postoji više različitih sistema koji prikupljaju podatke o zaposlenima u javnom sektoru koji međusobno nisu povezani, a rezultat toga je da je tačan broj zaposlenih još uvek nepoznat, jer različite institucije raspolažu različitim brojevima. Uvođenje jezera podataka bi omogućilo da se svi podaci sliju u jedno skladište, da se povežu i uporede, čime bi utvrdilo i u kom sistemu se javljaju greške i iz kojih razloga, a mogle bi potom i da se isprave.

Izazovi implementacije jezera podataka uključuju raznovrsnost formata podataka i nedostatak razumevanja značaja i mogućnosti elektronskih podataka i analitike. Osim različitih sistema i nejedinstvenosti u pogledu formata podataka, jedan od izazova u pogledu celokupnog procesa digitalizacije, pa tako i mogućnosti implementacije jezera podataka u Srbiji jeste nerazumevanje značaja podataka u elektronskom formatu i analiza koje se na osnovu tih podataka mogu izraditi. Poslovanje javnog sektora u Srbiji se i dalje u velikoj meri oslanja na papirnu formu. Često se dokumenta koja su izvorno u elektronskom obliku štampaju i arhiviraju u registratoru, u čemu se ogleda nerazumevanje mogućnosti koje digitalizacija pruža za analizu podataka i unapređenje poslovanja. Rukovodioci institucija su u razgovoru sa stručnjacima iz oblasti analize podataka iskazali oduševljenje saznanjem koje sve analize i na koji način mogu biti vršene, samo ako se podaci na adekvatan način čuvaju u digitalnom obliku. Dakle, potrebno je da se rukovodioci informišu i obuču o koristima u

poslovanju koje mogu ostvariti upotrebom naprednih analitičkih alata koji se zasnivaju na obradi velikih količina podataka.

Definisanje jasnih uloga svake institucije uključene u upotrebu jezera podataka kroz politike upravljanja podacima, ključno je za razvoj jezera podataka u Srbiji. Jasno definisane uloge svake institucije korisnika podataka bi doprinele bržem razvoju jezera podataka u Srbiji. Uloge se mogu definisati pomoću politika za upravljanje podacima koji su skladišteni u jezeru podataka. Ove politike treba da uključe instituciju odgovornu za prikupljanje, skladištenje, razmenu, agregaciju i analizu podataka. Razvijene zemlje članice OECD-a su već počele da razvijaju politike i principe upravljanja podacima. Srbija radi na tom planu, međutim ove politike je potrebno prilagoditi specifičnostima sektora u kome se planira gradnja jezera podataka.

Nedostatak stručnjaka iz IKT oblasti i niske plate u javnom sektoru su još jedan izazov koji je potrebno prevazići pre početka implementacije koncepta jezera podataka. Pored prilagođavanja politika sektorskim specifičnostima, izazov za Srbiju u pogledu implementacije koncepta jezera podataka mogu biti ljudski resursi. Upravljanje i korišćenje podataka skladištenih u jezeru podataka zahteva odlično poznavanje i razumevanje savremenih tehnologija.³³ U Srbiji ne postoji mnogo naučnika podataka. Takođe, javni sektor nije konkurentan poslodavac na tržištu kadrova iz oblasti informaciono-komunikacionih tehnologija (IKT), jer nudi značajno niže plate od privatnog sektora. Ovaj izazov naročito dolazi do izražaja imajući u vidu da je za svako jezero podataka koje bi se izgradilo potrebno imenovati jedno lice koje bi direktno bilo zaduženo za upravljanje tim podacima. Stoga, potrebno je prvo uložiti u razvoj ljudskih kapaciteta koji bi mogli da čine mrežu „upravljača“ podacima skladištenim u jezerima podataka. Potom, potrebno je osmisliti način na koji bi javni sektor mogao da konkuriše privatnom u segmentu privlačenja kadrova iz IKT oblasti, jer će razvoj ljudskih kapaciteta bez politike privlačenja kadrova rezultirati odlaskom svih stručnjaka u privatni sektor.

Postojanje Državnog centra za upravljanje i čuvanje podataka u Kragujevcu olakšava tehnički deo implementacije jezera podataka, ali su ipak potrebna dodatna ulaganja u hardverska rešenja za skladište podataka. Imajući u vidu da jezera podataka, u zavisnosti od svrhe korišćenja, ne zahtevaju nužno značajna dodatna ulaganja u digitalnu infrastrukturu, već se mogu izgraditi na postojećoj, Srbija ne bi trebalo da ima značajnijih izazova u tehničkom delu realizacije projekta implementacije jezera podataka. Postojanje Državnog centra za upravljanje i čuvanje podataka u Kragujevcu neznatno olakšava proces implementacije koncepta jezera podataka, imajući u vidu da postoji fizički prostor za skladištenje. Međutim, ono šta nedostaje jeste skladište podataka koje bi bilo smešteno u taj prostor. U tom segmentu su potrebna ulaganja.

Za uspešnu izgradnju jezera podataka u Srbiji potrebno je opremiti Državni centar za upravljanje i čuvanje podataka, kreirati politike za zapošljavanje i zadržavanje stručnjaka iz oblasti tehnologija i jasno definisati svrhu i ciljeve ovog projekta. Za izgradnju jezera podataka u Srbiji potrebno je da se fizički raspoloživ prostor u Državnom centru za upravljanje i čuvanje podataka popuni opremom potrebnom za rad jezera podataka, odnosno hardverom koji predstavlja skladišta podataka. Potom, bitno je kreirati odgovarajuće politike zapošljavanja i zadržavanja visokostručnog kadra iz oblasti savremenih tehnologija koji bi upravljali podacima. I poslednje, ali nikako manje bitno, jeste precizno definisati svrhu gradnje jezera podataka u zemlji, što će biti predstavljeno u nastavku.

³³ <https://www.frontiersin.org/articles/10.3389/fdata.2022.945720/full>

Primer upotrebe 1 - Jezero podataka u sektoru zdravstva

Digitalizacija zdravstva postaje sve aktuelnija tema koja zahteva prethodnu izgradnju adekvatne digitalne infrastrukture. Elektronski podaci postaju sve potrebni resurs u zdravstvu. Na primer, u Sjedinjenim Američkim Državama oko 85% bolnica je počelo sa digitalizacijom zdravstvenih dosijea pacijenata³⁴, a i u Srbiji se aktivno uvodi digitalizacija u sektor zdravstva. Međutim, za digitalizaciju zdravstva potrebno je izgraditi odgovarajuću digitalnu infrastrukturu, kojoj pripadaju i jezera podataka.

Portal eZdravlje u Srbiji nudi korisnicima interaktivnu platformu za komunikaciju sa lekarima, kao i pristup zdravstvenim informacijama i planiranje skrininga, doprinoseći efikasnijoj dijagnostici i prevenciji. Trenutno u Srbiji postoji portal eZdravlje³⁵ koji omogućava povezivanje i komunikaciju lica sa doktorom, zdravstvenim savetnikom, trenerom, prijateljem i osiguravajućom kućom. Povezivanje podrazumeva direktnu razmenu informacija o zdravstvenom stanju, rezultatima analiza poput laboratorijskih analiza, radioloških skrininga, kao i o različitim terapijama koje lice primenjuje. Ova komunikacija se može odvijati putem tekstualnih, audio ili video poruka, a cilj komunikacije je brza razmena informacija zarad dijagnostike i prevencije. Pored mogućnosti komunikacije, portal omogućava i pristup naučnim radovima iz oblasti medicine, preciznije, prevencije. Takođe, korisnici mogu proveriti svoje zdravstveno stanje i da li pripadaju rizičnoj grupi. Portal može da predstavlja i svojevrsni medicinski kalendar koji će korisnika obavestiti o zakazanim skriningima. Osim podsetnika, korisnik će dobiti informacije o načinu pripreme za odgovarajući skrining i vremenskom periodu potrebnom za pripremu.

Pored komunikacije sa medicinskim stručnjacima, eZdravlje nudi funkcije praćenja navika i fizičkih aktivnosti putem povezivanja sa pametnim uređajima, kao i deljenje tih rezultata sa trećim licima. eZdravlje nudi mogućnost praćenja navika i fizičkih aktivnosti, ali i postavljanje ciljeva u ovom domenu i njihovu realizaciju. Takođe, postoji mogućnost da se treće lice, odnosno doktor, savetnik ili trener ovlaste da oni postavljaju ciljeve i da prate njihovu realizaciju. Unosom zahtevanih medicinskih podataka u sistem, korisnik može dobiti analizu svog zdravstvenog stanja od analitičkog sistema. Analiza će biti praćena i odgovarajućim preporukama. Kako bi korisnik mogao da prati svoje navike bez fizičkog unosa podataka u sistem, data je mogućnost povezivanja sistema sa korisnikovim pametnim uređajima. Na ovaj način se kreira jedinstvena baza podataka o svakom korisniku.

Novi Zakon o zdravstvenoj dokumentaciji i evidencijama u oblasti zdravstva³⁶ postavio je osnov za povezivanje svih informacionih sistema u zdravstvu u jedinstveni Republički integrisani zdravstveni informacioni sistem (RIZIS). U ovom sistemu će biti integrisani podaci iz eKartona, informacionog sistema Republičkog zavoda za zdravstveno osiguranje, Instituta za javno zdravlje Srbije Dr Milan Jovanović Batut, Agencije za lekove i medicinska sredstva, kao i drugih koji tek treba da se uspostave.

eKarton će omogućiti lekarima i pacijentima digitalni uvid u lečenje, dok će eBolovanje i eUputi olakšati praćenje bolovanja i upućivanje na dalje lečenje, eliminišući potrebu za papirnom dokumentacijom i nepotrebni odlascima kod lekara. Sistem eKarton, koji treba da postane funkcionalan od januara 2025. godine, omogućiće izabranom lekaru kompletan uvid u tok lečenja pacijenta, ali i samom pacijentu. Sa druge strane, doktori medicine specijalisti će imati uvid samo u dokumentaciju koja je relevantna za oblast kojom se bave. Na ovaj način je ukinuta praksa nošenja papirne dokumentacije. Takođe, radi se na uspostavljanju sistema eBolovanje koji treba da omogućiti

³⁴ Roski, J., Bo-Linn, G. W., & Andrews, T. A. (2014). *Creating Value In Health Care Through Big Data: Opportunities And Policy Implications. Health Affairs, 33(7), 1115–1122*

³⁵ <https://www.e-zdravlje.gov.rs/landing/?v=20231018>

³⁶ Sl. glasnik RS, br. 92/2023

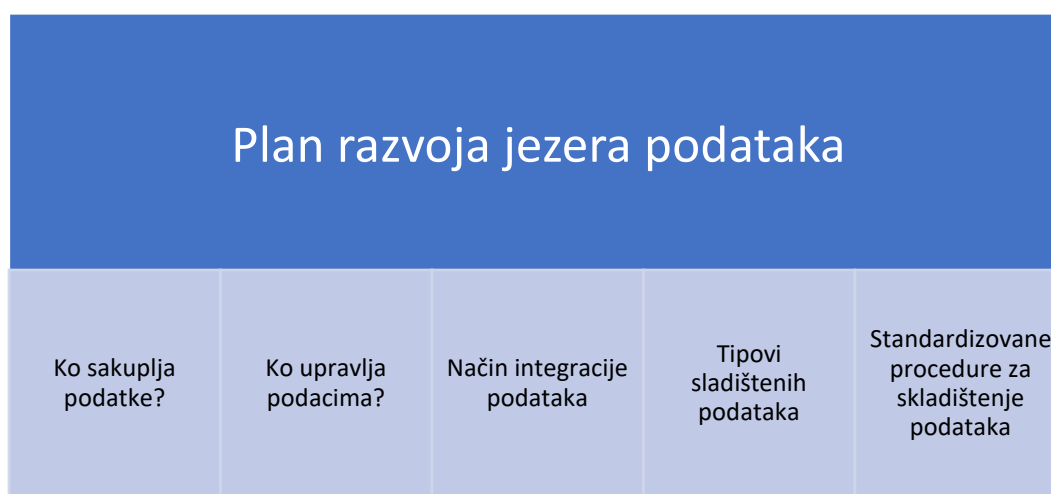
uvid u istorijat bolovanja i da prekine praksu odlaska kod izabranog lekara po doznake, kao i na razvoju eUputa koji treba da omogući da lekari specijalisti upućuju pacijente na dalje lečenje bez vraćanja pacijenata kod izabranog lekara po uput.

Pored navedenih sistema, u oblasti zdravstva se radi na uspostavljanju Registra genetičkih i biomedicinskih podataka, koji će omogućiti rano otkrivanje genetički uslovljenih bolesti, razvoj personalizovane medicine i moderne dijagnostike.

Sa razvojem brojnih zdravstvenih informacionih sistema, raste obim podataka, čije upravljanje i analize mogu biti olakšane uspostavljanjem jedinstvenog skladišta podataka, što treba da doprinese i unapređenju prevencije i predikcije bolesti. Može se zaključiti da trenutno postoje, kao i da se radi na uspostavljanju mnogobrojnih sistema koji generišu različite podatke pacijenata koji mogu da se koriste za predikciju i prevenciju bolesti, kao i da će količina podataka koji će se generisati i skladištiti rasti sa protokom vremena. U pitanju su podaci koji će poticati iz različitih sistema podataka i koji ukoliko se ne uspostavi jedinstveno skladište tih podataka i ne organizuje na adekvatan način, najverovatnije neće biti iskorišćeni na najbolji mogući način.

Implementacija jezera podataka u sektoru zdravstva u Srbiji zahteva detaljan plan koji će definisati upravljanje i integraciju heterogenih zdravstvenih podataka, kao i uvođenje radnog mesta administratora podataka za kontrolu procesa uvoza, skladištenja i manipulacije podacima. Jezera podataka se mogu koristiti u sektoru zdravstvene zaštite, kao što je već kroz više primera napomenuto. Prvi korak izgradnje jezera podataka u sektoru zdravstva bio bi donošenje plana koji bi propisao ko bi sakupljao i upravljao podacima, način integracije podataka u jezero podataka i tipove podataka koji bi bili skladišteni. Plan je veoma značajan, jer su podaci koji se generišu u ovom sektoru heterogeni, te je potrebno definisati način njihovog povezivanja. Na primer, zdravstveni podaci mogu da se povežu na osnovu vremenske ili mesne odrednice, ali samo ukoliko se sačini dobro skladište podataka. Međutim, izgradnja dobrog skladišta podataka može biti veliki izazov u slučaju zdravstvenih podataka, imajući u vidu da podaci potiču iz različitih izvora i da im se razlikuju pristupni protokoli. Kako bi se podaci skladištili na odgovarajući način i kako bi bili upotrebljivi u analitičke svrhe, odnosno kako bi se obezbedila realizacija ciljeva plana, potrebno je predvideti uvođenje radnog mesta administratora podataka koji bi kontrolisao i upravljao procesima uvoza podataka, njihovim skladištenjem i manipulacijom.

Ilustracija 2



*Izvor: Autor

Plan za izgradnju jezera podataka treba da proizađe iz analize postojećih protokola za prikupljanje i skladištenje podataka, koji treba da postave osnov za standardizaciju procedura i eliminaciju dupliranja podataka, čime se obezbeđuje efikasnost i ekonomičnost u pripremi, pristupu i analizi podataka. Osnov za izradu plana treba da bude analiza svih postojećih protokola za prikupljanje i skladištenje podataka, a što će kasnije omogućiti da u planu budu definisane standardizovane procedure za skladištenje podataka u jezeru podataka. Ovo unapređuje efikasnost pripreme podataka za istraživanje i analizu. U mnogim zemljama koje još uvek nisu uspostavile jezero podataka, što je slučaj i u Srbiji, podaci se prenose preko različitih vrsta eksternih memorija ili se kopiraju u tabele i šalju korisnicima. Ovaj pristup vodi do dupliranja podataka, s obzirom da svaki korisnik ima svoju lokalnu verziju podataka, što nije efikasno ni sa finansijskog aspekta, zbog potrebe za obimnim memorijskim prostorom, ni sa aspekta vremena potrebnog za prikupljanje i obradu podataka. Jezero podataka nudi rešenje za ove izazove, pružajući jedinstveno skladište za različite vrste podataka i omogućavajući brojnim korisnicima efikasan pristup i pretragu tih podataka. **Ključni preduslov za funkcionalno jezero podataka je uspostavljanje standarda za prikupljanje podataka i uspostavljanje kataloga podataka.**

Benefiti koje sektor zdravstva može imati od ovog tehnološkog rešenja su sledeći:

- integracija podataka iz javnog i privatnog sektora za bolje analize,
- finansijske uštede putem personalizovanih terapija i praćenja zdravlja populacije,
- unapređenje kliničkih ispitivanja,
- efikasnije upravljanje materijalnim i ljudskim resursima,
- detaljno ispitivanje i validacija istih podataka iz različitih izvora.

Uvođenje jezera podataka u zdravstveni sistem Srbije omogućilo bi povezivanje, a posledično i analizu, podataka javnog i privatnog zdravstva, što bi doprinelo napretku efikasnosti zdravstvenog sistema. Izgradnja jezera podataka u sektoru zdravstva neće koristiti samo državnom zdravstvu, već i sistem privatne zdravstvene zaštite može ostvariti značajne benefite od uvođenja jezera podataka, naročito imajući u vidu planirano povezivanje privatnog i javnog zdravstva. Jezera podataka predstavljaju odličan alat za skladištenje i analizu podataka iz oba ova izvora. U njima se mogu skladištiti podaci iz zdravstvenih kartona pacijenata, kliničkih ispitivanja, podaci o toku lečenja određenih bolesti, demografski podaci, podaci laboratorijskih istraživanja, podaci sa skenera, ali i podaci koje prikupe sami pacijenti putem pametnih uređaja. Svi ovi podaci mogu bez problema biti skladišteni u jezero podataka, u koje je moguće vremenom dodavati nove podatke, kako se drugi sistemi u zdravstvu razvijaju i kako se javljaju potrebe za njihovom analizom i povezivanjem sa postojećim podacima.

Implementacija jezera podataka u sektoru zdravstva donela bi dva glavna benefita: finansijske uštede i unapređenje kliničkih istraživanja³⁷. Navedeni benefiti se postižu primenom personalizovanih terapija za pacijente, koje su novim Zakonom predviđene, a čije uvođenje i funkcionisanje bi bilo efikasnije ukoliko bi se u zdravstvu izgradilo jezero podataka. Analizom podataka o pacijentima, poput medicinskih profila, rezultata rendgenskih snimanja, magnetne rezonance, kao i podataka iz medicinske literature, može se izraditi personalizovani tretman za svakog pacijenta, što uključuje procenu rizika i određivanje odgovarajuće terapije. Dodatno, integracija podataka sa pametnih uređaja pacijenata pruža uvid u navike pacijenata, što dalje doprinosi optimizaciji terapija. Takođe, postoji mogućnost slanja podsetnika pacijentima za uzimanje terapije ili povećanje fizičke aktivnosti, što poboljšava rezultate lečenja. Na ovaj način se detaljno prati tok terapije, njena učinkovitost i omogućava pravovremena reakcija nadležnog lekara, čime se sprečava rasipanje sredstava na

³⁷ Roski, J., Bo-Linn, G. W., & Andrews, T. A. (2014). Creating Value In Health Care Through Big Data: Opportunities And Policy Implications. *Health Affairs*, 33(7), 1115–1122

ponovljene terapije koje nemaju efekta. Takođe, može se pratiti delotvornost određenog leka. Na ovaj način su i unapređena ispitivanja, ali i ostvaruju se uštede.

Detaljna analiza velikih podataka omogućiće bolje razumevanje zdravstvenog stanja nacije i predviđanje budućih troškova lečenja. Kroz analizu podataka moguće je bolje razumeti zdravstveno stanje populacije i predvideti buduće troškove zdravstvene zaštite. Sprovođenje ovakvih analiza bi bilo značajno teže sa podacima koji su raspoređeni na više lokacija, ali jezero podataka omogućava centralizaciju i efikasnije upravljanje velikim i raznolikim skupom podataka, čime se olakšava sprovođenje ovakvih analiza.

Skladištenje podataka u jezeru za sektor zdravstva donosi koristi u vidu ponavljanja analiza nad istim podacima, a u cilju potvrđivanja dobijenih rezultata ili identifikovanja problema u dobijenim rezultatima. Dobar primer ovakve upotrebe podataka je analiziranje podataka uvođenja novih načina procene troškova lečenja pacijenata koji boluju od iste bolesti u slučaju javnog osiguranja. Konkretnije, Fond za zdravstveno osiguranje može koristiti jezero podataka za razvijanje različitih modela procene troškova, kombinujući kliničke i finansijske podatke. Ako postoji neslaganje između setova podataka, kao što je neusklađenost između identifikovanih pacijenata i dijagnoza u kliničkim istraživanjima i međunarodne klasifikacije bolesti, Fond može pristupiti originalnim podacima u jezeru da bi utvrdio uzroke tih neslaganja. Tako, jezero podataka omogućava detaljno ispitivanje i validaciju podataka, što je ključno za precizno donošenje odluka u zdravstvenom sektoru, posebno u kontekstu upravljanja troškovima i efikasnosti lečenja.

Optimizacija troškova zdravstvenog osiguranja, efikasno upravljanje resursima i reorganizacija sektora praćenjem opterećenja ustanova i planiranjem specijalizacija zdravstvenih radnika su još jedan vid koristi koje zdravstveni sektor može ostvariti. Analize prikupljenih podataka bi trebalo da imaju za cilj optimizaciju troškova zdravstvenog osiguranja, koja bi se postigla kroz bolju prevenciju i rano otkrivanje bolesti, aktivno praćenje grupa pacijenata koji su pod većim rizikom, optimizaciju broja kreveta u bolnicama i slično. Takođe, na osnovu prikupljenih podataka bi mogao da se reorganizuje sektor zdravstva prateći opterećenja različitih zdravstvenih ustanova kroz vreme, kao i da se planiraju potrebne specijalizacije i subspecijalizacije, odnosno da se bolje upravlja zdravstvenim radnicima.

Da bi građani bili spremni da podeli svoje podatke, mora postojati poverenje u instituciju koja upravlja i generiše zdravstvene podatke, ali i same politike privatnosti moraju detaljno i dobro biti razrađene. Politike privatnosti su od velike važnosti kada su u pitanju zdravstveni podaci. Stoga je bitno da institucija koja je zadužena za prikupljanje i upravljanje podacima u jezeru podataka bude od poverenja građanima, kako bi oni dali pristanak da se njihovi podaci koriste, jer kao što je već navedeno u regulatornom delu, osnov za prikupljanje podataka o ličnosti jeste pristanak. U Srbiji je već pomenutim Zakonom o zdravstvenoj dokumentaciji i evidencijama u oblasti zdravstva, po prvi put detaljno regulisano pitanje genetičkih podataka i omogućena je potpuna kontrola države u ovom segmentu, kao i poštovanje prava i obaveza u vezi sa ovim podacima.

Uzimajući u obzir koristi koje podaci od kliničkih ispitivanja mogu doneti celokupnom društvu, u ovom kontekstu je potrebno preispitati koncept pristanka i etička pitanja, kako bi bila uspostavljena ravnoteža između lične kontrole podataka i društvenih benefita. Međutim, u kontekstu zdravstva ne bi bilo loše preispitati koncept pristanka imajući u vidu da, na primer, podaci o kliničkim ispitivanjima i lečenjima nisu od koristi samo pacijentima nad kojima se ona vrše, već mogu koristiti i ostalima, odnosno da mogu doprineti efektivnijim terapijama za sve. Pitanje je kako očuvati privatnost pacijenata, a opet poštovati sve benefite za društvo. Trenutni koncept pristanka omogućava pojedincu da u potpunosti kontroliše deljenje ličnih zdravstvenih podataka. Ovaj koncept bi trebalo razmotriti tako da se uspostavi ravnoteža između lične kontrole i deljenja podataka koji mogu koristiti javnom

zdravlju. Kreatori politika u ovom kontekstu treba da razmotre potencijalnu vrednost upotrebe podataka i uporede je sa mogućim rizicima po privatnost pojedinaca, kao i da procene da li su trenutno važeća pravila za davanje pristanka dovoljno informativna građanima/pacijentima kako bi razumeli značenje datog pristanka ili nepristanka. U skladu sa tim, planirano je uspostavljanje tela koje će se baviti etičkim pitanjima i zahtevima za pristup podacima.

Iako je novi Zakon predvideo mogućnost deljenja podataka između zdravstvenih i naučnih institucija, zarad unapređivanja istraživanja trebalo bi razmotriti mogućnost izgradnje centralizovanog jezera podataka koje bi ujedinilo sve zdravstvene podatke pacijenata. Novi Zakon je predvideo mogućnost deljenja podataka između zdravstvenih institucija i naučnih instituta koji se bave istraživanjem u oblasti zdravstva. Ovo povezivanje je vrlo korisno za unapređenje istraživanja, jer jedna bolnica ili klinički centar ne može raspolagati svim potrebnim podacima za sprovođenje analiza. Bilo bi dobro urediti mogućnost da se podaci koje prikupe sve ustanove javnog zdravlja, odnosno tri klinička centra, bolnice, zavodi za javno zdravlje i domovi zdravlja, kao i naučne institucije slivaju u jedinstveno jezero podataka kojim bi upravljao Institut za javno zdravlje „Batut“ ili Republički fond za zdravstveno osiguranje, a potom uključiti i privatne bolnice i zdravstvene centre, imajući na umu da pacijenti u Srbiji često kombinuju privatno i javno zdravlje. Na ovaj način bi se stekao potpuni uvid u zdravstveno stanje pacijenta i kreirao njegov jedinstveni digitalni zdravstveni dosije. Međutim, za to je potrebno dobiti inpute iz privatne prakse. Nekada pacijenti nisu otvoreni ili iskreni da sve podatke podele sa lekarom, te dobijanje uvida u podatke iz drugih zdravstvenih ustanova može lekarima obezbediti bolji uvid u zdravstveno stanje pacijenta. Ovo je jedan od segmenata koji bi bilo dobro urediti, odnosno uspostaviti fazno deljenje podataka između javnog i privatnog zdravstva.



**Izvor: Autor*

Benefiti koje bi srpski zdravstveni sistem imao od uvođenja jezera podataka kreću se od obezbeđivanja boljeg pristupa zdravstvenim podacima za pacijente, preko boljeg upravljanja finansijskim, materijalnim i ljudskim resursima u sektoru, do unapređenja preventivne zdravstvene zaštite, koja je trenutno na niskom nivou.

Primer upotrebe 2 – Jezero podataka u Kragujevcu

Kragujevac se posebno ističe među jedinicama lokalne samouprave u Srbiji prema broju razvijenih digitalnih usluga. Od jedinica lokalne samouprave se očekuje da što više prilagode svoje usluge potrebama građana koji u njima žive i da odluke donose po participativnom principu, što bi moglo da se postigne uz pomoć digitalnih tehnologija. Za razliku od velikog broja lokalnih samouprava u zemlji koje su organizovane tako da njihova organizaciona struktura odgovara administraciji umesto građanima, Kragujevac je poslednjih godina razvio nekoliko pametnih usluga koje olakšavaju život građanima³⁸.

Javni prevoz u Kragujevcu je unapređen implementacijom Google Transit sistema, kao i praćenjem broja putnika u prevozu i vozila u realnom vremenu. Jedna od pametnih usluga koja je razvijena jeste sistem Google Transit koji podrazumeva unos podatka o svim linijama javnog prevoza na ovaj besplatni sistem. Međutim, osim na Google Transit, Kragujevac je iste podatke postavio i na nacionalni portal otvorenih podataka. Ovi podaci su ukršteni sa podacima brojača putnika iz linija gradskog prevoza, što je putnicima dalo uvid u popunjenost vozila. Tako je putnicima data mogućnost da odluče o ulasku u prevoz u zavisnosti od njegove popunjenosti. Kragujevac je razvio i sistem praćenja vozila u realnom vremenu, a koji kontroliše JP „Šumadija“. Pomoću ovog sistema se kontroliše da li svi prevoznici poštuju red vožnje i evidentiraju kvarove na mreži u cilju brže reakcije i smanjenja neprijatnosti za korisnike usluge.

Pomoću sistema eGrađanin, Kragujevčani mogu prijaviti probleme u funkcionisanju komunalnih usluga, pratiti potrošnju vode i davati predloge za unapređenje usluga. Građanima ove lokalne samouprave je dostupna usluga eGrađanin koju mogu koristiti za prijavu problema, slanje upita i predloga organima jedinice lokalne samouprave ili javnim preduzećima. Na primer, pomoću aplikacije je moguće prijaviti kvar na vodovodnoj mreži. Pored toga, pokrenut je projekat instalacije pametnih vodomera pomoću kojih se automatski učitava potrošnja vode korisnika i šalje automatski sistemu za izradu računa.

Saobraćaj u gradu će dodatno biti unapređen pomoću sistema snimanja i 3D vizuelizacije. Još jedan projekat koji Kragujevac sprovodi je snimanje i vizuelizacija saobraćajnica u 3D formatu, koje se radi u referentnom sistemu Republičkog geodetskog zavoda, a materijal je integrisan na jedinstvenu platformu. Predviđeno je da se podaci prikupljeni na ovaj način koriste za projektovanje, uvid i praćenje promena na saobraćajnicama, kao i formiranje katastra saobraćajne signalizacije.

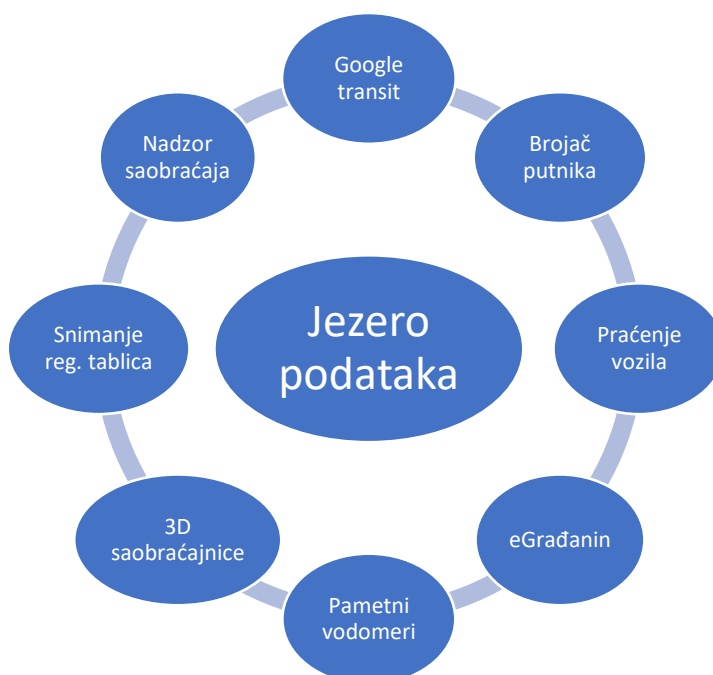
Uvođenje video nadzora će pomoći u sprečavanju rada nelegalnih taksi vozila i unapređenju bezbednosti. Grad trenutno radi na implementaciji sistema video nadzora saobraćaja, a u cilju sprečavanja rada nelegalnih taksi vozila i poboljšanja bezbednosti na saobraćajnicama. Sprečavanje rada nelegalnih taksi vozila treba da se postigne snimanjem registarskih tablica i njihovim optičkim prepoznavanjem, a potom će biti vršena provera da li je snimljena tablica registrovana u bazi registrovanih taksi vozila. Ukoliko nije, automatski će se kreirati prekršajni nalog.

Sledeći korak u razvoju pametnog grada bi trebalo da bude planiranje i izgradnja jedinstvenog skladišta podataka, poput jezera podataka, a koje bi pojednostavilo i unapredilo analize i dalo bolje inpute za razvoj novih pametnih usluga u cilju unapređenja života građana. Primetno je i pohvalno što je Kragujevac razvio različite pametne usluge. Sledeći korak, osim razvoja novih usluga, trebalo bi da bude povezivanje podataka koje generišu predstavljeni pametni servisi. Jezero podataka omogućava integraciju različitih izvora podataka na jednom mestu, a koristi koje bi grad Kragujevac mogao da ima

³⁸ <https://eid.rs/lat/pametn-grad.html>

od jezera podataka svode se na unapređenje usluga koje gradska uprava pruža, a koje direktno ili indirektno utiču na kvalitet života građana. Kako bi jezero podataka bilo implementirano i kako bi se podaci iz različitih servisa slivali u jedno centralizovano skladište podataka, potrebno je da uprava jedinice lokalne samouprave preduzme korake koji su već u prethodnom tekstu analize detaljno objašnjeni, a ovde će biti samo nabrojani:

1. definisati cilj i svrhu izgradnje jezera podataka,
2. analizirati postojeće procedure za generisanje podatka u različitim sistemima,
3. izraditi detaljan plan izgradnje jezera podataka, standardizovati procedure za skladištenje i upravljanje podacima i izraditi katalog podataka,
4. obezbediti potrebne resurse u vidu infrastrukture, materijalnih, ljudskih i finansijskih resursa,
5. uvoz podataka iz različitih sistema i testiranje funkcionalnosti jezera podataka,
6. puštanje u rad jezera podataka uz stalno praćenje i unapređenje rada.



**Izvor: Autor*

Usluge koje bi mogle da budu unapređene izgradnjom jezera podataka u Kragujevcu su:

- efikasniji javni prevoz,
- unapređenje saobraćaja kroz informisanje građana o radovima i unapređenje bezbednosti,
- unapređeno urbano planiranje (izgradnja vrtića, škola, bolnica, ...),
- ekonomsko planiranje zasnovano na kvalitetnijim analizama i bolja iskorišćenost resursa,
- promptnije reakcije na prirodne nepogode ili hitne slučajeve.

Integracija podataka različitih gradskih službi pruža osnov za bolje analize i informisanije donošenje odluka. To znači da bi se podaci iz različitih gradskih službi, kao što su transport, komunalne usluge, hitne službe, i urbanističko planiranje, mogli efikasnije analizirati i koristiti za donošenje odluka. Integracijom podataka, grad može da poboljša pružanje usluga građanima. Na primer, kombinovanjem podataka o saobraćaju, radovima na putu i stanju javnog prevoza, grad može da planira i obaveštava građane o alternativnim rutama tokom saobraćajnih zastoja ili radova na putu.

Povezivanje postojećih sistema dodatno bi doprinelo unapređenju funkcionisanja saobraćaja kroz efikasnije upravljanje saobraćajem i pravovremeno informisanje građana o promenama. Povezivanje podataka sa sistema eGrađanin o kvarovima na vodovodnoj mreži sa podacima o stanju saobraćajnica omogućili bi predviđanje i najave zatvaranja određene saobraćajnice usled kvara i obaveštavanje građana putem aplikacije o tome. Na taj način bi se uticalo na smanjenje gužvi u prevozu. Međutim, da bi se ovo realizovalo, potrebno da je da se podaci iz različitih servisa povežu u jednom silosu jezera podataka. Kasnije bi mogao da im se doda silos podataka koji prati gradski prevoz, kako bi mogle da se planiraju izmene linija gradskog prevoza u zavisnosti od stanja saobraćajnica i takođe, da se građani o tome obaveštavaju putem aplikacije eGrađanin.

Još jedan primer kako jezero podataka može da se koristi je unapređenje saobraćajne bezbednosti u gradu. U ove svrhe, bilo bi dobro povezati podatke iz sistema video nadzora, potom uključiti podatke kojima policija raspolaže i podatke iz saobraćaja, kako bi se identifikovale kritične saobraćajne zone. Nakon identifikacije, trebalo bi raditi na uspostavljanju preventivnih mera, poput bolje signalizacije ili osvetljenja tih delova saobraćajnica.

Podaci iz jezera se mogu koristiti za dalje planiranje urbanog razvoja i ekonomskog rasta. U domenu urbanističkog razvoja, to može značiti da se kroz razvoj grada, odnosno praćenje izgradnje novih stambenih objekata može planirati izgradnja predškolskih ustanova, škola, domova zdravlja, kao i parkova. Potom, ovi podaci mogu biti značajni input za planiranje reorganizacije saobraćaja, uvođenje novih linija gradskog prevoza, proširivanje saobraćajnica ili izgradnju novih javnih parkinga.

Jezera podataka mogu doprineti unapređenju procesa izrade lokalnog ekonomskog plana. U pogledu ekonomskog rasta, mogu se pratiti i analizirati podaci o lokalnom zapošljavanju, potrošnji, poslovanju preduzeća, pa čak i efekti stranih direktnih investicija na razvoj lokalne zajednice. Ovo može pružiti uvid u potencijale za generisanje budućeg lokalnog ekonomskog rasta, ali i inpute za lokalne ekonomske planove.

Podaci iz jezera podataka se takođe, mogu koristiti kako bi se unapredio lokalni turizam. Mogu se pratiti posete turista, identifikovati mesta koja najviše posećuju, obnoviti ta mesta, ali i osmisliti turističke ture koje bi uključile i neka nova mesta kako bi se podstakao lokalni turizam.

U situacijama prirodnih nepogoda ili hitnih okolnosti, jezero podataka može predstavljati neprocenjiv resurs za brzo i efikasno reagovanje. Integracijom ključnih informacija na jednom mestu, omogućila bi se trenutna razmena veoma važnih podataka između hitnih službi kao što su vatrogasci, policija i hitna pomoć. Na primer, u slučaju poplava, jezero podataka može obezbediti trenutne informacije o nivoima vode, lokacijama koje su najviše ugrožene, kao i o raspoloživim putevima za evakuaciju. Ovo omogućava hitnim službama da brzo razumeju obim situacije, koordiniraju napore za spasavanje i pružanje pomoći, kao i da optimizuju svoje resurse i najviše ih usmere na mesta gde su najpotrebniji. Sposobnost brzog pristupa i analize relevantnih podataka u realnom vremenu omogućava ne samo efikasnije upravljanje kriznim situacijama, već i potencijalno spašavanje života, smanjujući vreme odziva i poboljšavajući koordinaciju među ključnim službama.

Kragujevac raspolaže kapacitetima Državnog centra za upravljanje i čuvanje podataka, gde hardver potreban za jezero podataka može biti smešten, ali potrebno je investirati sredstva u ljudske resurse koji bi upravljali jezerom podataka. Navedeni su samo neki od predloga kako bi grad Kragujevac mogao da ima koristi od uvođenja jezera podataka. Daljim razvojem usluga, mogli bi da se dodaju novi silosi podataka u jezero podatka koji bi se povezivali sa prethodnim. Olakšavajuća okolnost za Kragujevac jeste činjenica da raspolaže kapacitetima Državnog centra za upravljanje i čuvanje podataka, gde postoji prostor za razvoj jezera podataka. Međutim, tu je prisutan i izazov u vidu nedostajućeg kadra u

svakoj od organizacija javnog sektora koja bi trebalo da koristi jezero podataka u svrhe sprovođenja analiza, vizuelizacije nalaza i predstavljanja tih nalaza donosiocima odluka.

Koristi koje bi građani Kragujevca imali od izgradnje jezera podataka koje bi bilo u nadležnosti jedinice lokalne samouprave počinju sa mogućnostima pružanja boljih usluga koje obezbeđuje lokalna samouprava građanima. Potom, raspoloživi resursi bi bili optimizovani. Saobraćaj bi bio unapređen smanjenjem gužvi i većom bezbednošću. Takođe, transparentnost, kao bitan element kontrole rada javnih službi bi bila veća.

Zaključak

Analiza jezera podataka ukazuje na njihov ključni značaj u savremenom informacionom društvu, gde podaci postaju sve važniji resurs za donošenje odluka u različitim sektorima, od državnih institucija do privatnih preduzeća. Jezera podataka se ističu po svojoj sposobnosti da skladište ogromne količine različitih tipova podataka, od strukturisanih do nestrukturisanih, u njihovom izvornom formatu. Ova fleksibilnost omogućava brzu i efikasnu obradu podataka, što je ključno za agilnost i inovativnost organizacija. Posebno se ističe njihova troškovna efikasnost, jer se mogu graditi na postojećoj infrastrukturi i koristeći softver otvorenog koda, što smanjuje troškove. Međutim, izazovi poput obezbeđenja kvaliteta podataka i upravljanja velikom količinom podataka zahtevaju pažljivo planiranje i implementaciju. U kontekstu javnog sektora, postoji potencijal za značajno poboljšanje analiza i donošenja odluka kroz efikasnije korišćenje podataka, što može dovesti do bolje organizacije i pružanja javnih usluga građanima.

Iako izgradnja i upravljanje jezerom podataka ne zahteva donošenje posebnih propisa, važno je poštovati postojeći Zakon o zaštiti podataka o ličnosti. Ovaj Zakon postavlja stroge zahteve u pogledu zakonitosti, transparentnosti, ograničenja čuvanja, tačnosti, integriteta i poverljivosti podataka. U skladu sa Zakonom o zaštiti podataka o ličnosti, svaki vlasnik jezera podataka preuzima ulogu rukovaoca ili obrađivača podataka, što sa sobom nosi obavezu imenovanja lica zaduženog za zaštitu podataka o ličnosti. Ova regulativa osigurava da se podaci u jezerima koriste na odgovoran i etičan način, vodeći računa o privatnosti pojedinaca. Pored toga, Zakon o informacionoj bezbednosti dodatno naglašava značaj sveobuhvatne zaštite, upravljanja rizicima, stručnosti i svesti o bezbednosnim aspektima u IKT sistemima.

Upravljanje jezerima podataka je složen proces koji zahteva precizno definisane ciljeve i efikasnu implementaciju kako bi se izbeglo stvaranje "močvare podataka", gde bi akumulirani podaci bili neorganizovani i neupotrebljivi. Ključ uspešnog jezera podataka leži u jasnoj viziji problema koji treba rešiti i specifičnim primerima upotrebe podataka. Kroz jezera podataka, organizacije mogu efikasno integrisati različite vrste podataka iz više izvora, što omogućava bolje informisano donošenje odluka i optimizaciju procesa. Jezera podataka doprinose unapređenju efikasnosti i transparentnosti upotrebe podataka, što im omogućava da postanu ključni alat za unapređenje poslovanja javnog sektora. Osim toga, omogućuju razvoj novih usluga i proizvoda zasnovanih na analizi podataka, doprinoseći ekonomskom rastu i društvenom blagostanju.

Primeri iz Evropske unije pokazuju da je integracija jezera podataka u javni sektor postala neizbežna potreba zbog rastuće količine podataka i potrebe za efikasnim upravljanjem podacima. Švajcarska, Finska, Francuska, Italija i Estonija su svedoci ovog trenda, gde svaka zemlja pristupa implementaciji jezera podataka na svoj način, u skladu sa specifičnim potrebama i ciljevima. Jezera podataka se pokazuju kao ključna komponenta u digitalnoj transformaciji javnog sektora, ne samo za unapređenje operativne efikasnosti, već i za omogućavanje informisanijeg donošenja odluka i pružanja kvalitetnijih usluga građanima.

Primena jezera podataka u javnom sektoru predstavlja transformacioni korak ka boljem upravljanju složenim i obimnim skupovima podataka. Državni funkcioneri, uključujući ministre i direktore institucija i ustanova, predstavljaju glavne zainteresovane strane koje mogu ostvariti koristi od složenih analiza zasnovanih na podacima iz jezera, a koje bi koristili za strateško donošenje odluka. Ove analize su od ključnog značaja za razvoj i unapređenje javnih usluga i politika koje direktno utiču na život građana. Pored toga, kontrolori transparentnosti i ispravnosti rada javnog sektora, kao što su Državna revizorska institucija i nevladine organizacije, takođe mogu imati koristi od podataka iz jezera. Oni mogu koristiti ove podatke za praćenje i ocenjivanje efikasnosti i efektivnosti rada javnog sektora,

obezbeđujući pritom da se Zakon o zaštiti podataka o ličnosti poštuje. Potom, akademska i istraživačka zajednica mogu imati značajnu ulogu, jer kroz partnerstvo sa javnim, privatnim i nevladinim sektorom mogu doprineti razvoju novih usluga i poboljšanju socijalnih, ekonomskih ili ekoloških politika.

Uvođenje jezera podataka u javni sektor Srbije predstavlja značajan izazov, zahtevajući migraciju podataka iz različitih sistema i formata u novu arhitekturu. Korišćenje jezera podataka može omogućiti pristup istorijskim podacima i njihovo korišćenje u analitičke svrhe, što može biti ključno za efikasno upravljanje podacima i donošenje odluka u javnom sektoru. Jedan od ključnih koraka u procesu implementacije jezera podataka u Srbiji jeste sistematizacija podataka i izrada kataloga podataka koji bi uključivao sve podatke koje javni sektor generiše. Ovaj proces bi omogućilo bolje upravljanje i maksimalno iskorišćavanje podataka. Takođe, potrebno je uspostaviti standarde u pogledu formata i kvaliteta podataka kako bi se omogućilo njihovo efikasno deljenje i analiza.

Ljudski resursi predstavljaju drugi izazov u procesu implementacije jezera podataka. Neophodno je ulagati u razvoj kadrova koji bi mogli efikasno da upravljaju ovim sistemima, kao i osmisлити načine za privlačenje stručnjaka iz oblasti informaciono-komunikacionih tehnologija da rade u javnom sektoru. Uz to, važno je da se rukovodioci informišu o koristima digitalizacije i analize podataka kako bi se u potpunosti iskoristile mogućnosti koje pružaju savremene tehnologije. U tehničkom smislu, implementacija jezera podataka u Srbiji ne bi trebalo da predstavlja značajan izazov, imajući u vidu postojanje Državnog centra za upravljanje i čuvanje podataka. Dva predloga primene jezera podataka u javnom sektoru Srbije detaljnije pokazuju mogućnosti i koristi od implementacije jezera podataka.

Implementacija jezera podataka u sektoru zdravstva u Srbiji donosi značajne benefite, uključujući poboljšanu integraciju i upravljanje velikim količinama zdravstvenih podataka iz različitih izvora. Ovo bi omogućilo efikasnije donošenje odluka, unapređenje preventivne medicine i personalizovane terapije zasnovane na detaljnijoj analizi podataka pacijenata. Jezero podataka bi takođe, doprinelo boljem upravljanju resursima u sektoru zdravstva, kao i unapređenju kliničkih ispitivanja i istraživanja. Za uspešnu implementaciju, ključno je formirati multidisciplinarni tim stručnjaka iz oblasti IT-a i zdravstva, definisati jasne protokole za prikupljanje i obradu podataka, te uspostaviti stroge mere zaštite privatnosti i sigurnosti podataka. Takođe, važno je obrazovati zdravstvene radnike o novim tehnologijama, kao i razmotriti pravne i etičke aspekte upotrebe zdravstvenih podataka.

Sa druge strane, implementacija jezera podataka u gradu Kragujevcu može značajno unaprediti pružanje lokalnih usluga, optimizovati upravljanje resursima i povećati transparentnost u radu javnih službi. Sa integracijom podataka iz različitih pametnih servisa, kao što su sistem javnog prevoza, upravljanje saobraćajem, eGrađanin za prijavu problema i praćenje rada komunalnih službi, jezero podataka bi omogućilo efikasniju analizu i bolje planiranje urbanih usluga. Pored toga, jezero podataka bi omogućilo optimizaciju ekonomskog planiranja. U slučaju prirodnih nepogoda ili hitnih situacija, integracija ključnih podataka bi doprinela efikasnijem reagovanju hitnih službi, čime se bi se povećala sigurnost građana. Kragujevac, sa svojom infrastrukturom i resursima, ima potencijal da postane primer dobre prakse u primeni jezera podataka za unapređenje lokalnih usluga i kvaliteta života građana.

Na kraju, može se zaključiti da implementacija jezera podataka u Srbiji predstavlja ključan korak ka unapređenju analitike i donošenja odluka na osnovu velikih podataka u javnom sektoru, doprinoseći efikasnosti, transparentnosti i boljem pružanju usluga građanima. Razvoj jezera podataka može značajno poboljšati upravljanje resursima, optimizovati operacije i pružiti osnovu za inovativne, personalizovane i preventivne pristupe u uslugama. Ali, nikako se ne sme zaboraviti da gradnja jezera podataka u Srbiji zahteva strateško planiranje, ulaganje u tehničku infrastrukturu i ljudske resurse.

Rečnik

ACID (Atomicity, Consistency, Isolation, Durability) – Atomiranost, doslednost, izolovanost, trajnost

BI alati (Business Intelligence tools) – Alati za poslovno istraživanje

DBMS sistemi (Database Management Systems) – Sistemi za upravljanje bazama podataka

ETL procesi (Extract, Transformation, and Load Processes) – Procesi ekstrakcije, transformacije i učitavanja

IoT (Internet of Things) – Internet stvari

OLTP transakcije (Online Transaction Processing) – Obrada transakcija u realnom vremenu

SQL database (Structured Query Language) – Strukturirani jezik upita

