
RAPPORT

Internationell jämförelse av transportmodeller och sekretessaspekter UPPDRAGSNUMMER 12601133



SLUTVERSION

2018-03-05

STRATEGISK JÄRNVÄG

Sweco Society AB

Ändringsförteckning

Uppdragsnummer: 12601133

Kund: Trafikanalys

Kontaktpersoner: Eva Lindborg

Organisation: Henrik Andersson, uppdragsledare

Linda Ramstedt, expert

Henrik Edwards, expert

Kajsa Lundborg, utredare

| Ver. | datum | | Granskad | Godkänd |
|------|------------|--------------------|---------------------------------------|---------|
| 1.0 | 2018-02-16 | Granskningsversion | Uppdragsledare | |
| 2.0 | 2018-02-23 | Granskningsversion | Uppdragsledare, Sweco intern, kund | |
| | 2018-03-05 | Slutversion | Uppdragsledare | |
| | | | | |

Sammanfattning

Trafikanalys har anlitat Sweco för att studera hur andra länder hanterar avvägningen mellan primärt företagets önskan om sekretess och behovet av att kunna använda detaljerade data över järnvägstransporter i beslutsunderlag. Fokus ligger på möjligheterna att validera och kalibrera modeller för prognostisering och analys av järnvägstransporter på detaljerad nivå.

Arbetet har drivits som en kombination av litteraturstudier och intervjuer med personer med god kunskap om hur transportmodeller används i deras respektive länder. De studerade länderna är Norge, Nederländerna, Danmark och Tyskland. Vidare har olika aktörers perspektiv på sekretessen i Sverige behandlats.

Analysen har gjorts utifrån en struktur för hur data som behövs för transportmodeller kan hanteras i en kedja från definition till insamling, överlämning, lagring, skattning, kalibrering och validering samt resultatuttag samtidigt som sekretessen bevaras och en tillräcklig detaljeringsnivå och transparens kan säkerställas.

Arbetet har bedrivits under en relativt kort period och med begränsad budget. Som en följd av detta har endast en kategori av brukare intervjuats i respektive land, även om intervjuerna kompletterats med olika typer av skriftlig dokumentation. Därför kan det inte uteslutas att det finns andra aspekter relaterat till sekretess och transportmodeller i de studerade länderna som inte kommit fram i studien.

Norge

Norge arbetar i hög grad med insamling av data med stöd av uppgiftsskyldighet. Samtidigt garanteras sekretessen i lagstiftning som i hög grad påminner om bland annat den svenska. Metoder som används för att säkerställa att enskilda personer inte kan identifieras är bland annat anonymisering och avidentifiering men också vanliga statistiska metoder för röjandekontroll. En intressant aspekt är att data som Jernbanedirektoratet samlar in från operatörerna och redovisar på en aggregerad nivå publiceras med högre detaljeringsgrad av Statistiska Sentralbyrån med hänvisning till "avtal med operatörerna". Detta avtal verkar inte finnas i skriftlig form, men kan ha anammats som praxis av branschen. Tillgång till grunddata ges framför allt för ändamål som kan definieras som forskning. Detta verkar ge en tillräckligt god tillgänglighet för att i synnerhet forskare skall kunna driva ett utvecklingsarbete av transportmodeller.

Nederländerna

Nederländerna kombinerar datainsamling av Statistics Netherlands och Dutch Railways med frivilliga överenskommelser med andra operatörer för att skapa exempelvis OD-matriser på detaljerad nivå. Det är särskilt intressant att man på så sätt får tillgång till detaljerade data för kalibrering av persontransportmodellen. I gengäld får alltså operatörerna tillgång till den nationella persontransportmodellen, vilket bidrar till den transparens som utgör en viktig grund för väl fungerande marknader.

Intressant är också att man avsiktligt väljer att bygga BASGOEDS-modellen på en aggregerad nivå, inte primärt av sekretessskäl men för att få en enklare hantering och

större transparens. Det skall samtidigt betonas att den geografiska tätheten utgör en viktig faktor här. När Nederländerna delas in i 40 zoner är varje zon i genomsnitt drygt 1 000 km². När Norge delas in i 90 zoner är varje zon drygt fyra gånger större. Skillnaden förstärks genom att den ekonomiska aktiviteten i varje nederländsk zon är större. Sammantaget blir zonindelningen i Nederländerna mer detaljerad ur ett geografiskt perspektiv, samtidigt som möjligheten att urskilja enskilda företag blir mindre. Detta illustrerar hur de givna premisserna för en nation kan påverka frågan om sekretess kopplat till transportmodeller.

Statistics Netherlands samlar in data med stöd i lagen och skyddar uppgifterna på olika sätt, exempelvis genom statistisk röjandekontroll. Organisationer som vill använda mikrodata måste ansöka om detta. Det har inte framgått i detta arbetet exakt hur de ansökningarna värderas. På det hela taget verkar dock möjligheten att få tillgång till data vara relativt god, vilket grundas på att den intervjuade organisationen är ett privat, kommersiellt forskningsinstitut. Däremot kan det valda sättet att säkerställa sekretessen rent praktiskt, genom att data endast finns tillgängligt på Statistics Netherlands hemsida, orsaka problem i form av ineffektiva arbetsprocesser.

Danmark

I Danmark omfattas många data som behövs för transportmodellerna av olika restriktioner kopplade till sekretess. Det framstår dock som att i alla fall Danmarks Tekniska Universitet lyckas hantera dessa genom en mångfald av lösningar, från lagstiftning till frivilliga överenskommelser i det specifika fallet. Behovet av sekretess verkar därför inte påverka den upplevda användbarheten av Landstrafikmodellen, i alla fall inte enligt DTU. I en genomgång av publikationer från de senaste sju åren på Transport-, Byggnings- och Boligministeriets hemsida verkar inte frågan om sekretess behandlas explicit. Det tas inte heller upp som någon direkt begränsning i dokumentationen till den dominerande Landstrafikmodellen.

DTU lyfter en intressant aspekt av sekretessen, det som gäller frågor om immaterialrätt. Detta är en fråga som möjligen har varit extra relevant i Danmark, där utvecklingen av den Landstrafikmodellen har gjorts i samarbete mellan Danmarks Tekniske Universitet och flera konsultbolag. Det är dock inte någon unik situation för Danmark och sannolikt har frågan reglerats på sedvanligt vis som i gemensamma utvecklingsprojekt generellt.

Tyskland

Arbetet med transportmodeller i Tyskland försvåras av att staten har valt att ta ett begränsat ansvar för utvecklingen av modellerna. Det gör det också svårt att få en tydlig bild av sekretessaspekter. Data finns hos olika aktörer och kan skyddas såväl av offentlig sekretess och affärssekretess. Det mest konkreta som kommit fram i detta arbete är den lagstadgade skyldigheten för järnvägsoperatörer att presentera motsvarande OD-matriser för såväl passagerare som gods, vilket tycks utgöra del av redovisningen till EU. Det är oklart vilken detaljeringsgrad som kan erhållas vid uttag från matriserna.

Det är också notabelt att det pågår ett departementsarbete för att göra det möjligt för forskare att få tillgång till mikrodata.

Sverige

Det främsta skälet att inkludera en beskrivning av Sverige är att jämföra hur forskare (representerade av VTI), Trafikverket samt konsulter upplever sekretessaspekter kopplat till järnvägsdata och transportmodeller.

Det framgår tydligt att i alla fall Trafikverket anser att olika sekretesskrav, framför allt i form av affärshemligheter, inverkar menligt på arbetet med transportmodellerna. VTI verkar inte uppleva problem i samma utsträckning. Det har dock varit svårt att inom ramen för detta arbete identifiera vad skillnaden beror på. En hypotes är att VTI som forskningsinstitut har lättare att komma åt data (jämför Norge). En annan förklaring kan vara att VTI gör andra typer av analyser som kan ligga närmare forskning och/eller ställa krav på andra data. Vid en översiktlig genomgång av aktuella lagar och förordningar (se referenslistan) framgår inget tydligt svar. SCB anger i sin sekretesspolicy att data kan lämnas ut till olika organisationer för statistik- och utredningsändamål, men att det är villkorat att enskild person inte riskerar att lida skada.

När VTI exempelvis ansöker om data från Eurostat framgår det tydligt att alla typer av organisationer kan ansöka, så länge det är en forskningsenhet som står för ansökan. Av detta följer två observationer eller kriterier för att Trafikverket skall kunna få tillgång till data från Eurostat. Det första gäller om Trafikverkets utvecklingsarbete av transportmodeller kan ses som forskning. Det andra är frågan om Trafikverket är organiserat på ett sådant sätt att man kan hävda att det är en forskningsenhet som gör ansökan.

Vidare kan man konstatera att Eurostat verkar ha besvärande långa handläggningstider och att man samtidigt begär en utförlig beskrivning av hur datas skall skyddas, dock utan att ange några kravnivåer.

Analys och slutsatser

Det finns två grundläggande utmaningar kopplat till sekretess när det gäller järnvägsdata och transportmodeller. För det första måste data produceras på en tillräckligt detaljerad nivå för att möjliggöra kalibrering och validering av modellerna. För det andra måste data göras tillgängligt för de berörda brukarna på en tillräckligt detaljerad nivå.

Det finns inte oväntat ett tydligt, omvänt samband mellan de två utmaningarna. Ju mer detaljerade data, desto större restriktioner på användningen.

I tabell A redovisas olika utmaningar kopplat till både sekretess och andra aspekter som identifierats i arbetet. Tabellen gör inte anspråk på att vara uttömmande.

Tabell A. Identifierade utmaningar.

| Moment | Utmaningar rörande sekretess | Andra utmaningar |
|-----------------------------------|---|---|
| Identifiering av databehov | - | <ul style="list-style-type: none"> Databehovet beror av vilka modeller som används och kan ändras med nya modeller |
| Datainsamling | <ul style="list-style-type: none"> Att få i synnerhet kommersiella aktörer att leverera potentiellt affärskänsliga data (Samtliga) | <ul style="list-style-type: none"> Insats för respondenter i form av arbete, kostnad (Norge, Sverige) Osäkerheter kring metod och kvalitet (Norge, Sverige) |
| Överlämning av data | <ul style="list-style-type: none"> Syftet med undersökningen definierar användningen (Nederländerna, Norge, Sverige) Byråkratisk och långsam rekvisering av data (EU) | <ul style="list-style-type: none"> Data ägs av de företag som samlat in dem (Tyskland) |
| Lagring av data | <ul style="list-style-type: none"> Data får bara lagras på den ansvariga myndighetens server, vilket gör arbete mer ineffektivt (Nederländerna, Sverige) Särskilda krav ställs på datasäkerhet (EU) | <ul style="list-style-type: none"> En fragmenterad lagring av data hos olika aktörer (Tyskland) |
| Nulägesbeskrivning | <ul style="list-style-type: none"> Beslutsfattare efterlyser detaljerade uppgifter om transportdata som kan avslöja enskilda personer (Norge, Sverige) | <ul style="list-style-type: none"> När data baseras på urvalsundersökningar kan felmarginalerna bli för stora (Norge, Tyskland) |
| Skattning/estimering | <ul style="list-style-type: none"> Modellernas algoritmer och annan kunskap skyddas av immaterialrättigheter (Danmark, Tyskland) Användningen begränsas till exempelvis forskare eller forskning. (Finland, Norge, Sverige) | <ul style="list-style-type: none"> Det kan finnas fördelar med en mer aggregerad modelluppbyggnad i form av enkelhet och transparens (Nederländerna) |

| Moment | Utmaningar rörande sekretess | Andra utmaningar |
|-----------------------------------|---|---|
| Kalibrering och validering | <ul style="list-style-type: none"> • Sekretess begränsar möjligheten att kalibrera och validera (Sverige) • Det är inte möjligt att ange överensstämmelsen mellan modell och verklighet (Sverige) | <ul style="list-style-type: none"> • Lagstiftning om rapportering anger inte krav att rapportera på detaljerad sträcka (Danmark, Norge, Sverige) |
| Resultatuttag | <ul style="list-style-type: none"> • Enskilda fysiska och juridiska personer skall inte kunna identifieras (Nederländerna, Norge, Sverige) • Önskan om högre detaljeringsgrad i redovisningen (Norge) | <ul style="list-style-type: none"> • Kvalitetsaspekter vid allt för detaljerade data vid urvalsundersökningar (Norge, Tyskland, Sverige) |

Under arbetet har ett antal olika metoder för att hantera kraven på sekretess identifierats. Nedan beskrivs dessa i relation till de grundläggande utmaningarna; att säkerställa att *detaljerade data produceras* och att *data finns tillgängliga för utveckling av transportmodeller*. Följande metoder tas upp:

- Uppgiftsskyldighet, frivilliga överenskommelser eller avtal
- Begränsad men tillräcklig tillgång till detaljerade data

Många av metoderna förekommer i flera länder, se tabell B.

Tabell B. Metoder för att hantera behovet av sekretess och samtidigt tillgodose krav på statistik om järnvägstransporter. Det garanteras inte att tabellen är uttömmande.

| Metod | Norge | Nederländerna | Tyskland | Danmark | Sverige |
|--|---|---|---|--|-------------------------------|
| Uppgiftsskyldighet | X | X | X | X | X |
| Frivilliga överenskommelser | | X (med operatörer) | | X | X |
| Avtal | X | | X (med operatörer om OD-matriser) | X (med operatörer om resandestatistik) | |
| Begränsningar i tillgång till grunddata - VEM | X (organisation) | X (syfte, specifikt projekt) | (arbete pågår - förslag att ge forskare tillgång till data) | | |
| Begränsningar i tillgång till grunddata - VARFÖR | X (syfte) | X (syfte, specifikt projekt) | | X (syfte) | |
| Begränsningar i tillgång till grunddata - HUR | X (praktisk hantering av data efter mottagande inkl radering) | X (behandling av data på myndighetens server) | | | X (åtkomst till data hos SCB) |

Samtliga studerade länder drar nytta av uppgiftsskyldighet för att säkerställa tillgång till data, även om data samtidigt omfattas av sekretesslagstiftningen. Nederländerna kompletterar systematiskt med frivilliga överenskommelser, där operatörerna får tillgång till transportmodellen i utbyte mot att man avslöjar data. Uppgifter från Danmark visar att frivilliga överenskommelser sluts per projektbaserat. I Sverige har de genomförts för att få tillgång till detaljerade resandestatistik, men metodmässiga utmaningar rörande bland annat täckning framhålls som en utmaning. I flera länder tillämpas också avtal med operatörerna för att säkerställa redovisning av data. I Tyskland och Danmark är det reglerat som ett grundkrav för operatörerna i samband med att de skall få trafikeringsrätt.

De studerade länderna använder också olika tekniker för att begränsa tillgängligheten till data, för att på så sätt kombinera dataskydd med nödvändig öppenhet för forskning och framtagande av statistik. Begränsningar sker utifrån någon eller en kombination av *vem* som önskar få tillgång till data, *varför* samt *hur* man får tillgång till uppgifterna. När det gäller vem som får tillgång till data har universitet/högskolor och forskningsinstitut typiskt lättare att få tillgång till data. Likaså ger syften som forskning eller forskningsnära verksamhet ofta enklare tillgång till data. Avseende hur man får tillgång till data tillämpar Nederländerna och Sverige ibland begränsningen att data endast finns tillgänglig på den nationella statistikbyråns server. Dessutom är utlämningen av data ofta förknippat med tydliga praktiska restriktioner, vilket bland annat illustrerats av Eurostats ansökanprocess som beskrivits i detta arbete.

Uppgiftsskyldighet har de bästa förutsättningarna att få in kompletta uppgifter enligt tydliga, enhetliga metoder. Den tydligaste nackdelen är den insats som krävs från uppgiftslämnarna och den insamlade organisationen. En nackdel är att förståelsen för uppgiftsskyldigheten kan bli lidande om motiven inte förklaras.

Frivilliga överenskommelser kan appliceras i Sverige och gör det redan för exempelvis hamnstatistik, men riskerna med tillvägagångssättet är stora vad avser viktiga kvalitetsaspekter, att säkra tillgång till data över tid samt att få täckande undersökningar. För att öka villigheten att rapportera data kan något erbjudas i gengäld, exempelvis tillgång till transportmodellerna eller marknadsdata.

Det övergripande intrycket från de studerade länderna är att sekretesskraven inte verkar innebära några allvarliga begränsningar i arbetet med transportmodeller. I arbetet har åtminstone fyra tänkbara förklaringar förts fram:

- De etablerade metoderna för att säkerställa att data både produceras och görs tillgängliga fungerar tillräckligt bra.
- Eftersom sekretesslagstiftningen har funnits under förhållandevis lång tid och dessutom är till stora delar internationellt implementerad genom bland annat EU, har ambitionerna för transportmodeller anpassats till vad som är möjligt med hänsyn till säkerställandet av sekretessen.
- Andra typer av utmaningar överskuggar sekretessaspekterna. Det gäller bland annat den statistiska osäkerheten för detaljerade data vid urvalsundersökningar.
- Fördelar med att använda mer aggregerade transportmodeller gör sekretessfrågan mer sekundär. Det gäller bland annat värdet av transparens/öppenhet och enklare uppdateringar.

Det är fruktbart att kombinera olika typer av åtgärder för att säkerställa ett gynnsamt utvecklingsarbete av transportmodellerna. I Sverige framstår det som en möjlighet att öka samarbetet med operatörer för att få tillgång till passagerardata på specifika sträckor, vilka kan användas för detaljerad kalibrering och validering. Här finns möjligheter att lära av pragmatismen i Danmark och det gemensamma intresset i Nederländerna, manifesterat i ett gemensamt ägande av den nationella transportmodellen.

Innehållsförteckning

| | | |
|----------|---------------------------------------|-----------|
| 1 | Inledning | 12 |
| 1.1 | Bakgrund | 12 |
| 1.2 | Metod | 12 |
| 1.3 | Avgränsning | 13 |
| 2 | En struktur för dataflöde | 13 |
| 2.1 | Identifiering av informationsbehovet | 14 |
| 2.2 | Insamling av data | 14 |
| 2.3 | Överlämning av data | 15 |
| 2.4 | Lagring av data | 15 |
| 2.5 | Nulägesbeskrivning | 15 |
| 2.6 | Skattning/estimering | 16 |
| 2.7 | Kalibrering | 16 |
| 2.8 | Validering | 17 |
| 2.9 | Resultatuttag | 17 |
| 3 | Beskrivning: Norge | 18 |
| 3.1 | Databehov | 18 |
| 3.2 | Datakällor och insamling | 19 |
| 3.3 | Utmaningar och hantering av sekretess | 21 |
| 3.4 | Sammanfattning | 23 |
| 4 | Beskrivning: Nederländerna | 23 |
| 4.1 | Databehov | 23 |
| 4.2 | Datakällor och insamling | 23 |
| 4.3 | Utmaningar och hantering av sekretess | 24 |
| 4.4 | Sammanfattning | 24 |
| 5 | Beskrivning: Danmark | 25 |
| 5.1 | Databehov, datakällor och insamling | 25 |
| 5.2 | Utmaningar och hantering av sekretess | 26 |
| 5.3 | Sammanfattning | 27 |
| 6 | Beskrivning: Tyskland | 28 |
| 6.1 | Databehov, datakällor och insamling | 28 |
| 6.2 | Utmaningar och hantering av sekretess | 29 |
| 6.3 | Sammanfattning | 29 |

10 (48)

| | | |
|----------|---|-----------|
| 7 | Beskrivning: Sverige | 30 |
| 7.1 | Databehov, datakällor och insamling | 30 |
| 7.2 | Utmaningar och hantering av sekretess | 31 |
| 7.3 | Sammanfattning | 33 |
| 8 | Analys och slutsatser | 34 |
| 8.1 | Utmaningar | 34 |
| 8.2 | Lösningar | 36 |
| 8.2.1 | Uppgiftsskyldighet, avtal eller frivilliga överenskommelser | 38 |
| 8.2.2 | Begränsa tillgängligheten till detaljerade data | 41 |
| 8.3 | Avslutande konklusioner | 43 |
| | Referenser | 45 |
| | Bilaga: Frågeformulär (svensk version) | 47 |

1 Inledning

1.1 Bakgrund

Inom ramen för regeringsuppdraget "Uppdrag att utreda hur kunskapen om järnvägstransporter kan förbättras" har Trafikanalys anlitat Sweco för att studera hur andra länder hanterar avvägningen mellan primärt företagets önskan om sekretess och behovet av att kunna använda detaljerade data över järnvägstransporter i beslutsunderlag. Fokus ligger på möjligheterna att validera och kalibrera modeller för prognostisering och analys av järnvägstransporter på detaljerad nivå. Utifrån hur studerade länder hanterar frågan ska slutsatser dras och möjliga lösningar att hantera sekretess och transparens i en svensk kontext diskuteras.

1.2 Metod

Arbetet har drivits som en kombination av litteraturstudier och intervjuer med personer med god kunskap om hur transportmodeller används i deras respektive länder. Intervjuerna har gjorts såväl via telefon som skriftligt.

Urvalet av länder har skett utifrån att de skall ha en väl utvecklad kultur och tradition av att använda transportmodeller för olika samhällsytten, men också baserat på närhet definierat utifrån nätverk hos Swecos och till viss del Trafikanalys medarbetare. Följande länder har studerats: Norge, Tyskland, Danmark och Nederländerna. Vidare har dialog förts med Finland och Väg- och Transportforskningsinstitutet i Sverige. Riktlinjer från EU har inkluderats. Dessutom har det svenska, övergripande systemet för officiell statistikproduktion studerats.

De intervjuade organisationerna i respektive land framgår av tabell 1.

Tabell 1. Intervjuade organisationer i arbetet.

| Land | Organisation | Typ av organisation |
|---------------|---|-----------------------------|
| Danmark | Danmarks Tekniske Universitet | Universitet |
| Finland | Trafikverket | Myndighet |
| Nederländerna | Significance | Privat forskningsinstitut |
| Norge | Transportøkonomisk Institutt | Privat, ideell stiftelse |
| Sverige | VTI, Statens Väg- och Transportforskningsinstitut | Statligt forskningsinstitut |
| Tyskland | DLR, Deutsche Zentrum für Luft- und Raumfahrt (DLR) | Federalt forskningscentrum |

Baserat på litteratur och intervjuer har analysen fokuserat på att skapa en struktur för att se hur data som behövs för transportmodeller kan hanteras i en kedja från definition till

insamling, överlämning, lagring, modellering och resultatuttag samtidigt som sekretessen bevaras och en tillräcklig detaljeringsnivå och transparens kan säkerställas. Strävan har varit att i görligaste mån klassificera eller sortera metoder som används i de studerade länderna utifrån en funktionell struktur eller systematik.

Arbetet har drivits i nära dialog med Trafikanalys för att säkerställa att material har varit relevant för det aktuella regeringsuppdraget.

1.3 Avgränsning

Denna studie fokuserar till största delen på sekretessbehovet för juridiska personer. Anledningen är att transportmodeller inte kan anses vara beroende av uppgifter om enskilda fysiska personer. Det kan dock inte uteslutas att man vid exempelvis insamling av data över resande kan komma att behandla persondata som kan kopplas direkt till enskilda individer. Hanteringen av sådan data regleras av personuppgiftsagen (PuL) som den 25 maj 2018 ersätts av Dataskyddsförordningen (GDPR).

Däremot har sekretessfrågor i vid bemärkelse inkluderats i studien, inklusive generella bestämmelser för att skydda enskilda personers identitet. Syftet med detta har varit att undersöka om det finns metoder och tekniker som kan vara användbara även för hanteringen av de uppgifter som behövs i transportmodeller.

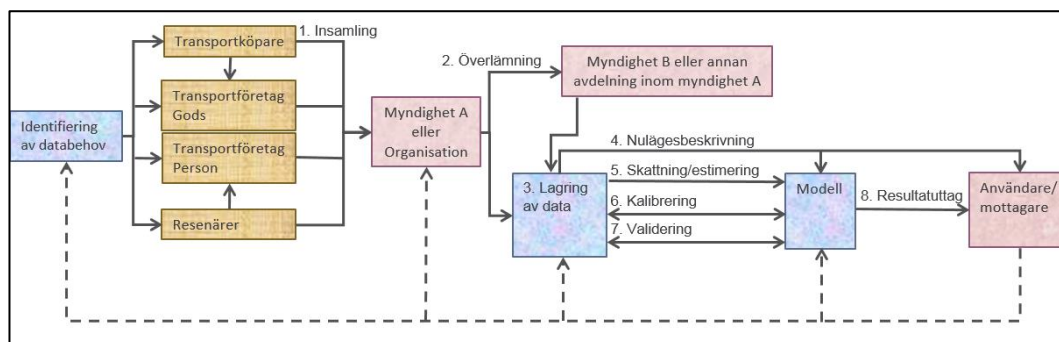
Arbetet har bedrivits under en relativt kort period och med begränsad budget. Som en följd av detta har endast en kategori av brukare intervjuats i respektive land, även om intervjuerna kompletterats med olika typer av skriftlig dokumentation. Därför kan det inte uteslutas att det finns andra aspekter relaterat till sekretess och transportmodeller i de studerade länderna som inte kommit fram i studien.

Likaså är urvalet av länder ganska litet och dessutom med fokus på länder i vårt närområde, där övergripande riktlinjer för sekretessaspekter till stor del styrs av riktlinjer framtagna av EU.

Även om en del lagar och förordningar citeras i rapporten är syftet med uppdraget inte att värdera hur exakta formuleringar i dessa påverkar möjligheterna att använda data.

2 En struktur för dataflöde

I figur 1 redovisas en övergripande struktur för hur data hanteras för framtagande av och användning i transportmodeller samt även statistiska beskrivningar av nuläget. Det finns starka samband mellan de olika stegen och även från de tidiga momenten till de avslutande. Som exempel på det senare kan nämnas att huruvida data som samlas in tillhör den officiella statistiken eller inte har stor påverkan på sekretesskraven under övriga delar av processen.



Figur 1. Dataflöde från identifiering av behov till resultatuttag.

2.1 Identifiering av informationsbehovet

Det första steget är att definiera informationsbehovet, i det här fallet primärt utifrån det som behövs för transportmodellerna. I princip kan en transportmodell beskrivas som matematiska eller logiska samband som använder indata för att presentera ett resultat. Indata kan handla om hur förutsättningarna för resor och transporter ser ut i form av exempelvis befolkning, arbetsställen, infrastruktur och trafikering. Resultaten är ofta framtida transportflöden med högre eller lägre detaljeringsgrad. I arbetet med att bygga upp en modell måste man alltså göra analyser för att beskriva de matematiska och logiska sambanden och dessutom måste man kalibrera sambanden och validera resultaten som modellen ger. En princip för validering är att denna skall göras mot uppgifter som inte används som indata i modellen eller som utgör grund för sambanden.

Det önskvärda informationsbehovet måste ställas mot aspekter som resursbehov för de som skall samla in och förvalta data, arbetsbördan och kostnader för de som skall lämna uppgifterna och även känsligheten i uppgifterna.

2.2 Insamling av data

När databehovet har fastställts skall data samlas in. Här finns ett behov av att säkerställa att i synnerhet juridiska personer är beredda att lämna ifrån sig uppgifterna. Den mest långtgående möjligheten att garantera detta är genom lagstiftning och uppgiftsskyldighet. I andra ändan av skalan har man frivilliga överenskommelser och någonstans mitt emellan dessa finns juridiskt bindande avtal.

Lagstiftning garanterar att data tas fram av uppgiftslämnarna, men samtidigt medföljer ett omfattande sekretesskrav. I Offentlighets- och sekretesslagen (2009:400) anges grundförutsättningen för sekretess vid framtagande av statistik inom myndigheter:

8 § Sekretess gäller i sådan särskild verksamhet hos en myndighet som avser framställning av statistik för uppgift som avser en enskilds personliga eller ekonomiska förhållanden och som kan hänföras till den enskilde.

Motsvarande sekretess gäller i annan jämförbar undersökning som utförs av Riksrevisionen, av riksdagsförvaltningen, av Statskontoret eller inom det statliga kommittéväsendet. Detsamma gäller annan jämförbar undersökning som utförs

av någon annan myndighet i den utsträckning regeringen meddelar föreskrifter om det.

Uppgift som behövs för forsknings- eller statistikändamål och uppgift som inte genom namn, annan identitetsbeteckning eller liknande förhållande är direkt hänförlig till den enskilde får dock lämnas ut, om det står klart att uppgiften kan röjas utan att den enskilde eller någon närstående till denne lider skada eller men. [...] ¹

I Offentlighets- och sekretessförordningen (2009:641) konkretiseras det vidare:

9 § Sekretess gäller, [...], i statliga myndigheters verksamhet, som består i utredning, planering, prisreglering, tillståndsgivning, tillsyn eller stödverksamhet med avseende på produktion, handel, transportverksamhet [...]

1. för uppgift om en enskilds affärs- eller driftförhållanden, uppfinningar eller forskningsresultat, om det kan antas att den enskilde lider skada om uppgifterna röjs, [...] ²

2.3 Överlämning av data

Överlämning av data mellan och även inom organisationer regleras av Offentlighets- och sekretesslagen (2009:400). En vanlig grundprincip är att sekretessen följer med data. Likaså styrs användning och överlämning av data ofta utifrån ändamålet som uppgavs vid insamlingen av uppgifterna, vilket kan försvåra överlämning av uppgifter som samlats in för exempelvis marknadsövervakning till uppgifter som rör statistikproduktion. ³

Det kan ställa till problem och/eller frustration när exempelvis en del av en myndighet samlar in uppgifter för den officiella statistiken och samtidigt inte kan använda uppgifterna för den egna verksamhetsuppföljningen. En kommun kan i sådana lägen råka ut för situationen att när man begär ut data från Statistiska Centralbyrån så redovisas inte data på samma detaljeringsnivå som kommunen själv samlat in. ⁴

2.4 Lagring av data

Lagring av data kan regleras på olika sätt, men utgör framför allt ett verktyg för att säkerställa sekretessen. Det kan handla om att data skall anonymiseras före lagring eller att uppgifter lagras av den statistikansvariga myndigheten och att brukare endast får tillgång till databasen genom exempelvis en hemsida med inloggning.

2.5 Nulägesbeskrivning

Ett vanligt syfte med datainsamling inom transportsektorn är att producera statistik över ett nuläge. Det kan gälla både flöden av gods och passagerare som görs tillgängliga för en användare/mottagare och förutsättningar för transportmodellen. Presentationen av

¹ Offentlighets- och sekretesslagen (2009:400), 24 kap. 8§

² Offentlighets- och sekretessförordningen (2009:641) 9 §

³ Statistiska Centralbyrån (2016), sid 3

⁴ Tindrebäck, Silke (2018)

nulägesbeskrivningar kan genomgå statistisk röjandekontroll för att säkerställa att uppgifter inte kan kopplas till enskilda personer. Som en följd av kontrollen kan data exempelvis presenteras på en aggregerad nivå med lägre detaljeringsgrad.⁵ Det skall betonas att det inte bara är sekretessen som styr publiceringen av detaljerade data, utan i hög grad också kvalitetsaspekter. När data härstammar från en urvalsundersökning kan den statistiska variansen vara större än medelvärdet om relationen är för liten och/eller detaljerad. Detta lyfts fram bland annat i dialogen med brukare i Tyskland.⁶

En utmaning med röjningskontrollen är att den inte differentierar mellan hur känsliga olika data är för uppgiftslämnarna.

2.6 Skattning/estimering

Skattning (estimering) innebär att man bygger upp sambanden mellan förutsättningar och resultat i modellen. Dessa kan bestämmas matematiskt eller logiskt. Med matematiskt avses exempelvis hur människor kan förväntas välja transportmedel utifrån den generaliserade reskostnaden som innebär att samtliga uppofringar som en resenär gör i form av tid och kostnader summeras för olika reskedjor. Det kan också gälla hur efterfrågan påverkas när restid, pris eller turtäthet förändras. Logiska samband kan gälla exempelvis begränsningar i infrastrukturen gällande kapacitet.

Skattning görs vanligen utifrån urvalsbaserade undersökningar av resenärers eller företags faktiska eller angivna beteenden, som resvaneundersökningar. Det är enklare att modellera hur resenärer agerar än hur företag väljer sina logistiska lösningar.⁷

Sekretess kan begränsa tillgången till data som behövs för att göra skattningarna.

2.7 Kalibrering

När man har skattat/estimerat sambanden kalibreras modellen. Syftet är att se om modellen beter sig som den är tänkt och programmerad att bete sig. I princip innebär detta att man observerar hur modellen reagerar på förändringar i indata och om resultaten som produceras är rimliga. Utgångspunkten för denna kontroll är verklighetsdata samt kunskaper, erfarenheter och kännedom om trafikprocesser.⁸

Exakt vilka data som används vid kalibrering och hur kalibreringen går till beror på modellen. I exempelvis den svenska godstransportmodellen Samgods sker kalibrering framför allt för att modellresultaten skall återspegla transportstatistiken. Med kalibreringen vill man i första hand fånga fördelningen av transportflödet på olika transportslag och även få en rimlig överensstämmelse på fordonstyper och varugrupper. I Samgods sker kalibreringen framför allt via kontroll av godsflöden vid kustområdena och justeringar av dessa med hjälp av kostnadsfaktorer.

⁵ Röjandekontroll beskrivs lättillgängligt i Trafikanalys (2017) och mer utförligt i Statistiska Centralbyrån (2015).

⁶ Liedtke, Gernot (2018)

⁷ Se exempelvis Trafikverket (2017b)

⁸ Trafikanalysforum (2018)

I den svenska persontransportmodellen Sampers sker kalibrering huvudsakligen genom rimlighetsbedömning av resultaten, exempelvis genom att undersöka om storleken på resandet är rimligt. Kalibrering sker ofta genom så kallad skaftning, vilket innebär att man för in och/eller justerar kopplingslänkar mellan modellens zoner och hållplatser/stationer.

Ju mer detaljerade data man har tillgång till, desto bättre kan man kalibrera modellen. Här kan typiskt sekretess sätta olika begränsningar.

2.8 Validering

Validering innebär att man kontrollerar att den skattade och kalibrerade modellen ger resultat som överensstämmer med observerade, faktiska fördelningar av resor och transporter uppdelade på transportmedel, transportväg, etc.

Validering av persontransportmodeller kan göras mot exempelvis trafikräkningar, biljettdata och resvaneundersökningar. För godstransportmodeller används ofta uppgifter om totala godstransportflöden fördelat på transportslag.

Trafikanalysforum beskriver sambandet mellan kalibrering och validering:

Till skillnad från kalibreringen där syftet är att kontrollera att modellens beteende och logik är korrekta, går valideringen ut på att se till att utdata från modellen överensstämmer med observerade värden. Valideringen görs på en kalibrerad modell och därför kan justering av vissa parametrar innebära att kalibrerings- och valideringsprocessen måste göras om. Ibland sker detta iterativt tills det att valideringskriterierna har uppnåtts.⁹

Olika sekretessaspekter kommer in även i valideringen, bland annat genom att det kan vara svårt att få tillgång till detaljerade data över transporter och resor i enskilda relationer, länkar och noder.

Det kan vara svårt att beskriva hur bra resultaten överensstämmer med verkligheten för att på så sätt skapa förtroende för modellerna. Genom att exempelvis ange att modellen har "90 %-ig överensstämmelse med observerade flöden för valideringsåret 2017" avslöjar man kanske indirekt uppgifter som möjligen kan vara känsliga.

2.9 Resultatuttag

Resultatuttag från transportmodeller kan göras på olika detaljeringsnivå beroende på modell. Det finns såväl övergripande, nationella modeller som regionala modeller.

Resultaten kan vara problematiska ur sekretessynpunkt om modellen är "för bra". För godstransporter på järnväg är exempelvis antalet kunder mycket få och det är inte särskilt svårt att identifiera enskilda aktörer om man bryter ner transporterna på länkar i systemet och varugrupp.

⁹ Trafikanalysforum (2018)

3 Beskrivning: Norge

3.1 Databehov

För modellering av godstransporter i Norge används ett nationellt modellsystem som består av varuflödesmatriser, en jämviktsmodell, kostnadsfunktioner, en nätverksmodell och en logistikmodul som beräknar frekvens och transportlösning.¹⁰

De viktigaste delkomponenterna i godsmodellen är:

1. Varuflödesmatriser för 39 varugrupper beskriver årliga varuflöden mellan norska kommuner samt mellan norska kommuner och utlandet.
2. Information om antal företag per geografisk zon som är leverantör eller mottagare av de olika varutyperna i matriserna. Inom Norge finns ungefär 90 geografiska zoner. Utanför Norge är huvudregeln en zon per land, med en finare indelning för de viktigaste länderna.
3. Nätverk som beskriver infrastrukturen och transportsystemet inklusive terminaler och omlastningspunkter.
4. Kostnadsfunktionerna beskriver kostnaderna för olika transportmedel kopplat till sträcka (exempelvis bränsleåtgång) och tid (bland annat personal), men också andra moment i transportkedjan som exempelvis omlastning. Kostnadsfunktionerna byggs upp "bottom-up" genom att arbetsmoment i transportkedjan definieras utifrån behov av arbetskraft, maskiner och andra resurser. Dessa prissätts därefter med hjälp av exempelvis löne- och prisstatistik. Med andra ord måste man inte samla in specifika kostnadsuppgifter från transportörer för att använda i transportmodeller.¹¹
5. Optimeringsrutiner för att välja storlek, sändningsfrekvens och transportkedja.¹²

För persontransporter använder man en nationell modell för långväga resor som kombineras med fem regionala transportmodeller samt specifika modeller för Osloregionen och Buskerudsregionen.¹³ Modellerna bygger på indata i form av exempelvis befolkningsmängd i dag och enligt prognoser, arbetsställen, infrastruktur och transportsystem samt kostnadsbild för exempelvis drivmedel och kollektivtrafiktaxor. Resultaten beskrivs i form av antal resor i olika relationer fördelat per transportslag och ärendetyp.

¹⁰ TØI (2017), sid 8

¹¹ Däremot samlar Statistiska Sentralbyrån in uppgifter om prisutvecklingen inom godsransporter för att kunna beräkna bland annat olika inflationsmått.

<https://www.ssb.no/innrapportering/naeringsliv#T> Uttag 180131

¹² TØI (2017), sid 8f

¹³ Hovi, Inger Beate (2018)

3.2 Datakällor och insamling

Huvudprincipen för insamling av data i Norge är att det görs av Statistiska Sentralbyrå (SSB). SSB har ett starkt mandat att begära in uppgifter med hjälp av statistikkloven.¹⁴ Samtidigt skyddas personuppgifter av sekretess på liknande sätt som i bland annat Sverige.¹⁵ Det är myndigheten Datatilsynet som säkerställer sekretessen genom personopplysningsloven.

En central datakälla för godstransporter är varuflödesundersökningen, som syftar till att få bättre kunskap om de viktigaste varuflödena i Norge.¹⁶ Undersökningen genomfördes senast år 2014 genom att ett urval om cirka 4 200 industriföretag och grossister fick rapportera varusändningar till norska kunder i Norge. SSB samlar dessutom in uppgifter om alla varutransporter som görs av de största transportföretagen i Norge, de så kallade "samlasterne". Till dessa räknas Bring, Schenker, DHL med flera. Syftet är att reducera rapporteringsbördan för industrier och grossister.

Verksamheterna skall ange bland annat följande uppgifter:

- Leveransadress med postnummer
- varumängd i kg eller m³,
- varuvärde,
- om transportkostnad ingår i varans pris eller inte,
- typ av gods,
- transportmedel samt
- om man upplever några flaskhalsar i transportsystemet.

Samtliga rapporter omfattas av tystnadsplikt (taushetsplikt) enligt statistikkloven § 2-4.

Den senaste varuflödesundersökningen bedöms dock inte ha gett tillfredsställande kvalitet i svaren.¹⁷ Bland annat därför arbetar man också med andra datakällor och metoder. Alla transporter inom jordbruk och fiske estimeras baserat på primärnäringstatistik. Det innebär att man fördelar transporterna från produktionsplatsen till närmaste kundföretag, oavsett om det är den faktiska kunden. För skogstransporter finns ett företag som heter Skog-data AS, där alla transporter sammanställs.

Varutransporter till och från utlandet samlas in från tulldeklarationer som lagras av Toll og afgiftsdirektoratet. Transporterna kan fördelas geografiskt utifrån uppgifter om produktionsplats och leveransadress.

¹⁴ Finansdepartementet (1989/2016). Lov om offisiell statistikk og Statistisk Sentralbyrå (statistikkloven), se <https://lovdata.no/dokument/NL/lov/1989-06-16-54>

¹⁵ SSB. Personvern og datasikkerhet <http://www.ssb.no/omssb/personvern/personvern-og-datasikkerhet> Uttag 180131

¹⁶ Dette stykke bygger på SSB. Varestrømsundersøkelsen. <http://www.ssb.no/innrapportering/naeringsliv/vsu2014> Uttag 180131

¹⁷ Hovi, Inger Beate (2018)

Validering och kalibrering av godstransportmodellen sker i hög grad mot den årliga statistiken över "transportytelser" (transporttjänster) i Norge.¹⁸ Uppgifterna för trafikarbetet på järnväg kommer från följande källor:¹⁹

- NSB persontrafikk
- Andra banor (Flytoget mellan Asker och Oslo Gardermoen Airport, Flåmsbanan, Bratsbergsbanen, Ofotbanen och Gjøvikbanen).
- Enskilda bolag
- NSB, SJ och Jernbaneverket (utlandsresande, sammanställdes dock enbart fram till år 2012)
- CargoNet (tidigare NSB gods) lämnar uppgifter till Bane Nor, som offentliggör dem.
- Övriga transportörer av gods på järnväg lämnar uppgifter till SSB.

Jernbanedirektoratet, tidigare en del av Jernbaneverket, publicerar trafikstatistik över den norska järnvägen.²⁰ Jernbanedirektoratet gör en totalundersökning bland operatörer.²¹ Antal person- och godståg per dygn publiceras dels för nationen som helhet, dels för olika sträckor. Siffrorna publiceras som intervall om cirka 20 tåg per dygn med lägsta nivå 1-50 tåg per dygn. Vidare publiceras antal tåg per dygn på olika banor i och mellan storstadsområdena. Dessutom redovisas antal passagerare per år per tågbolag, vilket blir tämligen avslöjande för exempelvis Flytoget AS, men däremot inte för NSB AS.

Gränsöverskridande trafik redovisas i tusentals resor per år för de olika operatörerna SJ AB och NSB AS. För godstransporter redovisas godsmängd i ton fördelat på CargoNet AS och "andra" operatörer. Även för godstransporter redovisas gränsöverskridande trafik per operatör. Vidare redovisas punktligheten per godstågsoperatör.

Statistiska Sentralbyån använder data från Jernbanedirektoratet och publicerar denna på mer detaljerad nivå. Exempelvis går det att söka godstransporter i ton efter region för lastning respektive region för lossning och på samma sätt interregionala resor mellan regioner. Norge delas in i sju regioner. Vidare redovisas godstransportarbetet och godsmängden utifrån transportenhet och om transporten är nationell, internationell, m m. Statistiken finns också tillgänglig på aggregerad nivå för 20 varugrupper. För personresande redovisas bland annat antal passagerare per bana, vilket är en hög detaljeringsgrad.²² Exempel på resandetal som redovisas gäller Flytoget, InterCity - Oslo-Halden, InterCity - Oslo-Lillehammer, InterCity - Oslo-Skien, Sørlandsbanen, Bergensbanen, med flera.

¹⁸ TØI (2016)

¹⁹ TØI (2016), sid 18, 22 och 24f

²⁰ Jernbanedirektoratet (2017).

²¹ SSB. Jernbanetransport. Om statistikken. <https://www.ssb.no/transport-og-reiseliv/statistikker/jernbane> Uttag 180131

²² SSB. Jernbanetransport. <https://www.ssb.no/statbank/list/jernbane?rxid=857a7e74-7670-4fa8-8c04-0b21f445c006> Uttag 180131

För att kalibrera och validera persontransportmodellerna använder man också nationella resvaneundersökningar. Samferdseldepartementet anger att den nationella resvaneundersökningen är avgörande för transportmodellerna.²³

Man kan också validera eller kalibrera godstransportmodellerna mot omlastningspunkter som hamnar. När det gäller privata respektive små hamnar kan det uppstå problem kopplat till sekretessen, eftersom flödena eventuellt kan härledas till enskilda företag.

3.3 Utmaningar och hantering av sekretess

Det framhålls som en generell utmaning med tanke på sekretessen att arbeta med allt för detaljerade flöden. När 39 varugrupper skall fördelas per kommun respektive zon i större städer samtidigt som det ofta kan vara få tillverkare av en vara så är det svårt att undvika att identifiera enskilda företag.²⁴ Det är inte endast av sekretessskäl som detaljerade data inte kan lämnas ut, utan också för att det finns en osäkerhet i statistiken om den bygger på ett urval av den berörda populationen. Genom att redovisa data på en mer aggregerad nivå reduceras den osäkerheten.

Det finns en upplevd efterfrågan på mer och mer detaljerade data från beslutsfattarnas/myndigheternas sida.²⁵

Generellt verkar det finnas en hög acceptans bland företag att man måste lämna ifrån sig potentiellt känsliga uppgifter, man förstår nyttan med statistiken. Företagen är mer oroliga för merarbete vid insamling/sammanställning än att uppgifterna skall undslippa sekretessen.²⁶

SSB skriver att man vid enstaka tillfällen lämnar individdata för forskning och offentlig planering, men data är då i merparten av fallen *avidentifierade* eller *anonymiserade* för att inte kunna härledas till enskilda personer. Undantagsvis kan det vara aktuellt att förmedla data till forskning som kan härledas till individer, men då skall forskaren och SSB få detta beviljat av Datatilsynet.²⁷

En intressant aspekt är att SSB publicerar mer detaljerade transportdata än Jernbanedirektoratet (som samlar in data) och samtidigt motiverar avsteget från sekretessen med "avtale med operatörerna." SSB skriver om sekretessen (konfidentialitet):

"Statistikklovens hovedregel er at opplysninger ikke skal offentliggjøres slik at de kan føres tilbake til oppgavegiver eller annen identifiserbar enkeltperson. Etter avtale med operatørene, er det i denne statistikken gjort unntak fra denne

²³ Samferdseldepartementet (2016). Prop. 1 S (2016 – 2017) Proposisjon til Stortinget (forslag til stortingsvedtak), for budsjettåret 2017. sid 33

²⁴ Hovi, Inger Beate (2018)

²⁵ Hovi, Inger Beate (2018)

²⁶ Hovi, Inger Beate (2018)

²⁷ SSB. Personvern og datasikkerhet <http://www.ssb.no/omssb/personvern/personvern-og-datasikkerhet> Uttaga 180131

regelen. Etter avtale med operatørene, er det i denne statistikken gjort unntak fra denne regelen.”²⁸

Avtalet är enligt uppgift inte skriftligt, men det skall ha etablerats enighet om att SSB kan publicera uppgifterna.²⁹ "Avtalet" implementeras sannolikt genom att det är angivet i inledningen av undersökningsblanketten att data kommer att publiceras.³⁰

Det är möjligt att ansöka hos SSB om att använda data för olika ändamål. Företag kan också få tillgång till data efter ansökan, men inte på samma detaljnivå som SSB och forskningsinstitutioner. I Norge skapas tillgängligheten till grunddata i flera steg³¹:

1. Den organisation som skall använda data skall vara godkänd som forskningsinstitut.
2. Avtalet om utlån av data definieras för den aktuella studien och signeras av behörig företrädare för verksamheten. Avtalet anger bland annat restriktioner för användning och publicering.³²
3. Personal i verksamheten som skall utföra arbetet signerar avtal om tystnadsplikt.³³
4. När arbetet är avslutat skall utlånade data förstöras, vilket skall bekräftas av behörig företrädare för verksamheten.³⁴

Om man skall arbeta tillsammans med en annan organisation så skriver man det i ansökan och därmed kan man lämna över data mellan sig. Verksamheter eller personer som inte följer reglerna kan nekas tillgång till data framöver.

Om man är användare av godstransportmodellen och gör resultatuttag eller andra uttag ur den så har man hittills bara använt sig av muntliga restriktioner som går ut på att säkerställa att data inte får publiceras så att enskild verksamhet kan identifieras.

Företrädare för Transportøkonomisk institutt menar att systemet fungerar väl. Det finns ett bra samarbete mellan SSB och TØI med stor grad av ömsesidig tillit gällande dataanvändning. SSB försöker så långt det är möjligt att tillmötesgå TØI:s behov av data. TØI å sin sida vet vilka behov uppdragsgivarna har och kommunicerar det till SSB.

I en värdering av det norska systemet för godstransportmodeller tas inte sekretessfrågor upp som något problem, däremot behandlas att en allt för hög detaljeringsgrad är förknippat med statistiska osäkerheter om grunden är en urvalsundersökning.³⁵ I en

²⁸ SSB (2018). Jernbanetransport. Om statistikken. <https://www.ssb.no/transport-og-reiseliv/statistikker/jernbane> Uttag 180123

²⁹ Grimstad, Kristian (2018)

³⁰ Hovi, Inger Beate (2018)

³¹ Hovi, Inger Beate (2018). TØI har efterfrågat och fått ut statistik från Sverige och har då gått igenom en liknande procedur.

³² SSB: "Avtale om utlån av anonyme statistikkopplýsingar til forskning mellom Statistisk Sentralbyrå (SSB) og [organisationens namn]"

³³ SSB: "Taushedserklæring"

³⁴ SSB: "Sletning av opplysninger"

³⁵ TØI (2017)

rapport för det norska samferdseldepartementet om kvaliteten på data för persontransporter tas sekretessaspekter upp, men framställs inte som något direkt problem.³⁶

3.4 Sammanfattning

Norge arbetar i hög grad med insamling av data med stöd i uppgiftsskyldighet. Samtidigt garanteras sekretessen i lagstiftning som i hög grad påminner om bland annat den svenska. Metoder som används för att säkerställa att enskilda personer inte kan identifieras är bland annat anonymisering och aidentifiering men också vanliga statistiska metoder för röjandekontroll. En intressant aspekt är att data som Jernbanedirektoratet samlar in från operatörerna och redovisar på en aggregerad nivå publiceras med högre detaljeringsgrad av SSB med hänvisning till "avtal med operatörerna". Detta avtal verkar inte finnas i skriftlig form, men kan ha anammats som praxis av branschen. Tillgång till grunddata ges framför allt för ändamål som kan definieras som forskning. Detta verkar ge en tillräckligt god tillgänglighet för att i synnerhet forskare skall kunna driva ett utvecklingsarbete av transportmodeller.

4 Beskrivning: Nederländerna

4.1 Databehov

I Nederländerna används en strategisk, nationell modell för godstransporter som heter BASGOED.³⁷ Modellens huvudsakliga syfte är att tillgodose behov av underlag för politiska beslut. En effekt av detta är att BASGOED nyttjar data som med stor sannolikhet bedöms finnas tillgängliga i framtiden. Modellen kan bland annat presentera transportflöden i ton per varutyp utifrån var flödena genereras respektive tas emot. Dessutom fördelar modellen flödena på de olika transportsätten. Varutyperna delas in i tio varugrupper och den geografiska dimensionen består av 40 inhemska zoner och 29 utländska zoner.³⁸

För personresor finns en nationell transportmodell (LMS) och dessutom fyra regionala modeller samt en särskild modell för Amsterdam (VMA). LMS ägs av Rijkswaterstaat, vilket är den ansvariga myndigheten för infrastruktur och vattenfrågor. Myndigheten sorterar under departementet för infrastruktur och miljö.³⁹ Tillämpningarna görs av ett begränsat antal konsulter.⁴⁰

4.2 Datakällor och insamling

Data för de nationella modellerna samlas in av Statistics Netherlands för både person- och godstransporter samt av Dutch Railways (NS) för persontransporter.

³⁶ TØI (2014)

³⁷ Detta avsnitt bygger i hög grad på deJong, Gerard (2018)

³⁸ deJong et al (2011)

³⁹ Dutch Ministry of Infrastructure and the Environment

⁴⁰ Van der Waard (

Kalibreringen av den aktuella versionen av BASGOED har gjorts utifrån den holländska godstransportdatabasen ("Basisbestand Goederenvervoer"), vilken skapades 2004 av Statistics Netherlands. Denna data finns tillgänglig på mikronivå, men används alltså endast på en aggregerad nivå för att beskriva transporter mellan zoner som är relativt stora. Den primära anledningen till detta har dock inte varit sekretessaspekter, utan en strävan efter att hålla modellen enkel, transparent och enkel att uppdatera.

Persontransportmodellerna estimeras baserat på mikrodata från flera år av nationella resdagsboksundersökningar som samlas in av Statistics Netherlands. Dessutom samlar Dutch Railways in data från tågpassagerare (KLIMAAT data). Vidare används vägtrafikdata (trafikeräkningar), bland annat för att estimeras och kalibreras en grundmatris för vägtrafiken som sedan kan räknas upp till det aktuella prognosåret. För tågtrafik skapas en motsvarande OD-basmatris av tågoperatörernas egna undersökningar. Detta arbete görs av tågoperatörerna och baseras på en frivillig överenskommelse. I gengäld blir operatörerna delägare av den nationella persontransportmodellen.

För validering av godstransporter används olika elasticiteter samt tidsvärdering från (forsknings-)litteraturen. För persontransporter används även basmatriserna för väg- och järnvägstrafik samt reslängdsfördelningar från olika källor.

4.3 Utmaningar och hantering av sekretess

Mikrodata från Statistics Netherlands måste hållas konfidentiella och kan bara användas för ett specifikt projekt. För att möta detta krav lagrar exempelvis det privatägda forskningsinstitutet Significance data på en intern server med begränsad tillgänglighet.

För att kunna använda mikrodata från resdagsboksundersökningen ansöker den aktuella organisationen om tillstånd från Rijkswaterstaat. Departementet är samtidigt beställare av modellutvecklingsuppdraget. Både departementet och Statistics Netherlands (som alltså samlar in och lagrar uppgifterna) ställer kravet på tillstånd från Rijkswaterstaat.

För att kombinera olika typer av mikrodata anger Significance att det finns mer stringenta krav. Exempelvis ville Significance kombinera data från resdagboksundersökningen med inkomstuppgifter som Statistics Netherlands samlar in från deklARATIONER. Eftersom skyddet av inkomstuppgifter är starkt kunde data inte föras över till Significance utan arbetet var tvunget att utföras på Statistics Netherlands server. Servern har begränsad kapacitet och det är svårt att få tillgång till den, beroende på att efterfrågan är stor från många användare (inte enbart inom transportområdet). Significance bedömde att tidsåtgången skulle bli allt för kostsam i termer av tid (en estimering bedömdes ta flera veckor) och avstod från arbetet.

4.4 Sammanfattning

Nederländerna kombinerar datainsamling av Statistics Netherlands och Dutch Railways med frivilliga överenskommelser med andra operatörer för att skapa exempelvis OD-matriser på detaljerad nivå. Det är särskilt intressant att man på så sätt får tillgång till detaljerade data för kalibrering av persontransportmodellen. I gengäld får alltså

operatörerna tillgång till den nationella persontransportmodellen, vilket bidrar till den transparens som utgör en viktig grund för väl fungerande marknader.

Intressant är också att man avsiktligt väljer att bygga BASGOEDS-modellen på en aggregerad nivå, inte primärt av sekretesskäl men för att få en enklare hantering och större transparens. Det skall samtidigt betonas att den geografiska tätheten utgör en viktig faktor här. När Nederländerna delas in i 40 zoner är varje zon i genomsnitt drygt 1 000 km². När Norge delas in i 90 zoner är varje zon drygt fyra gånger större. Skillnaden förstärks genom att den ekonomiska aktiviteten i varje nederländsk zon är större. Sammantaget blir zonindelningen i Nederländerna mer detaljerad ur ett geografiskt perspektiv, samtidigt som möjligheten att urskilja enskilda företag blir mindre. Detta illustrerar hur de givna premisserna för en nation kan påverka frågan om sekretess kopplat till transportmodeller.

Statistics Netherlands samlar in data med stöd i lagen och skyddar uppgifterna på olika sätt, exempelvis genom statistisk röjandekontroll.⁴¹ Organisationer som vill använda mikrodata måste ansöka om detta. Det har inte framgått i detta arbetet exakt hur de ansökningarna värderas. På det hela taget verkar dock möjligheten att få tillgång till data vara relativt god, vilket grundas på att den intervjuade organisationen är ett privat, kommersiellt forskningsinstitut. Däremot kan det valda sättet att säkerställa sekretessen rent praktiskt, genom att data endast finns tillgängligt på Statistics Netherlands hemsida, orsaka problem i form av ineffektiva arbetsprocesser.

5 Beskrivning: Danmark

5.1 Databehov, datakällor och insamling

Den viktigaste transportmodellen i Danmark⁴² är Landstrafikmodellen⁴³, som täcker både person- och godstransporter. Dessutom finns ett antal lokala och regionala modeller, vilka företrädesvis hanterar persontransporter. Det är Danmarks Tekniska Universitet (DTU) som har utvecklat Landstrafikmodellen tillsammans med flera stora konsultbolag.⁴⁴

DTU anger att bland annat följande data samlas in:

- Trafiktal som bygger på trafikräkningar från vägmyndigheter (primärt Vejdirektoratet), kollektivtrafikoperatörer, hamnar och flygplatser
- Detaljerad befolkningsstatistik från Danmarks statistik
- Resvanedata från den Danska Nationella Resvaneundersökningen och andra källor

⁴¹ Statistics Netherlands (2012)

⁴² Detta avsnitt bygger i hög grad på Naftchi, Yalda (2018)

⁴³ Landstrafikmodellen beskrivs av DTU på hemsidan www.landstrafikmodellen.dk

⁴⁴ Se exempelvis www.landtrafikmodellen.dk men också Trafikanalys (2011).

För kalibrering används primärt trafikräkningar och för validering jämför man modellresultat med annan relevant statistik, gör back-casting beräkningar för tidigare år samt jämför med andra modeller.

5.2 Utmaningar och hantering av sekretess

Många datauppgifter omfattas av någon form av sekretesskrav av ett eller flera av följande skäl:

- Persondata
- Företagskänslig information
- Intellektuellt kapital eller immateriella rättigheter.

Den sista kategorin är intressant eftersom den inte har lyfts fram i något av de andra studerade länderna. Det som avses är den kunskap som utvecklas i modellarbetet och som kan skyddas med immaterialrättslagstiftning, exempelvis algoritmer.

Landtrafikmodellen har utvecklats av DTU och flera konsultföretag tillsammans.

Företrädare från DTU menar att kraven på sekretess kan variera mellan företagen som berörs av undersökningen:

“Can be anything from ‘please publish for me’ to ‘please read, then eat the paper immediately’.”⁴⁵

DTU anger att man säkrar sekretessen på ett antal principiella och praktiska sätt:

- Data lagras så att tillgången skapas på en ”need-to-know”-basis.
- Data processas anonymiserat.
- Källdata används enbart för att förbereda den data eller de samband som ingår i modellen.

Över lag har data kunnat användas så som man har önskat i arbetet med Landtrafikmodellen, dock med vissa undantag. Dessa har framför allt gällt godstransportdata från företag. DTU anger att samtliga metoder för att få fram data från berörda parter som Sweco angett som exempel i frågeformuläret (se bilaga 1) har använts i arbetet:

- Lagkrav
- Krav i samband med upphandling av trafik
- Frivilliga överenskommelser
- Klassificering av data
- Nära dialog med branschorganisationer

⁴⁵ Naftchi, Yalda (2018)

När det gäller lagkrav så styrs en del av rapporteringen av Förordning (nr 1458) om rapportering av data för analys och planering etc. på järnvägsområdet.⁴⁶ Det framgår av förordningen att data undantas från gängse offentlighetsprincip avseende officiella dokument och rapporteringen verkar ske på verksamhetsnivå snarare än exempelvis fördelat per sträcka.

För passagerartrafik på järnvägen regleras tillgången till data genom statens upphandlingar med operatörer som DSB, Arriva med flera. I trafikavtalen slås det fast att operatören skall göra data tillgängliga för Trafikstyrelsen och Transport-, Byggnings- och Boligministeriet. Detta innebär att DTU och även konsultbolag som arbetar med Landtrafikmodellen eller i andra uppdrag för departementet i realiteten har fri tillgång till data, dock får man inte vidarebefordra uppgifterna.⁴⁷

För godstransporter på järnväg sker en inrapportering av data till Danmarks statistik.⁴⁸ DTU kan använda dessa data, men i övrigt måste det ske en förhandling mellan DTU och godsoperatören, det vill säga en frivillig överenskommelse där de exakta förutsättningarna definieras från fall till fall.

På infrastruktursidan anger DTU att man ofta möter en "betydande slutenhet" hos BaneDanmark när det gäller att få tillgång till data, men att man typiskt löser detta genom att upprätta sekretessavtal (eg. non-disclosure-agreements).⁴⁹

5.3 Sammanfattning

I Danmark omfattas många data som behövs för transportmodellerna av olika restriktioner kopplade till sekretess. Det framstår dock som att i alla fall DTU lyckas hantera dessa genom en mångfald av lösningar, från lagstiftning till frivilliga överenskommelser i det specifika fallet. Behovet av sekretess verkar därför inte påverka den upplevda användbarheten av Landtrafikmodellen, i alla fall inte enligt DTU. I en genomgång av publikationer från de senaste sju åren på Transport-, Byggnings- och Boligministeriets hemsida verkar inte frågan om sekretess behandlas explicit.⁵⁰ Det tas inte heller upp som någon direkt begränsning i dokumentationen till Landtrafikmodellen.⁵¹

DTU lyfter en intressant aspekt av sekretessen, det som gäller frågor om immaterialrätt. Detta är en fråga som möjligen har varit extra relevant i Danmark, där utvecklingen av den dominerande transportmodellen har gjorts i samarbete mellan Danmarks Tekniske Universitet och flera konsultbolag. Det är dock inte någon unik situation för Danmark och

⁴⁶ Trafik- og Byggestyrelsen (2016)

⁴⁷ Naftchi, Yalda (2018)

⁴⁸ Danmarks statistik (2017)

⁴⁹ På danska: "betydende lukkedhed". DTU menar att detta är "direkt komiskt", eftersom samma information omfattas av offentlighetsprincipen. Sweco har dock inte kunnat värdera om detta är en korrekt tolkning av den danska offentlighetsprincipen för de aktuella data.

⁵⁰ Transport-, Byggnings- og Boligministeriet (2018)

⁵¹ Dokumentationen finns tillgänglig på www.landtrafikmodellen.dk

sannolikt har den frågan reglerats på sedvanligt vis som i gemensamma utvecklingsprojekt generellt.

6 Beskrivning: Tyskland

6.1 Databehov, datakällor och insamling

Tyskland skiljer sig från övriga studerade länder genom att staten har tagit ett relativt litet ansvar för transportmodellernas tillkomst och utveckling.⁵² Därför har utvecklingsarbetet i hög grad gjorts av konsultföretag, delvis baserat på kommersiella uppdrag av transportdepartementet samt aktörer som Deutsche Bahn. Detta har resulterat i bristande transparens och dålig tillgänglighet till dokumentation kring tyska transportmodeller.⁵³

När det gäller godstransporter så finns det ett system av modeller som har utvecklats av företaget TTS TRIMODE Transport Solutions GmbH, som har bildats från bland annat företaget BVU Wirtschaft.⁵⁴ Ett exempel på modell är KaloDrom, vars tågmodul optimerar transporten för både gods och personer baserat på följande kriterier:

- Avstånd
- Tid
- Begränsningar i järnvägsnätet i form av lastprofiler, lutningar, viktbegränsningar och omkörningsmöjligheter
- Kostnader för olika tågtyper

Däremot hanteras inte frekvens och sändningsstorlek eller lagerkostnader för gods. Godsmodulen är beroende av uppdrag och data från Deutsche Bahn för sin utveckling. Med hjälp av såväl revealed preference- och stated preference-data skattar man modellen, i synnerhet transportvalen. De faktiska fallstudierna i form av konkreta transportupplägg uppgår till cirka 1 000.⁵⁵

För kalibrering av godstransportmodellerna används bland annat offentliga data över godstransporter på väg som rapporteras till Eurostat. Denna bygger på en urvalsundersökning med ungefär 1,5 miljoner transporter fördelade på olika varutyper. Data redovisas på den geografiska nivå som kallas LAU 2 och som motsvarar kommuner. Det går också att utläsa information om exempelvis typ av fordon samt om terminaler eller hamnar ingår i transportkedjan. Dessa data redovisas enbart på en aggregerad nivå av dels sekretesskäl, dels kvalitetsskäl. I en urvalsundersökning kan allt för detaljerade data ge felaktiga resultat om antalet observationer är för få.

⁵² Detta avsnitt bygger till stor del på Liedtke, Gernot (2018).

⁵³ Liedtke, Gernot (2018) men också Trafikanalys (2011).

⁵⁴ Intressant nog är namnet på den kommande europeiska transportmodellen för både gods och passagerare också TRIMODE. Uppdraget att utveckla modellen tillföll dock inte TTS TRIMODE Transport Solutions GmbH, utan PTV Group, som står bakom Visum-modellen.

⁵⁵ Liedtke, Gernot (2018)

Det finns liknande matriser som för vägtransporter (gods) tillgängliga även för järnvägstransporter, såväl för passagerare som för gods. Järnvägsoperatörerna är ålagda i lag att redovisa dessa, enligt uppgift som en del av Tysklands redovisning till EU.⁵⁶ Det skall vara möjligt att ta del av matriserna från ansvarigt departement. Det är oklart vilken upplösning som kan erhållas samt om det är förknippat med någon kostnad.⁵⁷

I Tyskland genomförs ingen varuflödesundersökning.

För persontransportmodeller är det ledande företaget Intraplan Consult GmbH.⁵⁸ Företaget har sammanställt data över personresor på uppdrag av tyska regeringen sedan 1985, bland annat i form av OD-matriser för olika transportslag.⁵⁹ Man har också gjort nationella resandeprognoser på uppdrag åt regeringen och har samarbetat med nämnda företaget BVU Wirtschaft, som då svarade för godstrafiken. Intraplan anger att man vid nulägesbeskrivningar av trafikflöden försöker kombinera grunddata som biljettförsäljning från Deutsche Bahn eller pendlingsstatistik med modellberäkningar. Intraplan beskriver sin kompetens att göra korrekta prognoser med hänvisning till bland annat externa, offentliga granskningar i samband med planering av olika infrastrukturprojekt.

Motsvarigheten till Statistiska Centralbyrån heter Destatis (Statistisches Bundesamt). De samlar in och förvaltar en del data. DLR⁶⁰ hanterar en del uppgifter och likaså olika konsultbolag.

6.2 Utmaningar och hantering av sekretess

För närvarande arbetar tyska transportdepartementet för att göra det möjligt för forskare att få tillgång till grunddata. Eftersom data i vissa fall lagras av konsultföretag måste departementet ha möjlighet att styra företagen att tillhandahålla data. En tydlig risk är att konsultföretagen får svaga incitament att tillmötesgå sådana önskingar på ett effektivt sätt.

6.3 Sammanfattning

Arbetet med transportmodeller i Tyskland försvåras av att staten har valt att ta ett begränsat ansvar för utvecklingen av modellerna. Det gör det också svårt att få en tydlig bild av sekretessaspekter. Data finns hos olika aktörer och kan skyddas såväl av offentlig sekretess och affärssekretess. Det mest konkreta som kommit fram i detta arbete är den lagstadgade skyldigheten för järnvägsoperatörer att presentera motsvarande OD-matriser för såväl passagerare som gods, vilket tycks utgöra del av redovisningen till EU. Det är oklart vilken detaljeringsgrad som kan erhållas vid uttag från matriserna.

Det är också notabelt att det pågår ett departementsarbete för att göra det möjligt för forskare att få tillgång till mikrodata.

⁵⁶ Liedtke, Gernot (2018)

⁵⁷ Liedtke, Gernot (2018)

⁵⁸ Intraplan: www.intraplan.de

⁵⁹ Intraplan: <http://www.intraplan.de/?p=77&l=1> , uttag 180212

⁶⁰ DLR är alltså det statliga, federala forskningscentrat för rymdfart.

7 Beskrivning: Sverige

Detta avsnitt om Sverige syftar inte att ge en detaljerad beskrivning av svenska modeller, databehov och källor, då detta behandlas i andra delar av myndighetens uppdrag till regeringen, inte minst i Trafikanalys delrapportering samt i underlagsrapporter från Trafikverket.⁶¹ I arbetet med denna rapport har dock sekretessaspekter kommenterats av Väg- och Transportforskningsinstitutet (VTI).⁶² Tillsammans med nämnda underlagsmaterial blir det därmed möjligt att skapa en bild av hur olika aktörer (forskare, myndigheter och konsulter) ser på sekretessfrågorna, vilket i sig är intressant.

7.1 Databehov, datakällor och insamling

Väg- och Transportforskningsinstitutet använder godstransportmodellen Samgods för forsknings, utvecklings- och utredningsprojekt. När det gäller persontransportmodellen Sampers har VTI ett uppdrag från Trafikverket att utveckla nästa generation av modellen, men tillämpar i dagsläget inte modellen i andra projekt.

VTI samlar vanligtvis inte in data, utan använder data från andra källor som Trafikverket, Trafikanalys och Statistiska Centralbyrån.

Efterfrågematriserna i godstransportmodellerna skattas m h a data från varuflödesundersökningar, handelsstatistik och, åtminstone delvis, sekretessbelagda uppgifter om produktion och konsumtion av varor i Nationalräkenskaperna och associerade data (SCB). Mikrodata från varuflödesundersökningarna förefaller kräva att det görs på universitet, men det bestäms eller blir tydligare när undersökningen från år 2016 ska användas.⁶³

VTI har genomfört kalibrering av Samgods-modellen och därvid använt ett antal datauppgifter gällande järnvägstransporter, se tabell 2.

Tabell 2. Data som VTI använder för kalibrering av Samgods. Källa: VTI

| Data | Datakälla |
|--|-------------------------------------|
| Totala tonkm för fraktat gods inom Sverige på väg, järnväg respektive sjöfart. | Officiell statistik. |
| Antal ton omlastat gods per kustområde i Sverige. | Trafikanalys undersökning Sjötrafik |
| Antal tonkm på jvg inom Sverige per STAN-varugrupp. | Trafikverket |
| Antal kTon som transporteras på järnväg över Öresundsbron per år | Trafikverket |
| Fördelningen (%) av tonkm på järnväg (inom Sverige) över de tre tågtyperna: Kombi, System och Vagnslast. | Trafikverket |

⁶¹ Trafikanalys (2017), Trafikverket (2017a) och Trafikverket (2017b)

⁶² Uppgifterna om och från VTI bygger på Vierth, Inge (2018)

⁶³ Källa: Sweco

Fokus i kalibreringen ligger på inrikes transportkm per transportslag, med fokus på järnväg där även fördelningen på 12 STAN-varugrupper hanteras. Vidare kalibreras mot flöden genom hamnar i 14 kustområden längs svenska kusten, även detta med styrning mot fördelningen på 12 STAN-varugrupper hanteras. En VTI-rapport finns som beskriver hur hamnflödena ser ut i verkligheten baserat på konfidentiell statistik från Trafikanalys (endast delar av rapporten med offentliga data publiceras).

För validering av Samgods används i ett särskilt projekt "KVAL"⁶⁴ sekretesskyddade mikrodata från Trafikanalys för valideringsändamål, detta gäller dock ej järnvägstrafik.

Vid validering av Sampers anger VTI att Trafikverket står för kalibrering och validering mot passagerarflöden och restider för olika färdmedel innan persontransportmodellen blir officiell. Trafikverket anger att följande data skulle vara till "stor nytta" för att bättre kunna utveckla, validera och kalibrera prognosmodeller för persontrafik⁶⁵:

- Statistik över resandet mellan olika start-/målpunkter
- Resandestatistik över antal påstigande/avstigande per station
- Medelkostnad som betalats för att resa mellan olika start-/målpunkter uppdelat på olika ärenden (arbete/privat/tjänste för långväga resor; tjänste/arbete/övrigt för regionala resor)
- Trafikeringskostnad för olika tågtyper

7.2 Utmaningar och hantering av sekretess

I hanteringen av uppgifter, inklusive sekretess och känsliga persondata, styrs VTI framför allt av tryckfrihetsförordningen (1949:105), yttrandefrihetsgrundlag (1991:1469) samt offentlighets- och sekretesslagen (2009:400) och personuppgiftslagen (1998:204).

I arbetena med Sampers och Samgods har VTI inte hanterat sekretesskrav direkt från företagen. När VTI rekvirerar resvanedata från Trafikanalys eller Statistiska Centralbyrån anges användningsområdet. Det finns inget särskilt sekretessavtal.

För godstransporter har sekretessavtal slutits vid beställning av sekretesskyddade mikrodata från Trafikanalys och Eurostat.

Sammantaget har VTI kunnat använda data så som man har önskat.⁶⁶

Vid ansökan till Eurostat har VTI varit tvungna att formulera egna förhållningsregler för att hantera sekretessen. Eurostat anger att följande skall framgå av ansökan⁶⁷:

1. Organisationens syfte,
2. Organisationens (bevisade) förmåga att producera forskning av hög kvalitet och att göra resultaten offentligt tillgängliga,

⁶⁴ I projektet ingår bland annat KTH, VTI och personal från Sweco.

⁶⁵ Trafikverket (2017a), sid 10

⁶⁶ VTI anger att det är en mycket tidsödande process att rekquirera mikrodata från Eurostat.

⁶⁷ Eurostat (årtal okänt): Application form for research entities. Notabelt är att sådana forskningsenheter kan vara en del av organisationer med andra, övergripande syften som exempelvis privata, kommersiella företag.

3. Hur forskningsverksamheten är organiserad samt (där så är relevant) hur man garanterar det vetenskapliga oberoendet och
4. Åtgärder för att garantera datasäkerhet ("the safeguards in place to ensure security of the data").

Eurostat efterfrågar en beskrivning av åtgärder för att garantera datasäkerheten som omfattar såväl fysisk (exempelvis brand) som digital säkerhet. Eurostat vill också ha en beskrivning av hur tillgången till data begränsas avseende antal personer samt hur detta övervakas. Om ansökan gäller konfidentiella data skall syftet med det specifika (forsknings-)projektet anges.

Eurostat anger att hanteringen av personliga data regleras av regleringen (EC) 45/2001.

VTI menar i princip att sekretesskraven kan hanteras utan allvarliga inskränkningar i verksamheten. I någon mening bekräftas kanske den bilden av att VTI i ett underlagsarbete till Trafikanalys regeringsuppdrag endast nämner sekretess en enda gång, om än direkt i inledningen:

*"Företagens behov av sekretess anges ofta vara en anledning till varför en mer detaljerad nivå på statistiken inte kan redovisas."*⁶⁸

Trafikverket anger att det ofta kan vara svårt att få tillgång till uppgifter som man menar behövs för olika *utredningar* (inte specifikt för transportmodeller). Enskilda företag nekar att lämna uppgifter med hänvisning till sekretess på grund av konkurrensskäl.⁶⁹ Detta är alltså ett exempel på när myndigheten efterfrågar statistik som inte ingår i den officiella statistiken och därmed inte omfattas av uppgiftsskyldighet.

Rörande persontransportmodeller (specifikt Sampers) anger Trafikverket att det finns betydande kvalitetsproblem i de resvanedata som används för estimering/skattning (nationella resvanedata) respektive validering och kalibrering (regionala resvanedata).⁷⁰ En möjlig alternativ datakälla för validering och kalibrering är uppgifter från trafikoperatörer och huvudmän. År 2009 och 2013 samlade Trafikverket in resandestatistik för järnväg utifrån frivillig basis, där uppgiftslämnaren fick ange om de önskade sekretess för lämnade data. Många men inte alla har lämnat data. Samtidigt innebär förfarandet en osäkerhet kring kvaliteten avseende metod och osäkerhet.

Trafikverket anger också att myndigheten ibland har tillgång till vissa data som dock inte kan användas fullt ut av sekretess. Därmed kan man inte redovisa hur bra överensstämmelse som uppnås.⁷¹

När det gäller godstransportmodeller anger Trafikverket följande:

En försvårande omständighet för Trafikverket är att det oftast saknas åtkomst till existerande data om godstrafik (eftersom de omgärdas av sekretess eller inte delas på grund av olika skäl så som affärshemligheter). En svårighet för

⁶⁸ VTI (2017), sid 11

⁶⁹ Trafikverket (2017a), sid 3

⁷⁰ Trafikverket (2017a), sid 7

⁷¹ Trafikverket (2017a), sid 10 samt kompletterande dialog med Sachse, Carsten (2018)

modellering av godstransporter är därför den nödvändiga insamlingen av omfattande och pålitliga data.⁷²

Vidare skriver Trafikverket att sekretessproblematik försvårar arbetet med att ta fram kvalitetssäkrade prognoser. Myndigheten anger som exempel att man blir ålagda att göra prognoser inom arbetet med de transeuropeiska transportnätverken och -noderna där data för kalibrering inte kan erhållas på tillräckligt detaljerad nivå.⁷³

Trafikverket tar också upp frågan om sekretess kopplat till kostnadsdata och verkar ta denna allvarigare än vad som framförts från norsk sida i detta arbete.⁷⁴

I ett annat underlags-PM för Trafikanalys regeringsuppdrag intervjuades andra brukare än enbart Trafikverket.⁷⁵ Frågan om sekretess lyftes framför allt vid två sammanhang. Dels som en hänvisning till Trafikverkets undersökningar bland tågoperatörer på frivillig basis (se ovan), dels kopplat till risken för att enskilda individers resande skall kunna identifieras när dataunderlaget från exempelvis en resvaneundersökning redovisas på för detaljerad nivå. I detta arbete anger Swecos experter dock följande: Med uppgifter om kollektivtrafik (antal i fordon på olika delsträckor, antal påstigande, antal avstigande) kan mer validering göras. Mycket sådana data är konfidentiella, i synnerhet operatörsdata.⁷⁶

7.3 Sammanfattning

Det främsta skälet att inkludera en beskrivning av Sverige är att jämföra hur forskare (representerade av VTI), Trafikverket samt konsulter upplever sekretessaspekter kopplat till järnvägsdata och transportmodeller.

Det framgår tydligt att i alla fall Trafikverket anser att olika sekretesskrav, framför allt i form av affärshemligheter, inverkar menligt på arbetet med transportmodellerna. VTI verkar inte uppleva problem i samma utsträckning. Det har dock varit svårt att inom ramen för detta arbete identifiera vad skillnaden beror på. En hypotes är att VTI som forskningsinstitut har lättare att komma åt data (jämför Norge). En annan förklaring kan vara att VTI gör andra typer av analyser som kan ligga närmare forskning och/eller ställa krav på andra data. Vid en översiktlig genomgång av aktuella lagar och förordningar (se referenslistan) framgår inget tydligt svar. SCB anger i sin sekretesspolicy att data kan lämnas ut till olika organisationer för statistik- och utredningsändamål, men att det är villkorat att enskild person inte riskerar att lida skada.⁷⁷

När VTI exempelvis ansöker om data från Eurostat framgår det tydligt att alla typer av organisationer kan ansöka, så länge det är en forskningsenhet som står för ansökan. Av detta följer två observationer eller kriterier för att Trafikverket skall kunna få tillgång till data från Eurostat. Det första gäller om Trafikverkets utvecklingsarbete av transportmodeller kan ses som forskning. Det andra är frågan om Trafikverket är

⁷² Trafikverket (2017b), sid 7

⁷³ Trafikverket (2017b) sid 25

⁷⁴ Trafikverket (2017b), sid 9

⁷⁵ Sweco (2018)

⁷⁶ Källa: Sweco

⁷⁷ Statistiska Centralbyrån (2016)

organiserat på ett sådant sätt att man kan hävda att det är en forskningsenhet som gör ansökan.

Vidare kan man konstatera att Eurostat verkar ha besvärande långa handläggningstider och att man samtidigt begär en utförlig beskrivning av hur datas skall skyddas, dock utan att ange några kravnivåer.

8 Analys och slutsatser

8.1 Utmaningar

Det finns två grundläggande utmaningar kopplat till sekretess när det gäller järnvägsdata och transportmodeller. För det första måste data produceras på en tillräckligt detaljerad nivå för att möjliggöra kalibrering och validering av modellerna. För det andra måste data göras tillgängligt för de berörda brukarna på en tillräckligt detaljerad nivå.

Det finns inte oväntat ett tydligt, omvänt samband mellan de två utmaningarna. Ju mer detaljerade data, desto större restriktioner på användningen.

I kapitel 2 strukturerades ett dataflöde. Efter att ha belyst de olika länderna framgår det att de grundläggande utmaningarna är snarlika i kedjan från definition av databehov till resultatuttag. Det är intuitivt rimligt, eftersom det handlar om samma potentiellt känsliga objekt och subjekt i grunden.

Det finns dock anledning att också specificera utmaningarna lite ytterligare för de olika momenten. En anledning till detta är att metoderna för att garantera sekretessen skilja sig åt och att de valda metoderna i sin tur kan generera utmaningar. I resultatuttag kan man t.ex. aggregera data för att inte avslöja känslig information, vilket är svårare under kalibreringen och valideringen, då faktiska, detaljerade uppgifter (eventuellt) måste användas. Vilket i sin tur leder till krav på att exempelvis data behandlas inom universitet/högskolor eller av forskningsinstitut av personer med något slags ackreditering.

Det kan samtidigt vara relevant att identifiera metoder i hela kedjan, eftersom dessa kanske kan överföras mellan de olika momenten.

Utmaningarna kan vara en direkt följd av sekretessbehovet, men också en konsekvens av det sätt som den ansvariga myndigheten har valt att hantera data.

En aspekt som kommit fram i detta arbete är att det också kan finnas andra utmaningar än de rent sekretessmässiga som kan vara väl så viktiga och avgörande. Det gäller exempelvis frågor rörande kvalitet, arbetsbelastning för uppgiftslämnarna och andra.

I tabell 3 redovisas en sammanställning av utmaningar som framkommit i arbetet, såväl rörande sekretess som andra aspekter. I tabellen anges också i vilka länder utmaningen har träffats på. Tabellen gör inte anspråk på att vara heltäckande givet avgränsningarna för arbetet.

Tabell 3. Utmaningar rörande bland annat sekretess.

| Moment | Utmaningar rörande sekretess | Andra utmaningar |
|-----------------------------------|--|---|
| Identifiering av databehov | - | <ul style="list-style-type: none"> Databehovet beror av vilka modeller som används och kan ändras med nya modeller |
| Datainsamling | <ul style="list-style-type: none"> Få i synnerhet kommersiella aktörer att leverera potentiellt affärskänsliga data (Samtliga) | <ul style="list-style-type: none"> Insats för respondenter i form av arbete, kostnad (Norge, Sverige) Osäkerheter kring metod och kvalitet (Norge, Sverige) |
| Överlämning av data | <ul style="list-style-type: none"> Syftet med undersökningen definierar användningen (Nederländerna, Norge, Sverige) Byråkratisk och långsam rekvisering av data (EU) | <ul style="list-style-type: none"> Data ägs av de företag som samlat in dem (Tyskland) |
| Lagring av data | <ul style="list-style-type: none"> Data får bara lagras på den ansvariga myndighetens server, vilket gör arbete mer ineffektivt (Nederländerna, Sverige) Särskilda krav ställs på datasäkerhet (EU) | <ul style="list-style-type: none"> En fragmenterad lagring av data hos olika aktörer (Tyskland) |
| Nulägesbeskrivning | <ul style="list-style-type: none"> Beslutsfattare efterlyser detaljerade uppgifter om transportdata som kan avslöja enskilda personer (Norge, Sverige) | <ul style="list-style-type: none"> När data baseras på urvalsundersökningar kan felmarginalerna bli för stora (Norge, Tyskland) |
| Skattning/estimering | <ul style="list-style-type: none"> Modellernas algoritmer och annan kunskap skyddas av immaterialrättigheter (Danmark, Tyskland) Användningen begränsas till exempelvis forskare eller forskning. (Norge, Sverige) | <ul style="list-style-type: none"> Det kan finnas fördelar med en mer aggregerad modelluppbyggnad i form av enkelhet och transparens (Nederländerna) |

| Moment | Utmaningar rörande sekretess | Andra utmaningar |
|-----------------------------------|---|---|
| Kalibrering och validering | <ul style="list-style-type: none"> • Sekretess begränsar möjligheten att kalibrera och validera (Sverige) • Det är inte möjligt att ange överensstämmelsen mellan modell och verklighet (Sverige) | <ul style="list-style-type: none"> • Lagstiftning om rapportering anger inte krav att rapportera på detaljerad sträcka (Danmark, Norge, Sverige) |
| Resultatuttag | <ul style="list-style-type: none"> • Enskilda fysiska och juridiska personer skall inte kunna identifieras (Norge, Sverige) • Önskan om högre detaljeringsgrad i redovisningen (Norge) | <ul style="list-style-type: none"> • Kvalitetsaspekter vid allt för detaljerade data vid urvalsundersökningar (Norge, Tyskland, Sverige) |

8.2 Lösningar

Under arbetet har ett antal olika metoder för att hantera kraven på sekretess identifierats. Nedan beskrivs dessa i relation till de grundläggande utmaningarna; att säkerställa att *detaljerade data produceras* och att *data finns tillgängliga för utveckling av transportmodeller*. Följande metoder tas upp:

- Uppgiftsskyldighet, frivilliga överenskommelser eller avtal
- Begränsad men tillräcklig tillgång till detaljerade data

Praktiska tekniker som anonymisering, avidentifiering, statistiska metoder för röjandekontroll med flera behandlas inte närmare här. Likaså behandlas inte rent praktiska restriktioner som åtaganden att radera data efter användning. Anledningen är att om dessa tekniker och restriktioner fungerar är problemet löst. Om de däremot inte fungerar återvänder grundproblemet att data inte finns tillgängliga för utveckling av transportmodeller. I detta arbete har det inte heller framkommit några praktiska tekniker som torde överraska berörda parter.

Många av metoderna förekommer i flera länder, se tabell 4.

Tabell 4. Metoder för att hantera behovet av sekretess och samtidigt tillgodose krav på statistik om järnvägstransporter. Det garanteras inte att tabellen är uttömmande.

| Metod | Norge | Nederländerna | Tyskland | Danmark | Sverige |
|---|---|---|---|--|-------------------------------|
| Uppgiftsskyldighet | X | X | X | X | X |
| Frivilliga överenskommelser | | X (med operatörer) | | X | X |
| Avtal | X | | X (med operatörer om OD-matriser) | X (med operatörer om resandestatistik) | |
| Begränsningar i tillgång till grunddata - VEM | X (organisation) | X (syfte, specifikt projekt) | (arbete pågår - förslag att ge forskare tillgång till data) | | |
| Begränsningar i tillgång till grunddata - VARFÖR | X (syfte) | X (syfte, specifikt projekt) | | X (syfte) | |
| Begränsningar i tillgång till grunddata - HUR | X (praktisk hantering av data efter mottagande inkl radering) | X (behandling av data på myndighetens server) | | | X (åtkomst till data hos SCB) |

Samtliga studerade länder drar nytta av uppgiftsskyldighet för att säkerställa tillgång till data, även om data samtidigt omfattas av sekretesslagstiftningen. Nederländerna kompletterar systematiskt med frivilliga överenskommelser, där operatörerna får tillgång till transportmodellen i utbyte mot att man avslöjar data. Uppgifter från Danmark visar att frivilliga överenskommelser sluts per projektbaserat. I Sverige har de genomförts för att få tillgång till detaljerade resandestatistik, men metodmässiga utmaningar rörande bland annat täckning framhålls som en utmaning. I flera länder tillämpas också avtal med operatörerna för att säkerställa redovisning av data. I Tyskland och Danmark är det reglerat som ett grundkrav för operatörerna i samband med att de skall få trafikeringsrätt.

Samtliga länder använder också olika tekniker för att begränsa tillgängligheten till data, för att på så sätt kombinera dataskydd med nödvändig öppenhet för forskning och framtagande av statistik. Begränsningar sker utifrån någon eller en kombination av *vem* som önskar få tillgång till data, *varför* samt *hur* man får tillgång till uppgifterna. När det gäller vem som får tillgång till data har universitet/högskolor och forskningsinstitut typiskt lättare att få tillgång till data. Likaså ger syften som forskning eller forskningsnära verksamhet ofta enklare tillgång till data. Avseende hur man får tillgång till data tillämpar Nederländerna och Sverige ibland begränsningen att data endast finns tillgänglig på den nationella statistikbyråns server. Dessutom är utlämningen av data ofta förknippat med tydliga praktiska restriktioner, vilket bland annat illustrerats av Eurostats ansökanprocess som beskrivits i detta arbete i kapitel 7.

8.2.1 Uppgiftsskyldighet, avtal eller frivilliga överenskommelser

Uppgiftsskyldigheten spelar en stor roll i bland annat Sverige och Norge.⁷⁸ I Sverige regleras skyldigheten i Lag (2001:99) om den officiella statistiken. Denna konkretiseras vidare i Förordning (2001:100) om den officiella statistiken och ytterligare av statistikansvariga myndigheter i föreskrifter. Trafikanalys anger exempelvis i "Trafikanalys föreskrifter om uppgifter till statistik om sjöfart m.m.":

Uppgifter till statistik om sjöfart avseende fartyg, gods- och persontransporter m.m. ska lämnas av näringsidkare och kommuner, som valts ut med statistiska metoder, och vilka bedriver rederiverksamhet, hamn- eller terminalverksamhet eller hanterar gods på lastageplats.⁷⁹

I Finland regleras skyldigheten för järnvägsoperatörerna i Järnvägslagen Lag 304/2011:

Järnvägsoperatörerna ska trots affärs- eller yrkeshemlighet ge Trafikverket de upplysningar som verket behöver för att bestämma banavgifter enligt 37 § samt för statistik och forskning.⁸⁰

Uppgiftsskyldigheten ger principiell möjlighet att kräva av trafikföretagen delar av de data som behövs för att utveckla och förvalta transportmodeller. Det skall samtidigt noteras att det inte är givet att den svenska lagstiftningen som säkerställer den officiella statistiken ger mandat att kräva exempelvis detaljerade trafikdata som kan behövas för kalibrering och validering av transportmodeller. Förordningen (2001:100) om den officiella statistiken anger vilka uppgifter som näringsidkare är skyldiga att lämna⁸¹. Här ingår uppgifter om tillhandahållande, försäljning och leveranser av tjänster. Det finns dock inget som säger att de uppgifterna skall kunna kopplas till specifika sträckor. Samma situation föreligger i lagstiftningen i Danmark, varför lagkraven kompletteras med avtalskrav för operatörerna.⁸²

⁷⁸ Hovi, Inger Beate (2018)

⁷⁹ Trafikanalys föreskrifter om uppgifter till statistik om sjöfart m.m, Kammarkollegiets författningssamling, KAMFS 2012:2, TRAFAFS 2012:2

⁸⁰ Lag 304/2011 Järnvägslag 94 §, tredje stycket. <https://www.finlex.fi/sv/laki/alkup/2011/20110304>

⁸¹ Förordning (2001:100) om den officiella statistiken, 5 §

⁸² Naftchi, Yalda (2018)

Uppgiftsskyldigheten innebär samtidigt ett långtgående uppgiftslämnarskydd som begränsar eller reglerar användningen av data längre fram i processen. I Finland statistiklag (280/2004) framgår exempelvis att:⁸³

Vid sammanslagning, bevarande, förstöring och övrig bearbetning av uppgifter som samlats in för statistiska ändamål ska det ses till att skyddet för någons privatliv, skyddet av någons personuppgifter eller någons affärs- eller yrkeshemlighet inte äventyras.

Vidare skall arbetsbelastningen för de berörda uppgiftslämnarna beaktas. I Förordning (2001:100) om den officiella statistiken framgår det att insamlingen skall ske så att det blir 1) så enkelt som möjligt, 2) står i proportion till användarnas behov och 3) innebär en rimlig arbetsbörda för uppgiftslämnarna.⁸⁴

Frivilliga överenskommelser om uppgiftslämning bygger på att berörda organisationer accepterar att uppgifterna publiceras. Överenskommelsen bygger i princip på att berörda uppgiftslämnare bedömer att värdet för den egna organisationen och/eller samhället i övrigt överstiger kostnaden för att de egna uppgifterna publiceras. Statistiken över trafik i svenska hamnar kan fungera som illustration över skillnaden mellan uppgiftsskyldighet och frivillighet. Här publicerar Trafikanalys som statistikansvarig myndighet uppgifter om den totala trafiken till och från Sveriges hamnar.⁸⁵ Dessa aggregerade uppgifter ingår i Sveriges Officiella Statistik. Av sekretesskäl publiceras inte trafikvolymerna i enskilda hamnar. Det gör dock branschförbundet Sveriges hamnar, baserat på en frivillig överenskommelse mellan medlemmarna (hamnarna).^{86, 87}

Ur sekretessynpunkt innebär datainsamling som baseras på frivilliga överenskommelser stora fördelar framför insamling som grundas i uppgiftsskyldigheten, vilket gör att fortsatt datahantering faller under offentlighets- och sekretesslagen.⁸⁸ Däremot innebär frivilligheten andra utmaningar, inte minst när det gäller att få nödvändig uppslutning bland aktörerna en enhetlig och konsistent metodik samt att garantera tillgängligheten till data över tid. Trafikverket i Sverige vittnar om sådana svårigheter vid insamlingar av data bland tågoperatörer baserat på frivilliga överenskommelser.⁸⁹ Med andra ord riskeras andra kvalitetskriterier av vikt för statistiken. Lag (2001:99) om den officiella statistiken lyfter bland annat noggrannhet, tillgänglighet och tydlighet, jämförbarhet och samstämmighet (möjligheten att kombinera uppgifter).⁹⁰

Danmarks Tekniska Universitet anger att man ofta är beroende av frivilliga överenskommelser för att få tillgång till data.

⁸³ Trafikanalys (2017) refererar till Finlands statistiklag (280/2004)

⁸⁴ Förordning (2001:100) om den officiella statistiken, 4 §

⁸⁵ Trafikanalys: <https://www.trafa.se/sjofart/sjotrafik/>

⁸⁶ Hanefalk, Christine (2018).

⁸⁷ I grunden regleras uppgiftslämningen av EU-krav, som i sin tur reglerar Sveriges Officiella Statistik.

⁸⁸ Naturligtvis måste skydd av personuppgifter för fysiska personer hanteras enligt gällande lagstiftning.

⁸⁹ Trafikverket (2017a)

⁹⁰ Lag (2001:99) om den officiella statistiken, 3a§

I Nederländerna sammanställs OD-matriser för resandet av tågoperatörerna själva. I utbyte mot att de redovisar passagerardata på en detaljerad nivå blir de delägare av den nationella persontransportmodellen. Detta är ett bra exempel på hur de kommersiella aktörerna bidrar till en transparens och informationstillgänglighet som främjar marknadens funktion.⁹¹

Mellanformen med juridiska avtal kan i teorin säkerställa tillgänglighet till data över tiden, men innebär samtidigt en avsevärd byråkrati som eventuellt eller till och med sannolikt har en kontraproduktiv påverkan på uppgiftslämnarnas vilja att bidra med data. En tydlig mellanform är dock möjligheten att ställa krav på järnvägsoperatörer att lämna över data som en del av trafikeringsavtalet. Detta görs bland annat i Danmark och Tyskland.⁹²

En ytterligare möjlighet är att kombinera uppgiftsskyldighet med en överenskommelse om att lätta på restriktionerna för användningen av data i senare moment i processen. Det bör dock betonas att när data väl omfattas av offentlighets- och sekretesslagen är det svårt att göra avsteg från denna. Uppgiftslämnaren äger inte längre uppgifterna som levererats till exempelvis Statistiska Centralbyrån.

I tabell 5 beskrivs för- och nackdelar med uppgiftsskyldighet relativt frivilliga överenskommelser i dataflödesstrukturen från figur 1.

Tabell 5. För- och nackdelar samt andra aspekter (+, -, 0) med uppgiftsskyldighet och frivilliga överenskommelser.

| Moment | Uppgiftsskyldighet | Frivilliga överenskommelser |
|-----------------------------------|---|---|
| Identifiering av databehov | 0 Det är vanligt att lagstiftning eller riktlinjer anger att myndigheten skall föra dialog med branschföreträdare. ⁹³ + Garanterar kvaliteten i definition. | + Överenskommelsen utgör ett naturligt motiv för dialog. 0 Kvalitet i definition måste säkerställas. |
| Datainsamling | + Garanterar data + Säkrar i görligaste mån kvaliteten i metodik och täckning + Säkrar tillgänglighet över tid | + Möjlighet att samla in data som annars inte samlas in - Osäker tillgång över tid |
| Överlämning av data | 0 Begränsas av offentlighets- och sekretesslagen | + Kan göras friare/enklare. |
| Lagring av data | 0 Utgör ett verktyg för att säkerställa sekretessen. | + Inga direkta begränsningar. |

⁹¹ deJong (2018)

⁹² Naftchi, Yalda (2018) och Liedtke, Gernot (2018)

⁹³ Se exempelvis Finlands statistiklag 280/2004, 6 §.

| | | |
|-------------------------------|--|--|
| Nulägesbeskrivning | - Kan endast göras om röjande av enskilda personer inte riskeras. - All data får samma sekretessklass | + Kan ge möjlighet att publicera mer detaljerade data, men risken för röjande av uppgifter för tredje part måste kontrolleras + Ger möjlighet till en differentiering av sekretess utifrån hur känsliga data är |
| Skattning/estimering | - Tillgängligheten till data begränsas på olika sätt, primärt som en följd av sekretesslagstiftning | + Inga direkta begränsningar |
| Kalibrering/Validering | 0 Kan typiskt göras mot data på aggregerad nivå, men kan begränsas på detaljerad nivå av sekretesslagstiftning | + Kan ge möjlighet att nyttja mer detaljerade data, men risken för röjande av uppgifter för tredje part måste kontrolleras |
| Resultatuttag | - Kan endast göras om röjande av enskilda personer inte riskeras. | + Kan ge möjlighet att publicera mer detaljerade data, men risken för röjande av uppgifter för tredje part måste kontrolleras |

Uppgiftsskyldighet har de bästa förutsättningarna både att få in kompletta uppgifter och enligt tydliga, enhetliga metoder (i alla fall efter en successiv harmonisering över tiden). Den tydligaste nackdelen är den insats som krävs från uppgiftslämnarna och den insamlade organisationen. Det är dock en betydande och viktig aspekt. En möjlig nackdel är också att förståelsen för uppgiftsskyldigheten kan bli lidande om motiven inte förklaras.

Frivilliga överenskommelser kan appliceras i Sverige och gör det redan för exempelvis hamnstatistik, men riskerna med tillvägagångssättet är stora vad avser viktiga kvalitetsaspekter, att säkra tillgång till data över tid samt att få täckande undersökningar. För att öka villigheten att rapportera data kan något erbjudas i gengäld, exempelvis tillgång till transportmodellerna eller marknadsdata. Detta kan bidra till att marknadernas funktion och kan på längre sikt även innebära att man breddar basen för att kompetensförsörja transportmodellområdet.

8.2.2 Begränsa tillgängligheten till detaljerade data

En metod för att säkerställa sekretessen är att begränsa tillgängligheten till grunddata eller omvänt, att ge tillgång till data utifrån vissa förutsättningar. Tillgängligheten kan begränsas eller erbjudas utifrån bland annat syftet med studien och vem som utför arbetet.

Syftet med undersökningen

I bland annat Finland, Norge och Sverige får data som annars är sekretessbelagd användas för forskningssyfte. Ansökan beskriver syfte, vilka data som behövs samt vilka som skall utföra arbetet och prövas därefter. Beslutet kan innehålla olika typer av restriktioner för handhavandet.

En utmaning med begränsningen är att under utveckling av (nya) modeller, det vill säga utveckling av ny metodik, skattning samt initial kalibrering och validering, går det sannolikt relativt lätt att hävda att arbetet är forskningsrelaterat. I senare förvaltningskede när modellen skall kvalitetssäkras och valideras mot aktuella data kan det vara svårare att hävda att arbetet är forskningsnära.

Utförare

I Norge och Sverige finns större möjligheter för forskare att få tillgång till data och i Tyskland arbetar ansvarigt departement med att skapa en motsvarande möjlighet.⁹⁴ De olika ländernas traditioner i uppbyggnad och förvaltning av transportmodeller har viss påverkan här. I Norge och Sverige har arbetet med modeller skett forskningsnära och Trafikverket (motsvarande) har tagit ett stort ansvar. I Tyskland har konsultfirmor utvecklat och förvaltat modellerna delvis med statlig finansiering, men också till en hög grad med hjälp av kommersiella uppdrag för bland annat Deutsche Bahn. Detta har skapat en något otydlig situation kring olika aktörers möjlighet att få tillgång till data.

I Norge skapas tillgängligheten till data i flera steg⁹⁵:

1. Den organisation som skall använda data skall vara godkänd som forskningsinstitut.
2. Avtalet om utlån av data definieras för den aktuella studien och signeras av behörig företrädare för verksamheten.
3. Personal i verksamheten som skall utföra arbetet signerar avtal om tystnadsplikt.
4. När arbetet är avslutat skall utlånade data förstöras, vilket skall bekräftas av behörig företrädare för verksamheten.

Ansökan ska utformas så att alla som skall arbeta med materialet anges.

Det är intressant att systemet verkar fungera bra för de som kan få behörigheten. Detta bekräftas såväl av Transportøkonomisk institutt som av Väg- och Transportforskningsinstitutet.⁹⁶ En utmaning gäller dock att andra aktörer fortfarande har begränsad tillgång till uppgifterna och att publicering av data fortfarande omfattas av sekretess. Brukare av modellerna på exempelvis Trafikverket riskerar att bli beroende av forskare för att kunna utveckla modellerna vidare. Konsulter kan utestängas från modellutveckling, vilket innebär att den samlade kompetensen på modellområdet inte kan

⁹⁴ Hovi, Inger Beate (2018) och Liedtke, Gernot (2018)

⁹⁵ Hovi, Inger Beate (2018)

⁹⁶ Hovi, Inger Beate (2018) och Vierth, Inge (2018)

användas. Även om detta kan lösas genom exempelvis fler forskare på området så innebär det en sämre koppling mellan utveckling och tillämpning.

En delmetod som bland annat tillämpas av SCB är att databearbetning utförs av dem och levereras till konsulten efter vederbörlig röjningskontroll. Detta gör dock arbetet trögriktigt och ineffektivt, eftersom dialogen sker mellan parter (personer) med stor skillnad i kunskap om det aktuella sakområdet.

8.3 Avslutande konklusioner

Det övergripande intrycket från de studerade länderna är att sekretesskraven inte verkar innebära några allvarliga begränsningar i arbetet med transportmodeller. I arbetet har åtminstone fyra tänkbara förklaringar förts fram:

- De etablerade metoderna för att säkerställa att data både produceras och görs tillgängliga fungerar tillräckligt bra.
- Eftersom sekretesslagstiftningen har funnits under förhållandevis lång tid och dessutom är till stora delar internationellt implementerad genom bland annat EU, har ambitionerna för transportmodeller anpassats till vad som är möjligt med hänsyn till säkerställandet av sekretessen.
- Andra typer av utmaningar bedöms överskugga sekretessaspekterna. Det gäller bland annat den statistiska osäkerheten för detaljerade data vid urvalsundersökningar.
- Andra typer av fördelar med att använda mer aggregerade transportmodeller gör sekretessfrågan mer sekundär. Det gäller bland annat värdet av transparens/öppenhet och enklare uppdateringar.

Metoder både för att garantera att data finns och att de kan användas är snarlika i de olika länderna. Det är i sig inte så förvånande, då det handlar om nationer med en snarlik och i grunden EU-harmoniserat förhållningssätt vad gäller offentlighets- och sekretesslagstiftning. Likaså är transportmodellområdet snarlikt mellan i alla fall Danmark, Nederländerna, Norge och Sverige, medan Tyskland avviker på flera sätt.

Det sammantagna intrycket från studien är att det är fruktbart att kombinera olika typer av åtgärder för att säkerställa ett gynnsamt utvecklingsarbete av transportmodellerna. I Sverige framstår det som en möjlighet att öka samarbetet med operatörer för att få tillgång till passagerardata på specifika sträckningar, vilka kan användas för detaljerad kalibrering och validering. Här verkar det finnas möjligheter att lära av pragmatismen i Danmark och det gemensamma intresset i Nederländerna, manifesterat i ett gemensamt ägande av den nationella transportmodellen.

Det verkar också finnas en potential att samnyttja data som finns tillgängligt inom Trafikverket, om än med olika syften och på olika enheter.⁹⁷ Denna möjlighet har dock inte studerats närmare.

⁹⁷ Trafikverket (2017a) och Sachse, Carsten (2018).

Utöver beskrivningen av identifierade utmaningar och lösningar finns det ett antal ytterligare kommentarer som är värda att lyfta fram.

- Traditionen när det gäller transportmodeller i det aktuella landet har betydelse både för utmaningarna och metoderna vad gäller sekretess. I Tyskland innebär det svaga offentliga ansvaret för modellerna att data är spridda på flera olika aktörer och skyddas såväl av offentlig som kommersiell sekretess. I Danmark, där den dominerande transportmodellen har utvecklats av akademi och konsultföretag tillsammans, finns det goda möjligheter för företagen att få tillgång till data. Inom EU (Eurostat) ställs krav på att det skall vara en forskningsenhet som ansöker om tillgång till data, men den enheten kan finnas inom alla typer av organisationer.
- Beskrivningen av Sverige och även indikationer från Norge visar att forskningsinstitut verkar ha lättare att få tillgång till data med nödvändig detaljeringsgrad.
- De geografiska och demografiska förutsättningarna spelar roll för hur detaljerade uppgifterna behöver vara för att vara funktionella. Det visar exempelvis en jämförelse mellan Nederländerna och Norge, där 40 zoner i Nederländerna ger en betydligt finare och tätare indelning än 90 zoner i Norge.

Referenser

Danmarks statistik (2017). Statistikdokumentation for Jernbanetransport 2017 3. kvartal

Danmarks Tekniske Universitet (2018). www.landstrafikmodellen.dk Uttag 180210

deJong, Gerard; Director, Significance, Nederländerna, E-postkorrespondens januari-februari 2018

deJong et al (2011), "Distribution and modal split models for freight transport in The Netherlands", abstract, European Transport Conference Glasgow 2011

Finansdepartementet (1989/2016). Lov om offisiell statistikk og Statistisk Sentralbyrå (statistikkloven), se <https://lovdata.no/dokument/NL/lov/1989-06-16-54> Uttag 180131

Finland: Järnvägslag, Lag 304/2011

Finland: Statistiklag, Lag 280/2004

Hanefalk, Christine; branschchef, Sveriges hamnar, E-postkonversation januari 2018

Grimholm, Kristian; Rådgiver, Seksjon for Transport, reiselivs- og IKT-statistikk, Statistisk sentralbyrå, E-postkonversation februari 2018

Hovi, Inger Beate; Forskningsledare, Transportøkonomisk Institutt, Norge (2018). Intervju 2018-01-19, E-postkorrespondens januari 2018

Jernbanedirektoratet (2017). Jernbanestatistikk 2016

Lahelma, Harri; Senior advisor, Trafikverket, Finland (2018). E-postkorrespondens januari 2018

Liedtke, Gernot; Head of Department, DLR, Deutsche Zentrum für Luft- und Raumfahrt, Tyskland, intervju 2018-01-24, E-postkorrespondens januari-februari 2018

Naftchi, Yalda (med flera); Leader, Center for Transport Analysis, Danmarks Tekniske Universitet, E-postkorrespondens januari-februari 2018

Sachse, Carsten; Trafikanalytiker, Trafikverket (2018), E-postkorrespondens februari 2018

Samferdseldepartementet (2016). Prop. 1 S (2016 – 2017) Proposisjon til Stortinget (forslag til stortingsvedtak), for budsjettåret 2017.

Statistics Netherlands (2012). Statistical disclosure control, Methods series 12

Statistiska Centralbyrån (2015). Handbok i statistisk röjandekontroll

Statistiska Centralbyrån (2016). Sekretesspolicy

Statistisk Sentralbyrå (årtal okänt): "Avtale om utlån av anonyme statistikkopplýsingar til forskning mellom Statistisk Sentralbyrå (SSB) og [organisationens namn]"

Statistisk Sentralbyrå (årtal okänt): "Sletting av opplysninger"

- Statistisk Sentralbyrå (årtal okänt): "Taushedserklaering"
- Statistisk Sentralbyrå (2018). Jernbanetransport. Om statistikken.
<https://www.ssb.no/transport-og-reiseliv/statistikker/jernbane> Uttag 180131
- Statistisk Sentralbyrå (2018). Personvern og datasikkerhet
<http://www.ssb.no/omssb/personvern/personvern-og-datasikkerhet> Uttaga 180131
- Statistisk Sentralbyrå (2018). Personvern og datasikkerhet
<http://www.ssb.no/omssb/personvern/personvern-og-datasikkerhet> Uttag 180131
- Statistisk Sentralbyrå (2018). Varestrømsundersøkelsen.
<http://www.ssb.no/innrapportering/naeringsliv/vsu2014> Uttag 180131
- Sweco (2018). Samgods och Sampers som kunskapsstöd om järnvägstransporter. En användarstudie
- Sverige: Offentlighets- och sekretesslagen (2009:400)
- Sverige: Förordning (2001:100) om den officiella statistiken
- Tindrebäck, Silke; Seniorkonsult, Sweco (2018).
- Trafikanalys (2012). Föreskrifter om uppgifter till statistik om sjöfart m.m, Kammarkollegiets författningssamling, KAMFS 2012:2, TRAFAPS 2012:2
- Trafikanalysforum (2018). <https://www.trafikanalysforum.se/grunder-i-mikrosimulering>
 Uttag 180216
- Trafik- og Byggestyrelsen (2016). Förordning (nr 1458) om rapportering av data för analys och planering etc. på järnvägsområdet
- Trafikverket (2017a).
- Trafikverket (2017b).
- Transport-, Bygnings- og Boligministeriet (2018). <https://www.trm.dk/da/publikationer>
 Uttag 180210
- TØI (2014). Bedre data for kollektivtransporten, TØI rapport 1305/2014
- TØI (2016). Transportytelser i Norge 1946-2015, TØI rapport 1544/2016
- TØI (2017). Vurdering av det nasjonale modellsystemet for godstransport, TØI rapport 1559/2017
- Van der Waard, Jan (årtal okänt). Regional Model application in The Netherlands, Netherlands Institute for Transport Policy Analysis
- Vierth, Inge (med flera); Forskningsledare, Väg- och Transportforskningsinstitutet, Sverige, E-postkorrespondens januari-februari 2018
- VTI (2017). Modellering av järnvägstransporter. En översikt kring datakällor, VTI Notat 29-2017

Bilaga: Frågeformulär (svensk version)

Arbetet syftar till att föreslå möjliga lösningar för att kombinera sekretess och transparens vid validering och kalibrering av modeller för prognostisering och analys av järnvägstransporter på detaljerad nivå.

Genom hela statistikkedjan från insamling via kalibrering, validering och till resultatuttag skall företagens (och individers) behov av sekretess hanteras. Företag är tågoperatörer av godstrafik eller persontrafik, transportköpare, (regionala) trafikmyndigheter/bolag, logistikföretag och speditörer, åkerier, m.m.

1. Vilka trafik-/transportmodeller använder ni i ert land och för vilka syften?
 - Godstrafik:
 - Persontrafik:
2. Vilken data till trafik-/transportmodellerna samlar ni in från företag, myndigheter, m.fl.?
 - Godstrafik:
 - Persontrafik:
3. Specifikt: vilken data använder ni för kalibrering?
 - Godstrafik:
 - Persontrafik:
4. Specifikt: vilken data använder ni för validering?
 - Godstrafik:
 - Persontrafik:
5. Finns det krav på sekretess för den data ni samlar in/använder? Vilka då? Har olika data olika krav på sekretess?
6. Vilka krav på sekretess framför företagen (uppgiftslämnarna)?
7. Hur möter ni företagens behov/krav av sekretess?
8. Har ni kunnat använda insamlade data så som ni har velat?
9. Har ni stött på några problem kring detta? Vilka då? I vilket moment i datakedjan (se tabell nedan)? 1. Insamling, 2. Överlämning, etc.
10. Hur har ni löst dessa? Lagtext, krav i upphandling, frivilliga överenskommelser, klassificering av data, tät dialog med branschorganisationer, etc.
11. Är ni nöjda med lösningen eller hade ni velat ha en annan? Vad har hindrat den önskvärda lösningen?

12. Vilka dokument beskriver hanteringen av sekretess hos er? Var kan vi hitta dem?
 Finns det dokument som beskriver regleringar kring hantering av känsliga data i
 ditt land?

| Moment i kedjan | Utmaning | Lösning | Önskvärd lösning och varför den inte kunnat genomföras |
|-----------------------------|----------|---------|--|
| 1. Insamling | | | |
| 2. Överlämning | | | |
| 3. Lagring | | | |
| 4. Nulägesbeskrivning | | | |
| 5. Skattning och estimering | | | |
| 6. Kalibrering | | | |
| 7. Validering | | | |
| 8. Resultatuttag | | | |