

# Nasjonal Vegdatabank: Stordatateknologi på små data

Jan Kristian Jensen<sup>1</sup>

*Jan Kristian Jensen: Large data technology on small data. The Norwegian national road database.*

KART OG PLAN, Vol. 77, pp. 127–128, POB 5003, NO-1432 Ås, ISSN 0047-3278

The National Road Data Base (NVDB) is the repository for all road data in Norway. More than 360 feature types are defined for a wide range of applications, ranging from car navigation, physical inventory to purely administrative purposes. All 253 million NVDB objects can be held within the same relational database, so NVDB is by definition not a big data problem. The Norwegian Public Road administration has successfully applied “Big Data” – technology to serve NVDB-data efficiently through our new NVDB API.

*Key words:* large data technology, road data base

*Jan Kristian Jensen.* Senior engineer, Statens Vegvesen, vegdirektoratet, seksjon NVDB og Geodata, Postboks 8142 Dep., NO-0033 Oslo. E-mail: Jan.kristian.jensen@vegvesen.no

## Alle kjørbare veger lengre enn 50 meter

Nasjonal vegdatabank (NVDB) er Norges register over alle kjørbare veger lengre enn 50 meter, pluss mange gang-/sykkelstier. I tillegg til vegnettet har NVDB ca 361 ulike typer fagdata definert i datakatalogen NVDB Datakatalog, 2017. Noen typer fagdata har kritiske detaljer for bilnavigasjon, som fartsgrenser, vegsperringer, restriksjoner på vekt m.m. Andre er inventaroversikt over fysisk utstyr langs vegen, som skiltplater, stikkrenner etc eller mer administrative forhold, som kontraktsområder og veg nummer. NVDB driftes av Statens Vegvesen, men i et tett samarbeid med Statens Kartverk og kommunene.

## NVDB er ikke stordata

Cirka 253 millioner fagdata-objekter er jo et respektabelt datavolum, men alt sammen får helt greit plass i samme relasjonsdatabase. Ergo er NVDB per definisjon intet «Big Data» – volum, men det kan diskuteres om effektiv søk uthenting og distribusjon av 253 millioner objekter kvalifiserer til «Big Bata» – stempelet.

NVDB kvalifiserer definitivt ikke på Velocity eller Variety (McAfee og Brynjolfsson,

2012), eller de litt nyere kriteriene Variability og Veracity knyttet til variasjon i datatyper og kvalitet (Tabell 1). Selvsagt finnes det flere «Big Data» definisjoner, med varierende antall «V»-er og ordlyd.

## NVDB bruker «Big Data» teknologi for effektive spørringer og distribusjon

Da vi bygget webløsningen Vegkart i 2012 (<http://vegkart.no>) fantes det ingen systemer i Statens vegvesen som kunne levere NVDB-data raskt nok, i store nok volum, for de spørringene som trengtes.

Hvis du for eksempel vil vite fartsgrensene på Fv33 i kontraktsområdet 0903 Arendal vest 2016-2021: For å finne de riktige 9 blant 859.286 fartsgrenser (per 20.4.2014) må vi kombinere med informasjon fra 1.4 millioner vegreferanser og 1573 kontraktsområder. Svaret må komme på sekundet, ellers går Vegkart-brukerne lei. De fleste relasjonsdatabaser vil slite med å svare raskt nok på slike spørringer, i hvert fall innenfor vårt budsjett. Men vi oppfordrer alle til selv å sjekke responstiden på Apache SOLR søkeindeks i webapplikasjonen Vegkart <http://vegkart.no>.

1. Statens Vegvesen, vegdirektoratet / seksjon NVDB og Geodata.

Tabell 1. NVDB passer ikke helt med de fem «V»-ene som særpreger «Big Data» (Det finnes litt avvikende «Big Data»-definisjoner).

De store «V»-ene i «Big data».	Hvordan passer dette på NVDB
Volume	Alle NVDB-data passer inn i en database – ergo per definisjon intet “Big Data” – volum. Men det kan diskuteres om effektiv søk og distribusjon kvalifiserer til «Big Data» – stempelet.
Variety	Hver objekttype med egenskapsverdi i NVDB har en entydig definisjon i datakatalogen, og ajourholdes der – altså motsatt av «variety» i «Big Data» sammenheng.
Velocity	NVDB-data har lav hastighet: Det tar tid å bygge ny veg, og fartsgrenser endres ikke så ofte.
Variability	Gjennom 20 års forvaltning har det vært ulik praksis på registrering til NVDB, hvilke objekttyper og egenskapsverdier som registreres, og hva disse betyr. Men denne variasjonen blekner mot det som menes med “variability” i “Big data” – sammenheng.
Veracity	Ja, datakvaliteten varierer i NVDB, og den største utfordringen er manglende fagdata, spesielt på det kommunale vegnettet. Men disse utfordringene gir ikke stor nok «Veracity» i «Big Data» sammenheng.

Man trenger altså ikke et «Big Data» – problem for å dra nytte av «Big Data» teknologi. Tvert imot er slik teknologi ofte et helt naturlig førstevalg. Slik var det i 2012, og denne utviklingen har selvsagt fortsatt. Og dette er på ingen måte «nytt og uprøvd»: Teknologi ingen hadde hørt om for noen få år siden brukes i dag til å servere film, musikk, nyheter og sladder til milliarder av mennesker over hele verden. Og det aller meste av denne teknologien er fritt tilgjengelig som åpen kildekode, uten lisenskostnader eller byråkrati rundt innkjøp.

### Referanser

- Andrew McAfee og Erik Brynjolfsson. Big Data: The Management Revolution  
Harvard Business Review, oktober 2012.
- NVDB 2017. Statens Vegvesen: Nasjonal Vegdatabank. <http://www.vegvesen.no/fag/teknologi/Nasjonal+vegdatatabank>
- NVDB datakatalog, 2017. Statens Vegvesen. <http://www.vegvesen.no/fag/teknologi/Nasjonal+vegdatatabank/Datakatalogen>