



id: SI_002

model: anthropic/claude-sonnet-4.6

seeds:

04_risiko: usikkerhet

03_relasjon: makt

05_handling: implementer

01_perspektiv: strategisk

02_tidhorisont: kortsiktig

Instans SI_002

Seeds: strategisk | kortsiktig | makt | usikkerhet | implementer

Realiteten ingen sier høyt

Komplett prisdata på alle varer er en maktillusjon. Det finnes ~10 milliarder unike produkter globalt (GS1-estimat). Å “samle alt” er ikke et teknisk problem — det er et **strategisk feilstilt mål**.

Det interessante spørsmålet er: komplett nok for hvem, og innen når?

Maktstrukturene som blokkerer deg

Tre aktører kontrollerer prisdata i dag:

- **Amazon** (~350M produkter, lukket API)
- **Alibaba/1688** (~200M produkter, delvis tilgjengelig)
- **GS1-nettverket** (barcodes, men uten live-priser)

De vil ikke gi deg dataen gratis. Prisdata er makt. Walmart bruker sanntidspriser som konkurransevåpen. Nielsen selger prisindekser for millioner. Du konkurrerer mot godt kapitaliserte aktører med eksisterende datafestninger.



Kortsiktig implementeringsstrategi (0-90 dager)

Ikke prøv å samle alt. **Velg en vertikal og domin den:**

1. **Skrp det åpne** — Google Shopping Feed (~2B produkter, offentlig tilgjengelig via structured data), PriceRunner-API, Prisjakt. Kostnad: ~\$500/mnd i infrastruktur.
2. **Kjøp det lukkede** — Distribuerte datasett fra Bright Data eller Oxylabs. ~\$2000-8000/mnd for kategorispecifikk scraping i skala.
3. **Bygg pipeline nå, perfekter senere** — En MongoDB-instans med daglig batch-innhenting på én kategori (f.eks. elektronikk i Norden) er mer verdifull enn et perfekt system som aldri lanseres.

Usikkerheten du må prise inn

- **Juridisk risiko:** GDPR + robots.txt-tvister. Ryanair vant mot scraping i EU 2022.
- **Data-decay:** 30% av e-handelspriser endres daglig. "Komplett" i går er foreldet i dag.
- **Definisjonsproblemer:** Er en bruktpriis på Finn.no "prisdata"? Er en forhandlet B2B-pris det?

Neste steg

1. **Definer scope aggressivt smalt** — én bransje, ett land, 30 dager. Lever noe konkret.
2. **Inngå datadeling-avtale med én stor aktør** (f.eks. Komplett.no eller Elkjøp) mot gjensidig verdi — unngå juridisk eksponering.
3. **Sett opp Scrapy + Kafka-pipeline denne uken** på 10.000 produkter og mål dataferskhet — det avslører raskt hvilke problemer som faktisk er kritiske.