

Ken Tore Tallakstad, PhDDeveloper / Data Scientist

OM MEG









VERKTØY



























DEMO FOR Å VISE:

- hvordan data kan brukes i Spark/Hadoop i en lambda arkitektur («real time» + batch)
- hvordan en ETL pipeline kan bygges med StreamSets
- hvordan Solr og Hue kan brukes til å få kjapp innsikt i batch data

Data prepp



Analyse

AIS DATA



Se skipsbevegelser i sanntid

AIS viser skipsbevegelsene til skip over 45 meter opptil 12 nautiske mil fra kysten. AIS-data hentes inn fra Kystverkets 50 basestasjoner langs kysten som fanger opp signaler fra alle skip over 300 bruttotonn som seiler i internasjonal fart.

KILDE 1:

```
(i) 153.44.253.27:5631
🔛 Apps 🖿 Fiske 🖿 Annet 🖿 Work 🖿 data 🖿 Sherpa 🖿 sysadmin 😸 Databricks Spark 🕻 bigSherpa - CM
!BSVDM,1,1,,A,H3n8`B1=E<4ppD000000000000000,2*07
!BSVDM,1,1,,A,13mC<L3P00PqjP`TkQm9f?vl28?8,0*19
!BSVDM.1,1,.B.13nK7T0001PL<@VRM<a;6V<l0<0<,0*06
!BSVDM,1,1,,B,H3mvrE4U@B?>1F0<4gogo00000000,0*76
!BSVDM,1,1,,B,H3mvrE4U@B?>1F0<4qoqo00000000,0*76
!BSVDM,1,1,,B,B3mtN?000`FPVJ:0w1:>swdUoP06,0*3B
!BSVDM,1,1,,B,13nRwN00010GRIvSML@AM5Fn00S>,0*4E
!BSVDM,2,1,4,B,53muq0400000hOSS3R0Lv1H4ddE:05=@62222215185147>h02QlQDlmRBsA,0*6E
!BSVDM,2,2,4,B,hCC1@VAC`80,2*2F
!BSVDM,1,1,,A,13mLQe0Pls1nk?<`W;q=G?vh0@?>,0*7C
!BSVDM,1,1,,A,13mVJQ00010vo:2Vv2I::6Nl00S3,0*12
!BSVDM,1,1,,A,H3m=SU4N@B?>1F0<;jkqo00H8220,0*74
!BSVDM,1,1,,A,13mDS:?000150uDWKaB0t9>l0<0D,0*1F
!BSVDM,1,1,,A,33FwQH50012>2fN`@mWKsn8l1DC:,0*11
!BSVDM,2,1,7,A,53e:iJ42A4QKTP7;C05LTi<tr04hLD<U85>2221?9@>756pA0=23ihCPDp88.0*08
!BSVDM,2,2,7,A,88888888880,2*3A
!BSVDM,1,1,,B,13ojQD0001Pkb;bU7<sJKC2jP@?E,0*72
!BSVDM,1,1,,A,13mFoV00011?i@vW9rTpSPTl00Rn,0*40
!BSVDM,1,1,,B,H3mlI34RC=D5PQn<=Plhno1P4420.0*0D
!BSVDM,1,1,,B,B3m8BS00006qeU`s2D7NKwdUoP06,0*56
!BSVDM,1,1,,A,13ogBr701B2:<R0`M4IdG9jh0<0;,0*7E
!BSVDM,1,1,,A,13mmHE?P02Po@0nU@r@>4?vl2@?I,0*70
```

NMEA – decoded: dekodere tilgjengelig på ulike språk

KILDE 1, DEKODING:

```
[n [1]: import ais
       ais.decode("13mLQe0P1s1nk?<`W;g=G?vh0@?>",0)
{u'cog': 342.0,
u'id': 1L,
u'mmsi': 257368500L,
u'nav status': 0L.
 u'position accuracy': 0L,
 u'raim': False,
 u'repeat indicator': OL.
u'rot': -731.386474609375,
 u'rot over range': True,
 u'slot number': 974L,
 u'slot timeout': 4L,
u'sog': 12.300000190734863,
 u'spare': 0L.
 u'special manoeuvre': OL,
u'sync state': 0L,
u'timestamp': 24L,
 u'true heading': 511L,
 u'x': 25.952383041381836,
 u'y': 70.97503662109375}
```

Nyttige parametere:

- 'mmsi': ID
- 'x': longitude
- 'y': latitude
- 'cog': kurs
- 'sog': hastighet
- 'rot': rotasjon

KILDE 2:

```
hbase(main):008:0> get 'mmsiShipInfo','259402000'

COLUMN

P:imo

P:mmsi

P:name

P:name

P:type

4 row(s) in 0.0330 seconds
```

HBase tabell med metadata

DATAFLYT

bigSherpa – Hadoop Cluster

