

Ontologibasert dataintegrasjon



Juni 2015 - Design/produksjon: FFI, Foto: Forsvaret

Informasjonssystemene i nettverksbasert forsvar er ofte utviklet og vedlikeholdt for separate formål uten felles grensesnitt eller datamodell. Ontologibasert dataintegrasjon bruker virtuelle grensesnitt for å løse utfordringene dette gir.

Forestill deg en militær analytiker i et operativt hovedkvarter som har som oppgave å planlegge og overvåke evakueringsflygninger inn i et stridsområde. Det er spesielt viktig å holde et våkent øye med evakueringsflygninger som er truet av fiendtlig aktivitet. Dersom en flygning er truet, er det analytikerens oppgave å lete etter vennlige styrker som er i stand til å nøytralisere trusselen i nærheten av landingsområdet.

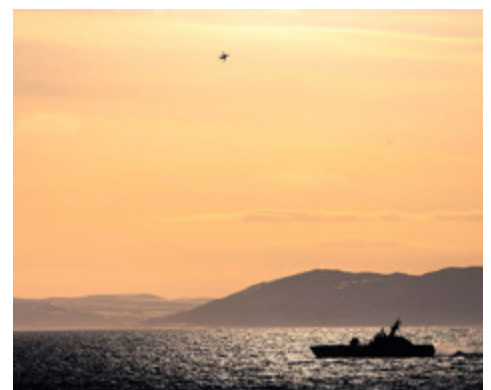
Analytikerens informasjonsbehov kan uttrykkes slik: finn alle evakueringsflygninger og vennlige enheter som er slik at flygningen klassifiseres som truet, og den vennlige enheten har kapabiliteter til å kunne bekjempe trusselen det er snakk om.

Behovet for informasjon

For å besvare dette informasjonsbehovet vil

analytikerens vanligvis måtte kombinere informasjon fra flere ulike systemer, som vi normalt vil finne i et operativt hovedkvarter. Analytikerens vil kanskje måtte konsultere en hendelseslogg for å følge utviklingen i stridsområdet, og et sporingssystem for evakueringsflygninger for å avgjøre hvilke landingssoner som er truet. I tillegg vil hun naturlig nok ha behov for et «blåprikkssystem» som viser hvor de vennlige styrkene står, og muligens også en Order of Battle-logistikk-database for opplysninger om kapabilitetene til de ulike vennlige og fiendtlige stridsenhetene.

Å kombinere alle disse opplysningene manuelt er en tidkrevende og skjør prosess, som fordrer at analytikerens kjenner de ulike informasjonssystemene og deres vanligvis ulike datamodeller og spørregrensesnitt.



▲ Samarbeid mellom sjø og luft i Porsangerfjorden var viktig under øvelsen Joint Viking i 2015.

Sammenstiller informasjon

Ontologibasert dataintegrasjon, kjent som Ontology Based Data Access (OBDA), tilbyr en dynamisk og fleksibel løsning på analytikerens problem med å sammenstille informasjon. I et OBDA-system er integrasjonen kun virtuell og derfor løst koblet til de underliggende datakildene. Dette betyr blant annet at de underliggende datakildene ikke behøver å være utviklet eller vedlikeholdt for å utveksle informasjon med hverandre.

Kjernen i et OBDA-system består av en begrepsmodell (også kalt en ontologi) som kan betraktes som en beregnbar spesifisering av hvordan typer av data i de underliggende kildene forholder seg til hverandre konseptuelt. Denne spesifiseringen fungerer som et abstraksjonslag som legges over de underliggende kildene for å presentere dem for analytikerens i vårt tenkte tilfelle, som om de skulle være én kilde. Begrepsmodellen designes gjerne slik at den reflekterer brukerens foretrukne vokabular, noe som tillater vår analytiker å uttrykke sitt informasjonsbehov med begreper som reflekterer hennes kompetanse.

Selve datainnsamlingen foregår ved at analytikerens formulerer sitt informasjonsbehov ved hjelp av et spørrespråk som reflekterer begrepsmodellen. Siden begrepsmodellen uttrykker forhold mellom typer av data i de underliggende kildene, gjør dette i sin tur det mulig å beregne hvilken informasjon som må hentes fra hvilke kilder, og hvordan den må kombineres for å svare på analytikerens informasjonsbehov. Denne oversettelsesprosessen gjør at et OBDA-system relativt enkelt kan tilpasses et vilkårlig antall kilder uten at kompleksiteten øker i brukerens øyne.

Begrepsmodeller som linser

I litteraturen for kommando- og kontroll skapes det ofte et inntrykk av at så lenge alle beslutningstakere får tilgang på den samme informasjonen og er kjent med den samme intensjonen, vil de konkludere likt. Empiriske studier viser at dette ikke er tilfellet. Informasjon må behandles og framstilles på relevant vis før den kan brukes. Informasjon om fienden foredles til etterretning, mens eksempelvis geografisk informasjon må tilpasses et egnet oppløsningsnivå for å sammenstilles med informasjon om egne og



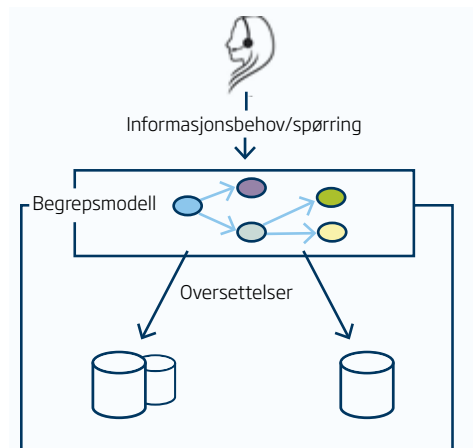
▲ Slik kan det se ut fra operasjonsrommet på Forsvarets operative hovedkvarter på Reitan.

► Begrepsmodell som søkegrensesnitt til heterogene datakilder.

andre styrker. Alt dette må i sin tur tilpasses avdelingens nivå og oppdrag. Det er derfor ikke noe enkelt svar på hvordan informasjon bør presenteres, det avhenger av formål, bakgrunnskunnskap og kontekst.

Med ontologibasert dataintegrasjon er det mulig å sette de samme underliggende datakildene i perspektiv på forskjellig vis avhengig av interesse og relevans. Siden integrasjonen av dataene er virtuell og formidlet av begrepsmodellen, er dette i bunn og grunn et spørsmål om å bytte til en annen begrepsmodell som uttrykker andre sammenhenger. Vi kan derfor tenke på en begrepsmodell som en «semantisk linse» som vi betrakter dataene gjennom. Forskjellige linser avslører forskjellige sammenhenger i de underliggende dataene.

Kontaktperson:
Audun Stolpe
audun.stolpe@ffi.no



Integrasjonssystemet Hårfagre

- FFI publiserte i 2014 første versjon av et dataintegrasjonsbibliotek kalt Hårfagre.
- Spesielt utviklet for å møte militærets krav til etterrettelighet og forutsigbarhet.
- Systemet implementerer en matematisk teori, også utviklet ved FFI, som garanterer at informasjonssøk returnerer korrekte og komplette resultater.
- Lisens: GNU 2, Open Source
- Språk: Scala
- Mer om Hårfagre: <http://bitbucket.org/auduns/harfagre>