Upravljanje digitalnim dokumentima

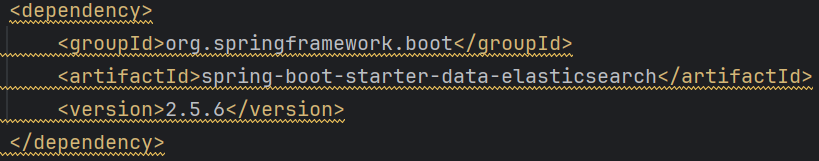
Predlog projekta, Nevena Gligorov E2 7/2024

# Arhitektura sistema i high-level pregled

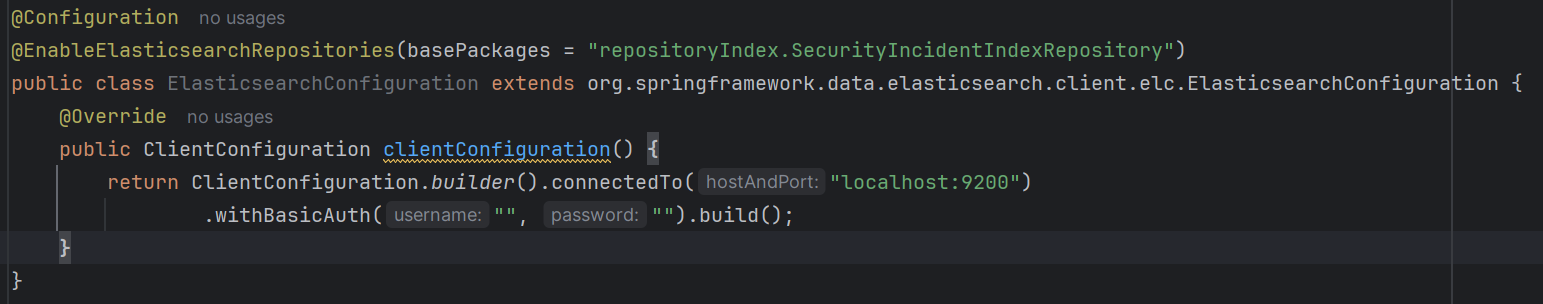
Sistem za monitoring i pretragu sigurnosnih incidenata će biti realizovan kao *SpringBoot* aplikacija na backend-u i *Angular* aplikacija na frontend-u. *Elasticsearch*, *Logstash* i *Kibana* koji predstavljaju osnovu *ELK stack*-a, zajedno sa *MinIO*-om, biće pokrenuti preko *Docker*-a uz pomoć *docker-compose.yml* fajla i odgovarajućih konfiguracionih fajlova. Za sve komponente *ELK stack*-a će se koristiti verzija 7.17, a ukoliko se naiđe na problem prilikom korišćenja ove verzije, preći će se na neku od novijih, gde bi poslednja solucija bila lokalna instalacija komponenti koje to budu zahtevale. Osim pomenutih, koristiće se i *PostgreSQL* baza koja će lokalno čuvati informacije vezane za korisnike i koja će predstavljati mesto prezistentnosti kada je reč o sigurnosnim incidentima.

## Komponente arhitekture

***Elasticsearch*** je distribuirani, *open source* mehanizam za pretragu i analitiku. *Dependency* i konfiguracija koja je potrebna da bi se *Elasticsearch* koristio u okviru aplikacije su dati na slikama 1 i 2.



Slika 1. Elasticsearch dependency

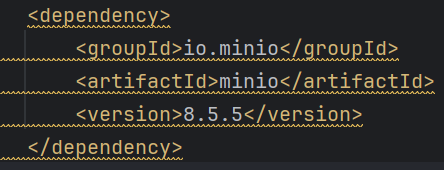


Slika 2. Elasticsearch konfiguracija

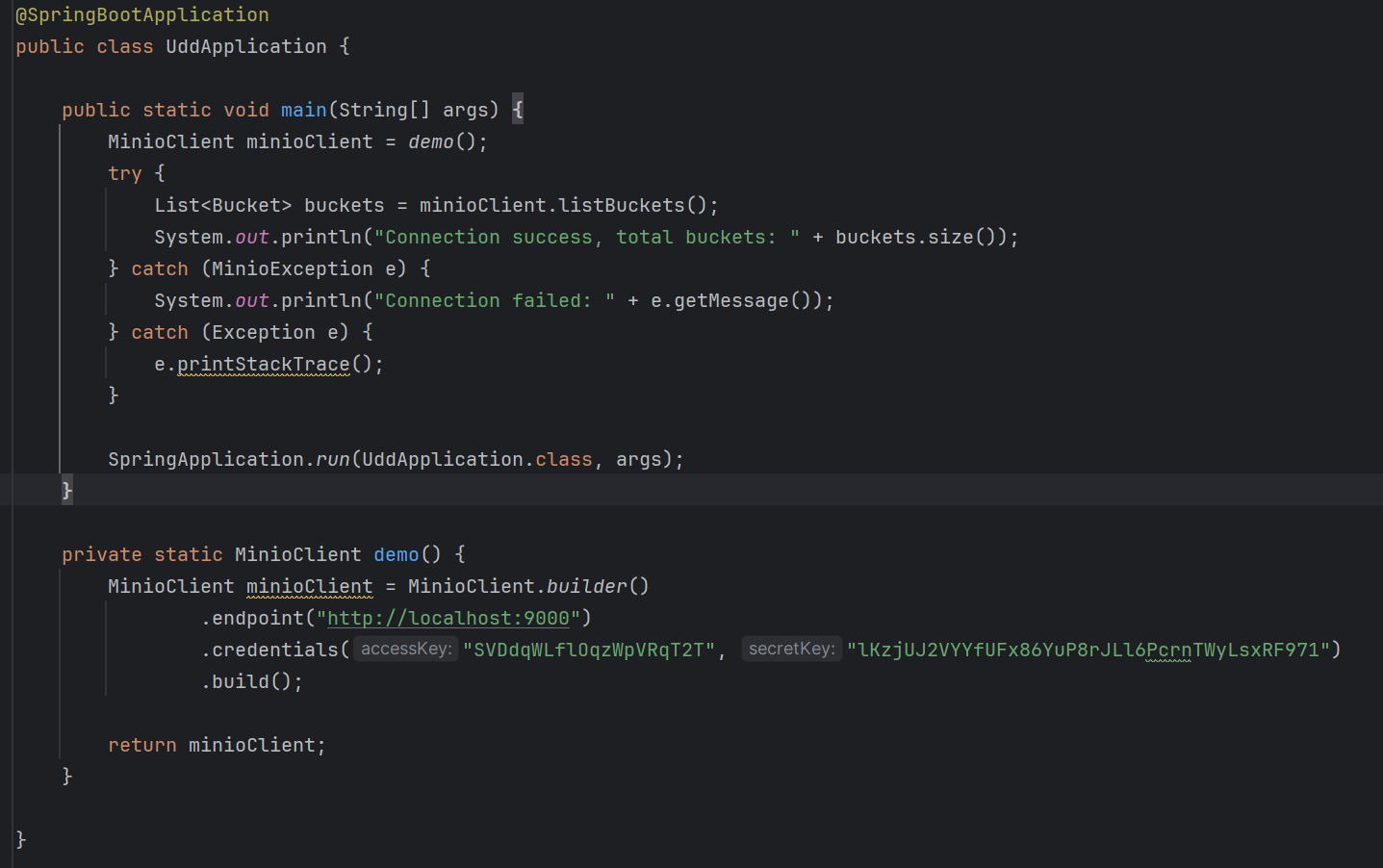
***Logstash*** je *open source pipeline* za obrađivanje podataka koji prihvata podatke sa jednog ili više izvora, transformiše ih i prosleđuje ka jednom ili više izlaza, odnosno odredišta. U slučaju sistema za monitoring i pretragu sigurnosnih incidenata, *Logstash* će podatke primati sa samo jednog izvora, koji predstavlja *.log* fajl u kome će se nalaziti svi logovi *SpringBoot* aplikacije. Nakon pristizanja, sigurnosni incidenti će biti sačuvani u PostgreSQL-u, prilikom čega će servis tu informaciju upisati i u *.log* fajl u formatu: *Timestamp LogLevel Attacked Org: AOName where Security Org: SOName with Member of incident team: fullName and severity: severity*. *Grok* filter podatke sa ulaza obrađuje i strukturirane ih šalje ka izlazu definisanom u konfiguracionom fajlu, odnosno ka Elasticsearch-u.

***Kibana*** omogućava vizualizaciju indeksiranih podataka, njihovo analiziranje i kreiranje različitih grafika. Povezivanjem sa *Elasticsearch*-om je realizovano uz pomoć komande koja je definisana u okviru *docker-compose.yml* fajla.

***MinIO*** predstavlja *open source* distribuirano skladište za objekte u okviru koga će se *PDF* dokumenti čuvati u sirovom obliku. Dependency koji će biti korišćen, kao i povezivanje sa SpringBoot aplikacijom je prikazani su na slikama 3 i 4.



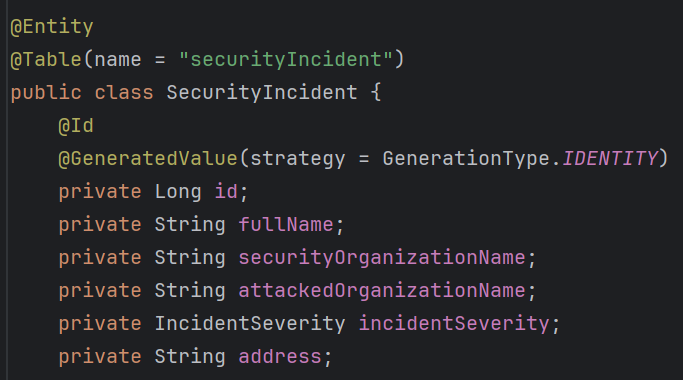
Slika 3. MinIO dependency



Slika 4. MinIO konfiguracija

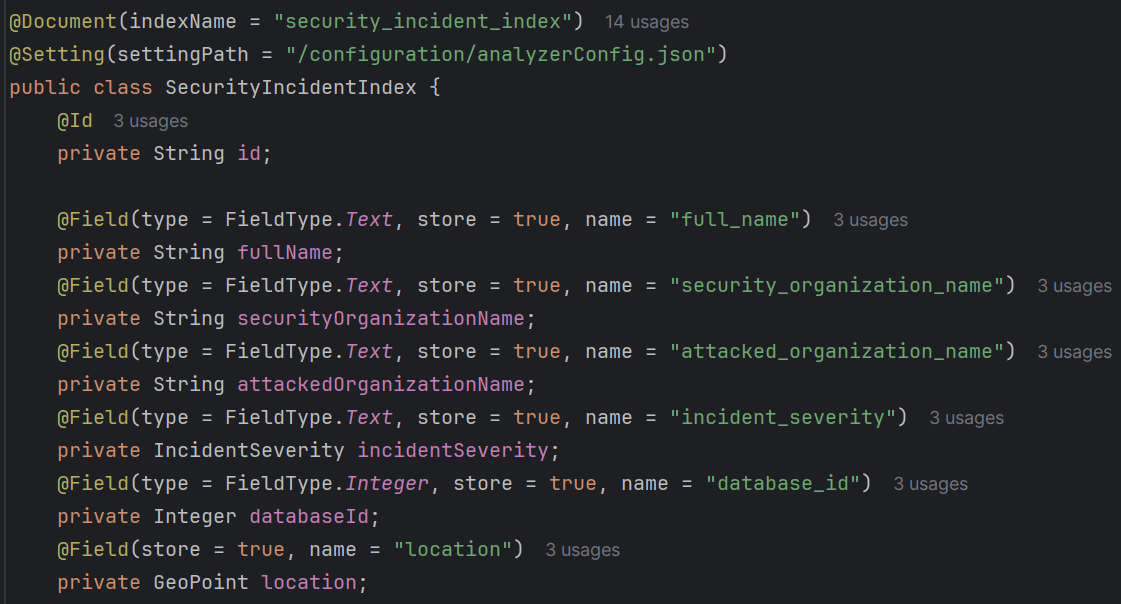
# Indeksiranje i pretrage

*SecurityIncident* klasa sadrži polja pobrojana u specifikaciji, kao što je prikazano na slici 5, i biće korišćena za čuvanje podataka u okviru PostgreSQL baze.



Slika 5. SecurityIncident

Kako je potrebno omogućiti geolokacijsku pretragu, umesto da u indeksu bude sačuvano polje za adresu pogođene organizacije tipa String, čuvaće se polje tipa *GeoPoint*. Prema tome, na slici 6 se nalazi primer klase sa poljima koja će se naći u indeksu. Uz pomoć [*LocationAPI*](https://locationiq.com/demo)-a je iz tekstualne adrese moguće dobiti informacije o geografskoj širini i dužini. *DatabaseId* predstavlja id entiteta sačuvanog u PostgreSQL bazi, kako bi postojala povezanost sa perzistentnom verzijom podataka. *@Document* anotacijom se definiše naziv indeksa u okviru koga će se čuvati podaci, a koji će se kreirati ukoliko već ne postoji. *@Setting* anotacija definiše putanju na kojoj se nalazi *analyzer*. Analyzer definisan nad čitavom klasom znači da će biti korišćen za svako polje definisano u klasi.



Slika 6. SecurityIncidentIndex

Struktura analyzera je prikazana na slici 7. Koriste se filteri koji će čitav tekst konvertovati u mala slova, na isti način će biti tretirana slova sa i bez ,,kvačica’’, poput č ć i c, ukoliko je tekst pisan na ćirilici, vršiće se konvertovanje u latinicu, izbacivaće se suvišne reči i reči će biti skraćivane na njihov koren.



Slika 7. Analyzer konfiguracija