

Geodata til mere end kort



Erik Kjems

Ph.d., civilingeniør, HD(A). Lektor ved Aalborg Universitet, Institut for Byggeri og Anlæg. Tidligere leder af Center for 3D GeoInformation og VR Media Lab ved Aalborg Universitet. Arbejder med rumlige interaktive modeller af by og landskab med særligt fokus på veje.

I takt med at megen af den information vi forsøger at fastholde med bl.a. geodata bliver tilgængelige som realtids information, må vores systemer og databeskrivelser nødvendigvis tilpasses disse nye udfordringer. Vores ønske om stadig flere informationer øger kravene til systemerne og organisationen omkring geodata. Når der er tale om "Big Data" og "Internet of Things" kan netop georeferencen spille en afgørende rolle, fordi den kan bidrage til at holde styr på datastrømmene. Er vores fælles nationale geodatagrundlag klar til denne rolle?

Keywords: Geografisk information, big data, bymodeller, landskabsmodeller

Indledning

Da jeg i februar-udgaven af Geoforums medlemsblad læste om FOTdanmark (Schielder, 2014), og om hvorvidt opgaven med at etablere et fælles nationalt datagrundlag nu var løst, fornemmede jeg en trang til at skrive den følgende artikel. Artiklen indeholder ikke nogen større refleksion om FOTdanmark, men har fokus på hvilke georelaterede opgaver der ligger lige forude. I nævnte artikel nævnes afslutningsvis, "at vore geodata med en ofte brugt frase er i en rivende udvikling, men hvad indebærer den". I centret for 3D GeoInformation på Aalborg Universitet har vi i snart 15 år arbejdet med koblingen mellem geodata repræsenteret i systemet og den virkelige verden. Denne tilgang har fokus på geodata som levende og foranderlige data, idet den verden, de beskriver, er foranderlige. I takt med, at megen af den information vi forsøger at fastholde med bl.a. geodata bliver tilgængelige som realtids information, må vores systemer og databeskrivelser nødvendigvis tilpasses disse nye udfordringer. Vores ønske om stadig flere informationer øger kravene til systemerne og organisationen omkring geodata. Når der er tale om "Big Data" og "Internet of Things" kan netop georeferencen spille en afgørende rolle, fordi den kan bidrage til at holde styr på datastrømmene. Er vores fælles nationale geodatagrundlag klar til denne rolle?

Fra 3D bymodeller til ?

Der har længe været fokus på de data, der ligger over "kort"fladen gerne omtalt som 3D bymodeller, idet det var byerne, der benyttede 3D modeller til visualiseringsformål. Mange af de modeller, der blev produceret på daværende tidspunkt,



er forældede og kræver en kærlig hånd for på ny at blive gjort anvendelige, mens andre byer har valgt løbende at vedligeholde de dyrekøbte bymodeller. Vi har desværre ingen fælles definition for de dele, der indgår i en by-og landskabsmodel. Hvor kunne det altså være fantastisk at have en fælles landsdækkende 3D grundmodel liggende med en beskeden CityGML level2 definition af indgående elementer. Men arbejdet er nærmest gået i stå og ikke kun her i landet. Generelt er produktionen af bymodeller stagneret internationalt (Morton, Horne, Dalton, & Thompson, 2012). En gang skulle en by af en hvis størrelse have en bymodel "So ein Ding muß ich auch haben"-agtigt, og sat lidt på spidsen blev der produceret bymodeller, uden at der var indtænkt nogen form for bagvedliggende organisation med kompetence indenfor anvendelsen af 3D bymodellen. Dem, der havde organisationen på plads og betalte omkostningerne, fik en værdifuld vedligeholdt model, som kunne indgå i byplanlægningen og benyttes til glimrende præsentationer ved borgerinddragelser og lignende. Man kan også finde en del anvendelser, der går ud over rene projektvisualiseringer, hvor modellen bliver benyttet til den stribe af anvendelsesmuligheder som er blevet skitseret gennem tiden og også blev peget på her i landet (Batty et al., 2001; Flemming et al., 2011). Således benyttes de bl.a. til rumlig støjkortlægning og visualisering af luftkvaliteten. For mange byer har bymodellerne været en dyr fornøjelse og anses måske ligefrem for en fejlinvestering, men her ligger fejlen ikke ved modellen, men ved manglende kompetence omkring den.

Ikke desto mindre er jeg af den opfattelse, at vi nu skal i gang for alvor. Dette kan virke besynderligt set i lyset af, at der reelt set ikke er kommet flere anvendelsesmuligheder end dem, der blev peget på for år tilbage. Men jeg vil hellere kigge frem og se på, hvilket billede der tegner sig i fremtiden og kan anes tydelige konturer af allerede nu. Jeg formoder, at rigtig mange har hørt om udtrykkene "Big Data" (Mashey, 1997), "Internet of Things" som allerede har fået sit faste acronym (IoT)

(Atzori, Iera, & Morabito, 2010; Sarma, Brock, & Ashton, 2000; Weiser, 1991), og måske også begrebet "Data Fusion" (Waltz, Llinas, & others, 1990), som er ved at komme på fleres og fleres læber. Udtrykkene er slet ikke nye og udviklingen indenfor områderne har pågået i et par årtier. Alle tre områder har potentiale til at få stor betydning for geodataområdet, men lad mig tage dem en ad gangen.

Big Data, store datamængder bliver oftest nævnt i forbindelse med alle de mange sociale sider, såsom Facebook, Twitter, Instagram, LinkedIn og mange flere, men også data i forbindelse med transport, vejret, biologiske fænomener osv. er en del af Big Data. Big Data er kendetegnet ved at data er ustruktureret i forbindelse med sin herkomst og ofte svære at organisere, hvilket gør dem krævende i datalogisk øjemed ("Big Data Definition," n.d.). Grunden til at udtrykket netop nu begynder at komme frem i lyset for alvor er bl.a. forbindelsen til IoT området.

IoT er kendetegnet ved, at ting får IP adresser og bliver koblet på internettet. Udvidelsen af adresseområdet i internet Protokollen ved skiftet fra version IPv4 til IPv6 har øget antallet af mulige enheder på nettet fra ca. 4×10^9 til $3,4 \times 10^{38}$. Formålet hermed var netop at få plads til de mange "ting" eller enheder, som man forventer vil blive tilsluttet til internettet de kommende år. Her er der ikke tale om dem, vi selv har styr på, såsom smartphones, tablets osv. men i enheder, som bliver koblet på internettet og indsamler information fra vores omverden, lige fra vaskemaskinen og varme anlægget i huset til avancerede vejrstationer eller trafikportaler. Populært bliver denne type information kædet sammen med "Sensor City" begrebet, som opstod på Harvard i 2007, hvor man havde placeret 1000 luftforureningsfølere omkring i byen Cambridge. Mængden af data, der kommer fra disse typer enheder, har en eksponentiel vækst disse år. Analysefirmaet Gartner postulerer, at IoT vil udgøre 26 milliarder enheder i 2020, hvilket er mere end 30 gange, hvad vi havde i 2009 (Middleton, Kjeldsen, & Tully,

2013). Bemærk, at tallet stadigvæk kun indeholder ting på nettet og ikke interaktive enheder. IBM fastslår, at netop "Smart City" konceptet med et groende antal sensorer i byen vil være en af de helt store spillere i denne udvikling (IBM, n.d.). En af dem, der virkelig må have kendskab til udviklingen i mængden af informationsstrømme er Eric Schmidt, formand for Googles bestyrelsen og daværende direktør. Han udtalte følgende på en Techonomy Conference (<http://techonomy.com/conf/>) i 2010: "Every two days now we create as much information as we did from the dawn of civilization up until 2003, that's something like five exabytes of data". Man kan kun gisne om, hvad tallet er i dag. Nu mangler vi kun at kombinere de mange data vi samler sammen, hvorfor "Data Fusion" kommer ind i billedet.

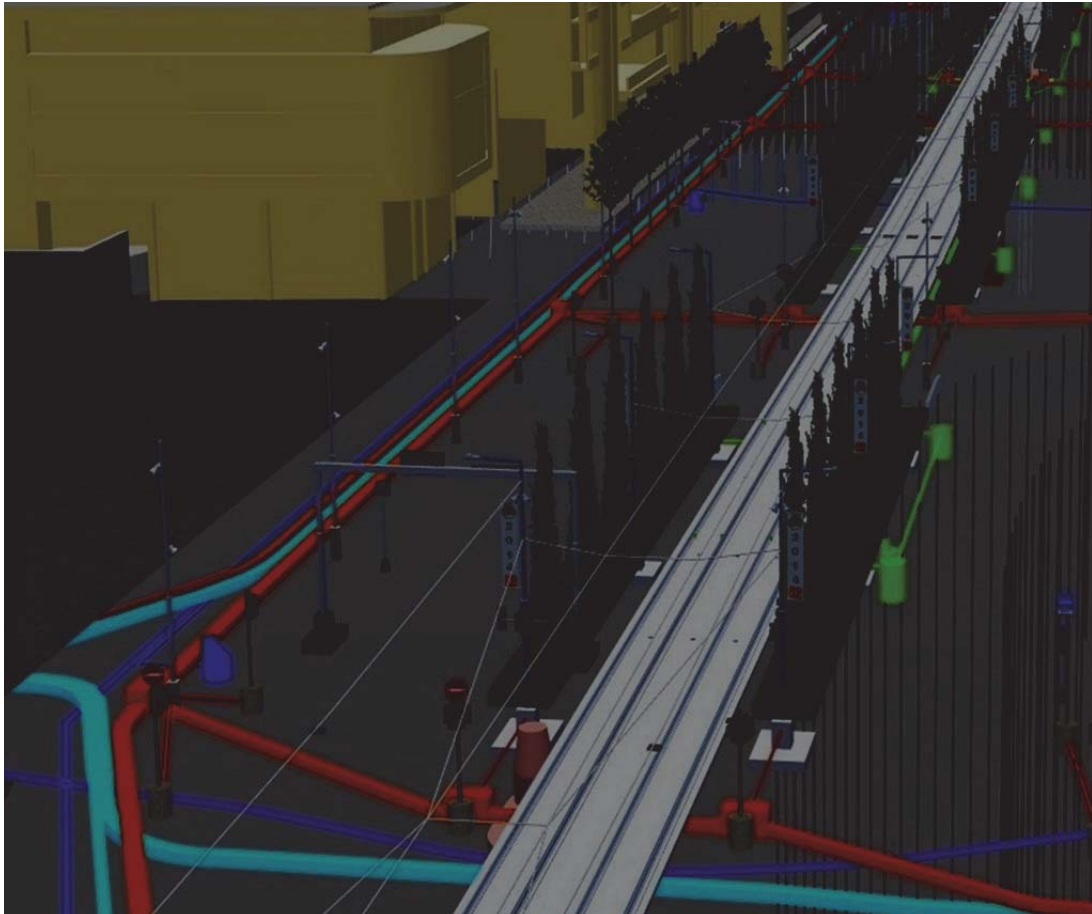
Data Fusion omtales i to forskellige sammenhænge, hvor især den ene berører et velkendt fænomen i geodata branchen. Samkøring af geodata eller digitale data generelt har altid været og er stadigvæk en stor udfordring, som der bruges enorme ressourcer på. De store kommercielle spillere holder benhårdt fast i deres dataformater, mens åbne formater ihærdigt prøver at vinde gehør og blive implementeret bredt, så også mindre spillere på markedet kan komme ind i varmen. Udfordringen er kun øget i takt med, at GIS og 3D systemerne er ved at smelte sammen, idet CAD verdenen begynder at producere geodata i form af virtuelle konstruktioner både indenfor byggeri og infrastruktur. Eller rettere 3D systemerne får attributdata tilknyttet objekterne, der indgår. Således vil man kunne klikke på et hvilket som helst objekt i den efterfølgende figur 1. og få indgående informationer frem om det pågældende objekt. Samkøring til produktion af styklister, kollision kontrol og meget andet er helt almindeligt i avancerede CAD systemer i dag.

De modelservere, der er ved at komme ind på markedet indenfor byggeri og infrastruktur, er systemer, der kan modtage og afgive data i åbne formater og kan styre adgang til forskellige dele af modellen med hensyn til opdateringer osv. Herved står disse i stærk kontrast til

de systemer, som blot kan indlæse data i 2D og 3D og samkøre dem i en vis udstrækning. Oftest med tab af information, pga. datakonverteringen fra ét format til et andet, til følge.

Således skal regulære 3D formater kommende fra f.eks. Autodesk og Bentley såsom de åbne formater IFC (Industrial Foundation Classes) kunne fungere sammen med bymodeller i CityGML og f.eks. ESRI's GIS produkter. Men er det realistisk? Formatkampen har især en taber, brugeren eller rettere forbrugeren.

Den anden og nok så interessante del af datafusionen ligger i samkøringen af datainformation. Dvs. information, som er samlet af primært sensorer. Et godt eksempel er trafiksystemerne, som samler data fra forskellige typer spoler i vejene, fra kameraer som aflæser nummerpladen og herved bestemmer den øjeblikkelige rejsehastighed på strækningen, registreringer langs vejen f.eks. vejtemperaturen eller også information fra køretøjerne selv, der ved, hvor de er, og med hvilken hastighed de bevæger sig. Alt sammen information, som kan bruges til at informere vejens trafikanter om evt. problemer på vejen. Den store udfordring her ligger i at samkøre data, altså fusionere data på en sådan måde, at de giver mening på tværs af systemerne. Dette er en meget stor udfordring, da de enkelte delelementer har forskellige producenter og derfor ofte kommunikerer med forskellige dataprotokoller eller "sprog" om man vil. Investeringer i vejene har som oftest en meget lang tidshorizont, hvorfor også denne problemstilling kan spille ind, idet en vedtaget standard kan nå at blive ændret adskillige gange inden, at hardwaren i vejen bliver udskiftet. Udviklingen indenfor datafusionen skal således agere som bindeled imellem de mange teknologiske øer, som er kommet gennem de seneste årtier. Men hvis det lykkes, vil ambulancen under udrykning kunne advare bilerne på vejen, omstille signalsystemet til grøn bølge, sørge for broen ikke hæves i utide og i øvrigt dirigeres udenom evt. forhindringer på vejen uden, at noget menneske behøver at være indblandet, blot for at udvide eksemplet lidt. Det siger sig



Figur 1. CAD model over Infrastukturanlægget i undergrunden ved Bjørvika i Oslo indeholdende oplysning om indgående elementer. (Vianova, Oslo)

selv, at hvis vi skal have 30 milliarder enheder på nettet kræver det helt andre systemer end dem, vi kender i dag til at håndtere disse. Det vil kræve, at systemerne bliver autonome, ikke blot med hensyn til registreringen af informationen men i høj grad til at tage beslutningerne for os. Registreres oversvømmelser i kloaksystemet, skal systemet have mulighed for at omdirigere spildevandet automatisk, så en større katastrofe undgås.

Geodata i fremtidens realtidssystemer

Det må være åbenlyst for de fleste, at det vil være over-

ordentligt vanskeligt at holde styr på de mange enheder i et abstrakt system bestående af et hav af tabeller, ID numre m.m. Især når man tænker sig 'Smart City'-konceptet ført ud i sin mest komplekse struktur. Systemerne, isoleret set, er i dag relative små størrelser og derved stadigvæk rimelig overskuelige, og de er samlet indenfor mere eller mindre standardiserede rammer, som tillader en rimelig kvalitet indenfor kommunikationen imellem enhederne. Men netop systemerne er begyndt at vokse alvorligt både i geografisk udstrækning og i kompleksitet. Herudover er man begyndt at få fokus på samko-

ring af data. Nu skal f.eks. togenes realtidssystem kobles sammen med bussernes og det øvrige trafiksystem. Men samkøringen er for øjeblikket besværliggjort af manglende konformitet systemerne imellem.

Der er mange udfordringer ved denne samkøring, men alle sensorer og øvrige informationsenheder har én fællesnævner, som kun i meget ringe grad bliver udnyttet, den geografiske reference. Alle sensorer er placeret et eller andet sted i den virkelige verden og næppe tilfældigt. Umiddelbart vil denne information være af stor værdig, fordi den nødvendigvis må være ganske unik. Uheldigvis vil en rumlig koordinat ikke være umiddelbar forståelig selv for en fagperson. Her er det så vores traditionelle tilgang til geodata kommer ind i billedet, bogstaveligt talt, idet der ikke er noget i vejen for, at de interaktive systemer tager udgangspunkt i et grundkort og også 3D modeller til at skabe overblik og til at vise den information, der skabes ude i virkeligheden. Interaktionsfladen vil umiddelbart give mening for de fleste og vil kunne danne et glimrende grundlag for styring og samkøring af data. Også de traditioner vi har med at tilknytte attributdata til geodata vil have stor værdi. Det vil således yderligere være muligt i nogle sammenhænge at benytte gængse GIS analysemetoder til at frembringe nyttig information. I (Kjems, Kolar, & Batty, 2005) kaldte vi dem for modelmaps.

Jeg mener, det er oplagt at udvide vore fælles FOT-data til også at kunne omfatte andet, end det vi bruger dem til i dag. Jeg vil gå så vidt som at sige, at hvis ikke vi får en fælles FOT-standard for de geodata, der beskriver den rumlige struktur, vil det på sigt blive meget dyrt at holde styr på den nødvendige samkøring af geodata, uden dog at ville sætte en størrelse på beløbet. Udviklingen af det digitale kortværk som for alvor startede i 80'erne havde en primus motor i gasleverandørerne, som skulle lægge gasledninger i jorden i Danmark. Det digitale kortværk er blevet suppleret, fornyet og forædlet mange gange siden og ligger nu tilgængelig i en fælles standard - 30 år senere. 3D data er langt mere komplice-

ret og forefindes i adskillige matematiske beskrivelser og datarepræsentationer. Hvis vi lader stå til, vil der komme mange forskellige systemer og pseudo objektstandarder, som vil koste dyrt på sigt især for det offentlige at køre sammen.

Der bliver produceret store mængder 3D data af by og landskab, og der udvikles hele tiden forbedrede metoder til at øge kvaliteten herunder den geografiske præcision i de data, der bliver genereret. Men der sker kun meget lidt på datarepræsentationen. Hvordan skal data indgå i vore systemer? Her tænker jeg ikke på visualiseringsdelen. Den er der rigelige muligheder for at håndtere på en hensigtsmæssig måde. Nej, det er geoobjektet som interaktiv informationsbærer. Indenfor by og landskab er CityGML p.t. bedste bud på en beskrivelse, der på den ene side indeholder muligheden for udveksling af data på en struktureret måde med mulighed for at videreføre semantisk information og som samtidig benytter en skalerbar objektstruktur, der kan håndtere geografien på en elegant måde. CityGML løser dog ikke umiddelbart problemet med at håndtere objekterne som informationsbærer i systemerne, men vil være et godt skridt på vejen.

Så måske bør vi starte med at få lagt hele landet ud i en CityGML udgave og se på, hvordan data hertil kan forædles med tiden. GML er hierarkisk opbygget, og der er principielt ingen grænser for de detaljer, man vælger at inddrage. Begrænsninger ligger for øjeblikket i "City"-delen og de definitioner, der er vedtaget her. Der er i dag en grænseflade mellem IFC definitionen og CityGML, og sådanne vil der komme flere af, men alt andet lige bør FOT data være en del af dette og danne et udvidet grunddatasæt indeholdende en rumlig beskrivelse, som kan benyttes til interaktive systemer foruden visualisering og meget andet.

Alternativet

På faldrebet vil jeg nævne et alternativ, som dog vil kræve en artikel mere at beskrive uddybende. I Center



for 3D GeoInformation på AAU har vi i perioden 2008-2012 deltaget i et forskningsprojekt ved navn InfraWorld finansieret af Norges Forskningsråd, som tog udgangspunkt i den datamodel, der har været under udvikling i centret siden 2006. Slutproduktet var en brugergrænseflade bestående af en rumlig bymodel, som kunne kobles til realtidsinformationer med data fra den virkelige verden. Omdrejningspunktet var systemets objektbeskrivelse, der giver en løsning på alle de problemstillinger, der blev nævnt ovenfor, og lidt til – intet mindre. Forestil dig at modtage en bunke virtuelle æsker med et ukendt indhold. Du henter æskerne ind i systemet, f.eks. en webbrowser, hvorved alle æskerne aktiveres og begynder at pakke sig selv ud og give sig til kende f.eks. som et hus. Alle dele, der kommer ud af æskerne, er geografisk refereret og bliver lagt på de korrekte steder. Når huset er på plads, begynder det at kommunikere med det virkelige hus, det repræsenterer, og henter f.eks. aktuelle forbrugsdata. Der er nærmest uendelig mange anvendelsesmuligheder, og i projektet blev nogle stykker afprøvet.

REFERENCES

- Atzori, L., Iera, A., & Morabito, G. (2010). *The Internet of things: A survey*. *Computer Networks*, 54(15), 2787–2805.
- Batty, M., Chapman, D., Evans, S., Haklay, M., Kueppers, S., Shiode, N., ... Klosterman, R. E. (2001). *Visualizing the City: Communicating Urban Design to Planners and Decision Makers*. Redlands: ESRI Press.
- Big Data Definition. (n.d.). *MIKE2.0, the open source methodology for Information Development*. Retrieved March 14, 2014, from http://mike2.openmethodology.org/wiki/Big_Data_Definition
- Flemming, L., Schack Madsen, P., Sørensen, M., Lindeneq Johansen, R., Nielsen, T., Bodum, L., ... Hjorth, F. (2011). *Geoforums vejledning i 3D-bymodeller* (p. 105).
- IBM. (n.d.). *IBM Intelligent Operations Center*. Retrieved from <http://www-03.ibm.com/software/products/en/intelligent-operations-center/>
- Kjems, E., Kolar, J., & Batty, S. E. (2005). *From Mapping to Virtual Geography*. In *Proceedings of CUPUM 2005*. London.
- Mashey, J. R. (1997). *Big Data and the Next Wave of Infra-*
- Stress*. In *Computer Science Division Seminar*, University of California, Berkeley.
- Middleton, P., Kjeldsen, P., & Tully, J. (2013). *Forecast: The Internet of Things*, Worldwide, 2013.
- Morton, P. J., Horne, M., Dalton, R., & Thompson, E. M. (2012). *Virtual City Models: Avoidance of Obsolescence. Education and Research in Computer Aided Architectural Design in Europe-eCAADe*, Prague, Czech Republic.
- Sarma, S., Brock, D. L., & Ashton, K. (2000). *The networked physical world*. Auto-ID Center White Paper MIT-AUTOID-WH-001.
- Schielder, R. (2014). *Hvad nu FOTdanmark?* GEOFORUM. Retrieved February 28, 2014, from <http://www.geoforum.dk/GEOFORUM-151-10612.aspx>
- Tufte, E. R. (1990). *Escaping Flatland*. *Envisioning Information*, 12–36.
- Waltz, E., Llinas, J., & others. (1990). *Multisensor data fusion* (Vol. 685). Artech house Boston.
- Weiser, M. (1991). *The computer for the 21st century*. *Scientific American*, 265(3), 94–104.