

Jan Sarivuo

Full Stack Developer | Web-kehittäjä | API-integraatiot | WordPress & FastAPI

 jan.sarivuo@gmail.com - +358 40 839 5777

 Portfolio: janstech.github.io  GitHub: github.com/janstech

 LinkedIn: linkedin.com/in/jan-sarivuo

Tuotehaku-arkkitehtuuri – DataStoppi Oy

1. Tausta

Tämä projekti on osa työssäoppimisjaksoani DataStoppi Oy:ssä (08/2025–11/2025). Tavoitteena oli suunnitella ja toteuttaa tuotehaku- ja tuotetietojärjestelmä, joka yhdistää useiden tukkureiden tuotetiedot, hinnat ja varastosaldot yhteen, helposti käytettävään näkymään. Ratkaisun tuli integroitua olemassa olevaan WordPress-pohjaiseen extranetsivustoon ja toimia luotettavasti ilman manuaalista päivystämistä.

2. Arkkitehtuuri – yleiskuva

Järjestelmä perustuu kolmikerros-arkkitehtuuriin:

- Frontend: WordPress (extranet), jossa omat proxy-lisäosat ja shortcodet tuotehalle ja tuoteryhmien selailulle.
- Backend: FastAPI-palvelu, joka tarjoaa REST-rajapinnat hakuun, kategorioihin ja data-reload-toimintoihin.
- Tietokanta: MariaDB, johon toimittajien syötteet normalisoidaan (tuotteet, hinnat, varastot, kategoriat).

Toimittajasyötteet (esim. Aurdel ja Also) noudetaan ja muunnetaan Python-skripteillä muotoon, josta FastAPI voi palvella dataa WordPressille. Päivitykset ajetaan ajastetusti.

3. Frontend – WordPress ja käyttökokemus

WordPressiin toteutettiin omat shortcodet:

- [tuotehaku] – vapaa hakusivu, jossa tuotteita voi hakea nimellä, merkeillä ja muilla hakusanoilla.
- [tuoteryhma] – tuoteryhmien selauslista hakutoiminnolla ja sivutuksella.
- [tuoteryhmahaku] – valitun tuoteryhmän tuotteet, suodatettuna esimerkiksi toimittajan ja varastosaldon mukaan.

Käyttöliittymässä keskityttiin selkeään taulukkomaiseen listanäkymään, oletuksena pääällä olevaan “näytä vain varastossa” -suodattimeen sekä siihen, että haku on nopea ja reagoi ilman raskaita sivunlatauksia.

4. Backend – FastAPI ja hakurajapinnat

FastAPI-sovellus tarjoaa mm. seuraavat rajapinnat:

- /search – tuotehaku
- /categories – tuoteryhmien listaus
- /category-products – rajattu listaus tuoteryhmän tuotteista
- /reload-data – data-reload toimittajasyötteistä
- /healthz – terveys- ja tilannekysely valvontaa varten

Hakutoiminnallisuus toteutettiin siten, että haku on aksenttivapaa ja kirjainkoon suhteen joustava, monisanaiset haut käsitellään “AND”-periaatteella ja taustalla käytetään indeksejä ja tarvittaessa fulltext-hakua suorituskyvyn varmistamiseksi.

5. Tietokanta – normalisoitu tuotetieto

MariaDB:hen suunniteltiin kevyt mutta laajennettava skeema:

- products – yksi rivi per tuote/toimittaja, sisältäen mm. tuotetiedot, hinnat ja varastosalon.
- categories – hierarkkinen tuoteryhmärakenne.
- supplier_feed_meta – tietoa toimittajan syötteen noutokerrasta, rivimääristä ja onnistumisesta.

Tavoitteena oli, että uudet toimittajat on helppo lisätä myöhemmin ilman, että hakurajapinnat tai WordPress-osa täytyy kirjoittaa uusiksi.

6. Ajastus, varmuuskopiot ja valvonta

Järjestelmään toteutettiin cron-ajastukset mm. päivittäiselle data-reloadille (toimittajasyötteet → tietokanta), päivittäisille ja viikottaisille varmuuskopioille sekä palvelun automaattiselle käynnistykselle esimerkiksi uudelleenkäynnistykseen jälkeen.

Lisäksi järjestelmään lisättiin prosessin watchdog, joka tarkistaa taustapalvelun toiminnan ja käynnistää sen tarvittaessa uudelleen, lokivahti, joka etsii virheitä lokeista ja ilmoittaa niistä, sekä Slack-hälytykset, jotka tuovat tärkeimmät virheet ja tilannepäivitykset näkyviin reaalialjassa.

7. Tietoturva ja toimintavarmuus

Tietoturvassa huomioitiin mm. rajatut tietokantaoikeudet, ympäristömuuttujien käyttö (esim. yhteydet ja avaimet eivät ole versiohallinnassa), rajapintojen suojaus yhdistelmällä API-avainta, IP-rajausta ja WordPress-kirjautumista sekä resurssien ja parametrien rajaaminen (esim. hakujen sivutusrajat).

Valvonta ja automaattinen hälytys yhdistettynä varmuuskopioihin mahdollistavat sen, että järjestelmä pystyy palautumaan häiriöstä ilman jatkuvaan käsin tehtävää valvontaa.

8. Oma roolini ja opit

Projektissa vastasin kokonaisuudessaan teknisestä toteutuksesta: arkkitehtuurisuunnittelusta, integraatioista, tietokantamallista, ajastuksista ja valvonnasta käyttöliittymään asti.

Keskeisiä oppeja olivat kokonaistason arkkitehtuurin suunnittelu (frontend-API-tietokanta-integraatiot), tuotannon ajastusten, varmuuskopioinnin ja valvonnan käytännön toteutus, suorituskykyisen ja käyttäjäystävällisen haun rakentaminen normalisoidun datan päälle sekä yhteistyö asiakkaan ja toimittajien suuntaan teknisiä integraatioita toteuttaessa.