



id: SI_003

model: openai/gpt-5.4-mini

seeds:

04_risiko: sårbarhet

03_relasjon: tillit

05_handling: implementer

01_perspektiv: operasjonell

02_tidhorisont: langsiktig

Instans SI_003

Seeds: operasjonell | langsiktig | tillit | sårbarhet | implementer

Å samle komplett prisdata på **alle varer i verden** er ikke først og fremst et data-problem, men et **operasjonelt nettverksspørsmål**. Den vanligste feilen er å tenke "build a giant scraper". Det skalerer dårlig, skaper juridisk friksjon og gir lav tillit. Den riktige linsen er å bygge et **globalt pris-innsamlingssystem med flere kilder, standarder og valideringslag**.

Kompletthet er umulig som absolutt mål, men mulig som **grad av dekning**:

1. **Handlekurv-varer**: dagligvarer, elektronikk, klær, helseprodukter
2. **B2B/engros**: katalogpriser, kontraktspriser, volumrabatter
3. **Lokale markeder**: butikker, markedsplasser, apper, distributører
4. **Tidsdimensjon**: pris per tidspunkt, ikke bare "nåpris"

Det mest robuste designet er en **trelagsarkitektur**:

- **Kildeinnsamling**: API-er der de finnes, partnerfeeds, EDI, POS-integrasjoner, nettbutikker, offentlige prisregistre.
- **Normalisering**: enhet, valuta, pakningsstørrelse, moms, geografi, kampanje vs listepreis.
- **Tillitsscore**: hver pris får en kildevekt basert på observasjonstype, recency, historikk og kryssbekreftelse.

Sårbarheten er å stole på én type datakilde. Nettpreise er ofte dynamiske, skjulte eller personaliserte. Derfor må systemet bruke **redundans**: minst 2-3 uavhengige observasjoner per vare når mulig. For eksempel: Amazon, Walmart og en lokal forhandler for samme SKU/GTIN. For varer uten GTIN må man bruke produktidentitetsmatching via bilde, tekst og attributter.



På langs sikt bør målet være en **global prisgraf**:

- identitet: GTIN, MPN, produsent, ingrediens/attributter
- observasjon: pris, dato, geometri, kanal
- kontekst: kampanje, lagerstatus, frakt, avgifter

Dissens: Jeg ville **ikke** starte med “alle varer”. Start med **20-50 millioner høyt omsatte SKU-er** og bygg tillit gjennom dekning, ikke totalitet. Når operasjonen er stabil, utvides nettverket land for land.

Neste steg:

1. Definer en prisstandard: SKU-identitet, enhetsnormalisering og tillitsscore.
2. Piloter 3 kilder per marked i 2 land og mål dekning/feilrate.
3. Bygg en global master for 1 million produkter før du skalerer til 10+ millioner.