

# Sistem za identifikaciju minerala i njihovih nalazišta u Srbiji

## Član tima

- Branko Marić SV70-2021

## Motivacija

Srbija poseduje bogato, ali široj javnosti nedovoljno poznato mineralno nasleđe, uključujući nalazišta raznovrsnog poludragog kamenja. Interesovanje za amatersku gemologiju, sakupljanje minerala i planinarenje sa ciljem pronalaska uzoraka je u porastu. Međutim, informacije o vrstama minerala, njihovim karakteristikama i tačnim lokacijama nalazišta su rasute, često se nalaze u stručnoj literaturi, na forumima ili su deo usmenog predanja, što otežava ulazak u hobi novim entuzijastima.

Motivacija za projekat je stvaranje centralizovanog, inteligentnog sistema koji bi korisnicima omogućio laku identifikaciju minerala na osnovu vizuelnih karakteristika, ali i pružio precizne informacije o potencijalnim nalazištima, geološkom kontekstu i srodnim primercima. Ovaj sistem bi, za razliku od statičnih resursa, koristio **ekspertski sistem** i **mehanizme rezonovanja** kako bi pružio personalizovano i pouzdano iskustvo, rangirajući potencijalne kandidate i dajući preporuke za dalje istraživanje.

## Pregled problema

Specifičan problem koji se rešava je nedostatak jedinstvenog, interaktivnog alata za podršku amaterima i hobistima u potrazi za dragim i poludragim kamenjem na teritoriji Srbije.

### Postojeća rešenja se uglavnom svode na:

- **Statične veb-sajtove i blogove:** Pružaju opise minerala, ali bez interaktivne pomoći pri identifikaciji ili dinamičkog pretraživanja lokacija.
- **Geološke karte i stručna literatura:** Precizne su, ali su kompleksne za korišćenje laicima, zahtevaju stručno znanje za tumačenje i nisu lako dostupne.
- **Forumi i društvene mreže:** Informacije su nepouzdanе, subjektivne i neorganizovane.

### Nedostaci postojećih rešenja:

- Ne postoji sistem koji na osnovu vizuelnih parametara može da izvrši rezonovanje i ponudi listu mogućih minerala.
- Ne postoji alat koji povezuje identifikovani mineral sa poznatim nalazištima u Srbiji na inteligentan način.
- Ne postoji sistem koji bi upozoravao korisnike na zaštićena područja.

**Prednost ovog rešenja** bi bila u implementaciji baze znanja i ekspertskog sistema koji bi korisniku pružio personalizovano iskustvo. Sistem bi pomagao korisniku da identifikuje kamen koji je pronašao na osnovu unetih karakteristika, rangirajući kandidate, i usmeravao ga ka lokacijama gde bi mogao pronaći slične primerke, uzimajući u obzir i zakonska ograničenja.

## Metodologija rada

### Očekivani ulazi u sistem (Input)

Korisnik će unositi podatke o pronađenom uzorku kamena kroz interaktivan upitnik. Pored osnovnih vizuelnih karakteristika, dodajemo i opcione, naprednije ulaze koji omogućavaju kompleksnije rezonovanje:

- **Lokacija pronalaska:** Korisnik unosi lokaciju putem teksta ili odabirom na mapi.
- **Vizuelne karakteristike:**
  - **Boja:** Odabir jedne ili više boja sa palete.
  - **Prozirnost:** Odabir opcije (neproziran, poluproziran, proziran).
  - **Sjaj:** Odabir opisa sjaja (staklast, sedefast, mastan, metalan, bez sjaja).
- **Opcione napredne karakteristike:**
  - **Tip stene pronalaska:** Odabir tipa stene (npr. vulkanska, sedimentna, metamorfna) ili konkretnog naziva (serpentinit, granit, krečnjak).

### Očekivani izlazi iz sistema (Output)

Na osnovu unetih podataka, sistem će generisati sledeće informacije, rangirane po verovatnoći:

- **Lista mogućih minerala rangirana po bodovima:** Lista kandidata, npr. "Opal (95 poena)", "Kalcedon (80 poena)".
- **Detaljan opis identifikovanog minerala:** Karakteristike, geološki kontekst, tipovi stena u kojima se nalazi, zanimljivosti.
- **Spisak ili mapa nalazišta u Srbiji:** Prikaz poznatih lokacija gde se identifikovani mineral može pronaći, sa geografskim i geološkim područjem (npr. "Fruška gora - Neogeni vulkanski kompleks").
- **Preporuke za srodne minerale:** Predlozi drugih minerala ili varijeteta koji se mogu naći na istoj lokaciji ili u sličnom geološkom kontekstu.
- **Saveti:**

- Informacije u kom periodu je prijavljeno najviše pronalazaka

## Baza znanja projekta

Baza znanja će biti srce sistema i sastojaće se iz dva dela:

1. **Baza činjenica (Facts):** Sadržaće formalno opisane podatke o mineralima, geološkim oblastima i lokacijama u Srbiji. Svaki mineral će biti objekat sa atributima:
  - naziv
  - boja (lista mogućih boja)
  - providnost (lista opcija)
  - sjaj (lista opcija)
  - tvrdoća\_min, tvrdoća\_max (Mohsova skala)
  - boja\_ogreba
  - tip\_stene (lista tipova)
  - nalazista (lista lokacija)
2. **Baza pravila (Rules):** Sadržaće ekspertska pravila za rezonovanje, napisana u **Drools (DRL)** formatu. Pravila će služiti za bodovanje, filtriranje, klasifikaciju lokacija i generisanje preporuka i upozorenja na osnovu korisničkog unosa.

# Primeri kompleksnih pravila i korišćenje baze znanja

## Forward-Chaining

**Scenario:** Korisnik je na Zlatiboru, u steni zelenkaste boje (serpentinit), pronašao kamen sledećih karakteristika:

- **Lokacija:** Zlatibor
- **Boja:** Zelena
- **Prozirnost:** Poluproziran
- **Sjaj:** Mastan
- **Tip stene pronalaska (opciono):** Serpentinit

**Korak 1: Učitavanje i Kategorizacija Kandidata:** Sistem prima korisnički unos i kreira objekat Uzorak. Zatim, učitava **sve** minerale iz baze znanja u radnu memoriju, ali ih odmah logički deli u dve grupe na osnovu lokacije.

- **Činjenica:** Uzorak(lokacija="Zlatibor", ...)
- **Akcija:** Sistem kreira dve liste kandidata:
  1. PrimarniKandidati: Svi minerali iz baze za koje je Zlatibor naveden kao poznato nalazište. (npr. Serpentin, Magnezit, Hromit, Epidot...).
  2. SekundarniKandidati: Svi ostali minerali iz baze. (npr. Ametist sa Fruške gore, Malahit sa Rudnika, itd.).
- Svi kandidati u obe liste imaju početni skor 0.

**Korak 2: Osnovno Bodovanje za Sve Kandidate:** Sistem primenjuje osnovna pravila za poređenje vizuelnih karakteristika na obe liste podjednako.

- **Pravilo "MatchColor", "MatchLuster", "MatchTransparency":** Pravila se izvršavaju za sve kandidate.
- **Rezultat nakon ovog koraka (primer):**
  - U listi PrimarniKandidati: Serpentin (25 poena), Epidot (25 poena).
  - U listi SekundarniKandidati: Pretpostavimo da postoji neki drugi, vizuelno sličan zelenkasti mineral, npr. Žadeit (koji nije na Zlatiboru), takođe bi dobio oko 25 poena.

**Korak 3: Bonus za lokaciju i geološki kontekst:** Uvode se dva pravila koja daju značajne bodove mineralima u odnosu na lokaciju na kojoj su pronađeni i za geološki kontekst.

- **Pravilo "LocationPriorityBonus":** AKO Kandidat pripada listi PrimarniKandidati, dodeli **značajan bonus za geolokacijsko poklapanje**.
  - **Akcija:** Svi kandidati iz prve liste (Serpentin, Epidot...) dobijaju npr. **+30 poena**. Kandidati iz druge liste ne dobijaju ove poene.
- **Pravilo "ConfirmGeologicalContext":** Ako Kandidat.tip\_stene sadrži Uzorak.tip\_stene, dodeli bonus za geološki kontekst.
  - **Akcija:** Serpentin dobija dodatnih **+20 poena** jer se vezuje za serpentinit.
- **Stanje poena nakon ovog koraka:**
  - Serpentin: 25 (vizuelno) + 30 (lokacija) + 20 (stena) = **75 poena**.
  - Epidot: 25 (vizuelno) + 30 (lokacija) = **55 poena**.
  - Žadeit (sekundarni kandidat): 25 (vizuelno) + 0 (lokacija) = **25 poena**.

**Korak 4: Analiza Rezultata i Generisanje Pametnog Izlaza** Sistem sada analizira finalne skorove i donosi zaključak, ali je sada sposoban da prepozna i anomalije.

- **Scenario A (Standardna identifikacija):**
  - Pošto je Serpentin, primarni kandidat, ubedljivo na prvom mestu, sistem generiše standardni izlaz.
  - **Izlaz:** "Najverovatniji mineral je **Serpentin** (skor: 75/100). Identifikacija je potkrepljena poklapanjem vizuelnih karakteristika, tipa stene i poznatog nalazišta na Zlatiboru."
- **Scenario B (Potencijalno novo otkriće):**
  - Zamislimo da je korisnik uneo savršene karakteristike za Malahit (boja, sjaj, prozirnost...), ali je kao lokaciju naveo Zlatibor, gde Malahit nije dokumentovan.
  - Malahit bi kao SekundarniKandidat dobio izuzetno visok skor na osnovu vizuelnih karakteristika (npr. 45/50), ali ne bi dobio bonus za lokaciju. Njegov konačan skor bi bio 45 što je visok skor kad uzmemo u obzir da nije na dobroj lokaciji i sistem će ovo prepoznati kao anomaliju.
  - **Izlaz (u ovom slučaju):** „Vaš uzorak se po karakteristikama savršeno poklapa sa mineralom **Malahit**, koji do sada nije bio dokumentovan na području Zlatibora. Ovo bi moglo predstavljati novo nalazište. Vaš unos je zabeležen i biće prosleđen administratorima na dalju analizu.”

## CEP (Obrada kompleksnih događaja)

Ova tehnologija omogućava sistemu da prepozna značajne obrasce u realnom vremenu, analizirajući prijave više korisnika.

### Pravilo: "DetectUndocumentedHotspot" (Detekcija nedokumentovanog nalazišta)

- **Svrha:** Automatsko prepoznavanje klastera prijave korisnika koji sa velikom verovatnoćom ukazuju na postojanje novog, do sada neevidentiranog nalazišta (hotspota) određenog minerala.
- **Logika procesa:**
  1. **Definisanje događaja (Event):** Sistem generiše specifičan događaj, **PotencijalnoOtkriceEvent**, svaki put kada se u Forward-Chaining procesu desi "Scenario B" (visoka podudarnost na neočekivanoj lokaciji).
    - **Okidač:** Događaj se kreira kada je najbolje rangirani mineral SekundarniKandidat (nije sa poznatog nalazišta), ali je njegov skor poklapanja sa vizuelnim i drugim karakteristikama izuzetno visok.
    - **Sadržaj događaja:** Svaki PotencijalnoOtkriceEvent nosi ključne podatke: mineralID, korisnikID, lokacija\_GPS, skor\_poklapanja.
  2. **Definisanje CEP pravila (Rule):** U sistemu postoji pravilo koje neprekidno analizira tok PotencijalnoOtkriceEvent događaja.
    - **Naziv pravila:** "DetectUndocumentedHotspot"
    - **Logika pravila (opisno):** **KADA** se u vremenskom prozoru od **45 dana** akumulira najmanje **5** PotencijalnoOtkriceEvent događaja, pri čemu Svi događaji moraju da zadovolje sledeće uslove:
      - **Grupisanje po mineralu:** Svi događaji se moraju odnositi na isti mineral (npr. Ametist).
      - **Diverzitet izvora:** Događaji moraju poticati od najmanje **3 različita korisnika**.
      - **Geografska blizina:** Lokacije prijavljene u događajima moraju se nalaziti unutar radijusa od **5 kilometara** jedna od druge.
  3. **Akcija (Action):** Ukoliko su svi uslovi ispunjeni, sistem automatski izvršava sledeće korake:
    - **Kreiranje upozorenja:** Generiše novi objekat **HotspotUpozorenje** u sistemu.
    - **Sadržaj upozorenja:** Objekat sadrži agregirane podatke:

- Naziv minerala.
- Prosečne koordinate centra hotspot-a (izračunate na osnovu svih lokacija).
- Ukupan broj prijava.
- Lista ID-jeva korisnika koji su prijavili nalaze.
- Prosečna pouzdanost (skor) poklapanja.
- **Notifikacija:** Šalje obaveštenje administratorima sistema na dalju verifikaciju i potencijalno ažuriranje baze znanja sa novim, potvrđenim nalazištem.

#### **Pravilo: "DetectSeasonalPattern" (Otkrivanje sezonskih obrazaca pronalazaka)**

- **Svrha:** Pružanje korisnih saveta korisnicima o najboljem vremenu za potragu za određenim mineralima.
- **Logika procesa:**
  1. **Definisanje događaja (Event):** Svaka uspešna identifikacija je FindingEvent.
  2. **Definisanje CEP pravila (Rule):**
    - **Naziv pravila:** "IdentifySeasonalFindingSpike"
    - **Logika pravila (opisno):** Ovo je pravilo koje analizira podatke dugoročno. Kada je prosečan mesečni broj FindingEvent događaja za određeni mineral (npr. Opal) u specifičnom geološkom području (npr. Neogeni vulkanski kompleksi)
      - tokom određenog perioda (npr. septembar-oktobar),
      - za više od 250% veći od proseka istog minerala u istom području tokom ostatka godine,
      - i ovaj obrazac se ponovi najmanje dve uzastopne godine.
  3. **Akcija (Action):** Sistem automatski dodaje SeasonalTip (sezonski savet) u opis minerala i/ili lokacije. Na primer: *"Podaci pokazuju da korisnici najviše uspeha u pronalaženju Opala na Fruškoj gori imaju u septembru, verovatno usled smanjene vegetacije i erozije izazvane letnjim kišama."*

#### **1. Pravilo: "DetectCorrectionPattern" (Otkrivanje obrasca za korekciju podataka)**

- **Svrha:** Automatsko identifikovanje potencijalno netačnih ili zastarelih informacija u bazi znanja na osnovu konzistentnih prijava korisnika sa terena.

- **Logika procesa:**

1. **Definisanje događaja (Event):** Sistem generiše događaj PotentialCorrectionEvent svaki put kada korisnik unese podatke za poznati mineral na poznatoj lokaciji, ali se jedna od opcionih karakteristika (npr. tip stene, boja ogreba) ne poklapa sa onom u bazi činjenica.
2. **Definisanje CEP pravila (Rule):**
  - **Naziv pravila:** "SuggestFactUpdate"
  - **Logika pravila (opisno):** Kada se u vremenskom prozoru od 90 dana akumulira najmanje 10 PotentialCorrectionEvent događaja, pri čemu SVI događaji moraju da zadovolje sledeće uslove:
    - Da se odnose na isti mineral (npr. Epidot).
    - Da se odnose na isto nalazište (npr. Kopaonik).
    - Da potiču od najmanje 5 različitih korisnika.
    - Da svi prijavljuju istu novu vrednost za isti atribut (npr. svi unose tip\_stene = 'skarn' dok u bazi stoji tip\_stene = 'amfibolit').
3. **Akcija (Action):** Sistem kreira objekat DataCorrectionAlert i šalje notifikaciju administratoru. Objekat sadrži informacije: mineral, nalazište, atribut\_za\_ispravku, stara\_vrednost, nova\_vrednost, broj\_prijava, ID\_korisnika. Time se omogućava da se baza znanja održava ažurnom uz pomoć kolektivne inteligencije korisnika.

## Backward-Chaining: Pretraga Minerala po Hijerarhijskoj Pripadnosti

Ovaj mehanizam se aktivira kada korisnik želi da pronađe sve specifične tipove minerala koji pripadaju nekoj široj geološkoj grupi na određenoj lokaciji. Sistem tretira hijerarhijsku pripadnost kao vezu (npr. "Ametist je vrsta Kvarca, a Kvarc je vrsta Silikata") i koristi logiku backward chaining-a sa rekurzivnim query-jem bi pronašao sve članove date grupe.

### Logika Procesa

#### Korak 1: Definisanje Baze Znanja o Hijerarhiji

U sistemu postoje činjenice (MineralHierarchy) koje definišu direktnu vezu "pod-tip pripada nad-tipu". Ove činjenice predstavljaju osnovno znanje o klasifikaciji minerala.

- MineralHierarchy("Kvarc", "Silikati")
- MineralHierarchy("Serpentin", "Silikati")



## Korak 2: Korisnički Upit (Cilj)

Korisnik postavlja upit sistemu sa ciljem da pronađe sve minerale iz grupe "Silikati" na lokaciji "Fruška Gora".

- **Ulaz (Input):** mineralGroup = "Silikati", location = "Fruška gora"

## Korak 3: Interni Proces Sistema (Rekurzivni Backward Chaining)

Sistem pokreće pravilo čiji je zadatak da pronađe sve minerale (\$m) na datoj lokaciji (\$loc) za koje važi hipoteza isMemberOfGroup(\$m.getName(), \$group). Da bi dokazao ovu hipotezu, Drools endžin koristi rekurzivni query koji poziva sam sebe:

```
query isMemberOfGroup(String subType, String superType)
// Osnovni slučaj: Da li postoji direktna veza?
MineralHierarchy(subType, superType;)
or
// Rekurzivni korak: Da li postoji veza preko posrednika?
(
    MineralHierarchy(subType, $intermediate;) and
    isMemberOfGroup($intermediate, superType;) //
)
end
```

**Kako rekurzija radi u praksi za isMemberOfGroup("Amethyst (Quartz)", "Silikati"):**

1. **Prvi poziv:** isMemberOfGroup("Amethyst (Quartz)", "Silikati")
  - **Osnovni slučaj:** Ne postoji direktna činjenica MineralHierarchy("Amethyst (Quartz)", "Silikati"). Ne uspeva.
  - **Rekurzivni korak:**
    - Pronalazi posrednika: MineralHierarchy("Amethyst (Quartz)", \$intermediate="Kvarc"). Uspeva.
    - Sada mora da dokaže novi, manji cilj: **poziva sam sebe** sa isMemberOfGroup("Kvarc", "Silikati").
2. **Drugi poziv (rekurzivni):** isMemberOfGroup("Kvarc", "Silikati")
  - **Osnovni slučaj:** Pronalazi direktnu činjenicu MineralHierarchy("Kvarc", "Silikati"). **Uspeva!**
3. **Povratak iz rekurzije:** Pošto je drugi poziv uspeo, ceo and uslov u prvom pozivu je zadovoljen.
4. **Zaključak:** Hipoteza isMemberOfGroup("Amethyst (Quartz)", "Silikati") je dokazana kao **tačna**.

## Korak 4: Finalni Izlaz Sistema

Glavno pravilo sakuplja sve minerale sa lokacije "Fruška gora" za koje je rekurzivni query uspeo da dokaže da pripadaju grupi "Silikati".

- **Izlaz (Output) za dati primer:** Lista minerala koja sadrži Amethyst (Quartz) i Rock Crystal (Quartz).