



LUNDS UNIVERSITET

Ekonomihögskolan

Institutionen för informatik

Mobildata i kampen mot COVID-19

En kvalitativ studie om hur mobildata kan användas för att bromsa smittsamma sjukdomar

Kandidatuppsats 15 hp, kurs SYSK16 i Informationssystem

Författare: Nils Gunnarson
Simone Lewenhaupt

Handledare: **Christina Keller**

Rättande lärare: Björn Svensson
Benjamin Weaver

Mobildata i kampen mot COVID-19: En kvalitativ studie om hur mobildata kan användas för att bromsa smittsamma sjukdomar

ENGELSK TITEL: Mobile data in the fight against COVID-19: A qualitative study how mobile data can be used to curb infectious diseases

FÖRFATTARE: Nils Gunnarson, Simone Lewenhaupt

UTGIVARE: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Osama Mansour, PhD

FRAMLAGD: maj, 2022

DOKUMENTTYP: Kandidatuppsats

ANTAL SIDOR: 100

NYCKELORD: Smittsamma sjukdomar, COVID-19, big data, mobildata, rörelsedata

SAMMANFATTNING (MAX. 200 ORD):

Under COVID-19 samarbetade Folkhälsomyndigheten med Telia för att genom rörelsedata få insikter som skulle bidra till att begränsa pandemins framfart. Detta är första gången i historien ett folkhälsoinstitut samarbetar med ett telekomföretag, en möjlighet som uppstått tack vare att man kan samla in och analysera stora mängder data (big data). Syftet med denna studie är att genom en kvalitativ datainsamling undersöka hur mobildata har använts för att begränsa smittspridningen av COVID-19 samt att belysa dess möjligheter och utmaningar. Det som framkommit från den empiriska undersökningen är att rörelsedatan indirekt har varit ett verktyg för att bromsa pandemin genom att ligga till grund för rekommendationer, sprida kunskap till allmänheten samt prediktera sjukhusinläggningar. De utmaningar som framkommit är främst etiska utmaningar i form av att anonymisera datan och att arbeta transparent. Vidare identifieras tekniska utmaningar i form av exakthet och hur mycket man kan lita på datan. Framtida möjligheterna som har identifierats syftar till att man nu har en baslinje för en befolknings rörelsemönster som kan byggas ut. Vidare kommer man med hjälp av mobildata kunna precisera epidemiologiska modeller och prediktera sjukhusinläggningar mer omfattande i framtiden.

Innehåll

1	Introduktion.....	4
1.1	Bakgrund	4
1.2	Problemformulering	5
1.3	Forskningsfråga.....	6
1.4	Syfte	7
1.5	Avgränsningar	7
2	Litteraturgenomgång.....	8
2.1	Smittsamma sjukdomar	8
2.2	Big data	9
2.2.1	Begreppet big data.....	9
2.2.2	Big Data och folkhälsa	9
2.2.3	Utmaningar med big data inom folkhälsoområdet.....	11
2.3	Mobildata	12
2.3.1	Mobildatas funktionalitet	12
2.3.2	Användningsområden för mobildata.....	13
2.3.3	Utmaningar och begränsningar med mobildata	14
2.3.4	Andra länders användning av mobildata.....	15
3	Metod.....	17
3.1	Litteraturgenomgång	17
3.2	Val av metod	17
3.2.1	Kvalitativa studier	17
3.2.2	Semistrukturerade intervjuer	18
3.3	Intervjuguide och frågeguide	18
3.3.1	Intervjuguide	18
3.3.2	Frågeguide.....	19
3.4	Urval.....	19
3.5	Datainsamling.....	20
3.6	Analys.....	21
3.7	Etiska överväganden	21
3.8	Validitet och reliabilitet.....	22
4	Empiriskt resultat	24
4.1	Presentation av empiriskt resultat	24

4.2	Användning och nytta med mobildata	24
4.2.1	Användningen av mobildata.....	24
4.2.2	Nyttan med mobildata	25
4.2.3	Samarbetet mellan Telia och Folkhälsomyndigheten	28
4.2.4	Alternativa metoder.....	29
4.3	Utmaningar och risker	30
4.3.1	Tekniska utmaningar	30
4.3.2	Utmaningar med datakvalitet	31
4.3.3	Etiska utmaningar.....	32
4.3.4	Övriga utmaningar.....	33
4.4	Möjligheter och framtid	33
4.4.1	Mobildatans möjligheter	33
4.4.2	Mobildata med högre precision.....	35
4.4.3	Fortsatt arbete med mobildata	36
5	Diskussion.....	37
5.1	Resultatdiskussion.....	37
5.1.1	Användning och nytta med mobildata	37
5.1.1.1	Användning av mobildata	37
5.1.1.2	Nyttan med mobildata	38
5.1.2	Utmaningar och risker.....	39
5.1.2.1	Tekniska utmaningar	39
5.1.2.2	Utmaningar med datakvalitet	40
5.1.2.3	Etiska utmaningar.....	40
5.1.2.4	Övriga utmaningar.....	41
5.1.3	Möjligheter och framtid	42
5.2	Metoddiskussion.....	43
6	Slutsats	45
6.1	Förslag till vidare forskning	46
	Appendix 1 – Intervjuguide.....	47
	Appendix 2 – Frågeguide	48
	Appendix 3 - Intervjuguide (engelska).....	51
	Appendix 4 – Frågeguide (engelska)	52
	Appendix 5 – Transkribering intervju 1	56
	Appendix 6 – Transkribering intervju 2.....	64
	Appendix 7 – Transkribering intervju 3	71
	Appendix 8 – Transkribering intervju 4.....	80
	Appendix 9 – Transkribering intervju 5.....	88

Referenser..... 94

Tabeller

Tabell 3.4: Presentation av respondenter	20
--	----

Definitioner

Begrepp	Definitioner
Big Data	<i>“Big Data consists of extensive datasets—primarily in the characteristics of volume, variety, velocity, and/or variability—that require a scalable architecture for efficient storage, manipulation, and analysis.”</i> (Chang & Grady, 2019, p.6).
Mobildata	All typ av data som kan extraheras från en mobiltelefon såsom signaldata från telefonmaster, GPS-platsdata och Bluetoothdata (Urbaczewski & Lee, 2020).
Rörelsedata	Data som baseras på mobildata och specifikt används för spårning av rörelse. Kritisk datatyp för att sammanställa rörelsemönster och rörelsetrender (Telia, n.d.a).
Rörelsemönster/rörelsetrender	En analys av ett samhälle/en grups beteende och rörelsebehov baserat på sammanställd rörelsedata (Telia, n.d.a).
Dataset	En uppsättning av insamlad data (Jansen, Kovacs, Esko, Saluveer, Söstra, Bengtsson & Magpantay, 2021), till exempel, insamlad data över antalet hemarbetande under en specifik månad i Göteborgs kommun.
Anonymiserad data	Data som inte innehåller information om individer och tar bort identifierbarheten av personuppgifter för att omfattas av dataskyddslagarna (Ågren, Bjelkmar & Allison, 2021).
Aggregerad data	Sammanställda dataset beträffande en specifik tid eller plats, till exempel det totala antalet personer som reser från Stockholm till Malmö under en specifik vecka (Ågren, Bjelkmar & Allison, 2021).

1 Introduktion

1.1 Bakgrund

Den 30:e januari 2020 beskrev Dr Tedros Adhanom Ghebreyesus, generaldirektören för World Health Organization (WHO) utbrottet av COVID-19 i Wuhan, Kina som en folkhälsokris av högsta grad. Redan den 11 mars 2020 hade smittan ökat drastiskt globalt och virusinfektionen COVID-19 karaktäriserades av WHO:s generaldirektör som en pandemi (WHO, n.d). WHO underströk vikten av att alla länder aktivt behövde förbereda sig för sjukdomen och att COVID-19 med stor sannolikhet skulle komma att drabba hela världen. Folkhälsomyndighetens generaldirektör, Johan Carlson, hävdade att flera relevanta åtgärder redan hade vidtagits med fokus att bromsa fortsatt smittspridning och skydda de grupper som låg inom riskzonen (Folkhälsomyndigheten, 2020b). Under en presskonferens i april 2020 uttryckte FN:s generalsekreterare António Guterres att coronapandemin var den största krisen världen hade stått inför sedan andra världskriget och att ansvaret för att motverka smittan låg hos alla världens länder (HBL, 2020). Den stora frågan världen stod inför var dock hur spridningen av COVID-19 skulle motverkas.

Avsaknaden av vaccin vid utbrottet av COVID-19 i kombination med virusets förmåga att sprida sig snabbt gjorde det nästintill omöjligt att stoppa spridningen helt (Thunström, Newbold, Finnoff, Ashworth & Shogren, 2020). Istället lades tidigt fokus på att “flatten the curve”, vilket är en strategi som syftar till att bromsa samt fördröja spridningen av viruset mellan personer (Harris, 2020). “Flatten the curve” applicerades tidigt världen över främst för att reducera belastningen på sjukvården i väntan på vaccin (Thunström et al., 2020). En åtgärd som vidtogs för att platta ut smittkurvan var social distansering. Om befolkningen utövar social distansering genom att undvika allmänna ytor, publika sammankomster, stanna hemma, hålla avstånd och minimera sin rörelse menade experter att spridningen av smittan kunde saktas ner (Harris, 2020). Harris (2020) beskriver effekten av social distansering och klargör att utmaningen är att begränsa rörelsen hos världens befolkning. Harris (2020) menar dock att strategin enbart fungerar om en tydlig majoritet följer den.

För att kunna utvärdera en pandemis utveckling, restriktioner och prediktioner är det kritiskt att kunna följa upp människors rörelse. En stor mängd aggregerad data behöver samlas in och analyseras för att kunna kartlägga en befolknings rörelsemönster (Telia, n.d.a). I Sverige inledde Folkhälsomyndigheten tidigt ett samarbete med den svenska telefonoperatören Telia kring aggregerad och anonymiserad rörelsedata för att följa hur människor förflyttar sig (Telia, n.d.a). Telia Crowd Insights lanserades 2017 och är en tjänst som analyserar anonymiserad och aggregerad mobil nätverksdata och har innan pandemin använts för samhälls- och stadsplanering. Under COVID-19 fick Telia Crowd Insights uppgiften att undersöka hela Sveriges befolknings rörelse genom att spåra mobildata. Detta gör att man kan se beteenden och rörelsebehov för hela Sveriges befolkning (Telia, n.d.b). Det som mäts är människors rörelse mellan telefonmasters upptagningsområden (Folkhälsomyndigheten, 2022). Data Insights på Telia nämner att verksamheten förser Folkhälsomyndigheten med

precis och säker mobildata som gör att man kan se mönster i människors rörelser vilka annars hade varit svåra att kartlägga. Datan bidrar till viktiga insikter för att vidta åtgärder mot smittspridning (Telias, n.d.a). Telias data kan användas på olika geografiska nivåer, i Folkhälsomyndighetens fall har datan i huvudsak använts på region- och kommunnivå (Folkhälsomyndigheten, 2022). Med hjälp av Telias mobildata har Folkhälsomyndigheten veckovis rapporterat Sveriges befolknings rörelser och gjort jämförelseanalyser. Som exempel kan man se att resandet i Sverige under februari och mars 2022 ökade men fortfarande låg på en nivå ca 10 % under normalåret 2019 (Folkhälsomyndigheten, 2022).

Utöver Telia har även Google bidragit med mobildata extraherat från Google Maps som har varit till hjälp för sjukvården när viktiga beslut har fattats i kampen mot COVID-19 internationellt (Google, 2022). Rapporterna som Google har publicerat innehåller rörlighetstrender baserade på geografisk plats och olika platskategorier. Bland annat har rörelsemönster inom detaljhandeln, matbutiker, parker, stationer i kollektivtrafiken, apotek, bostadsområden och arbetsplatser kunnat åskådliggöras genom extrahering av data för antalet besök på platserna (Google, 2022). Det statistiska underlaget från Google består av anonymiserade dataset från personer som har aktiverat platshistorik i Google Maps inställningar (Google, 2022). Likt Telia har Google presenterat data för människors rörelsemönster och har kunnat följa utveckling av rörelsetrender (Edsberg Møllgaard, Lehmann & Alessandretti, 2021). I kontrast till Sverige har Singapore, Sydkorea, Israel och Kina använt sig av mobildata för övervakningspåring på individnivå. Ländernas myndigheter har genom en applikation kunna följa om individer har varit i kontakt med smittade personer (Mansour, Kajatzi & Ghazawneh, 2022).

Apple har också försett myndigheter och hälsoorgan med mobildata. Rapporterna speglar förfrågningarna om vägbeskrivning i Apple Maps och publiceras dagligen (Apple, n.d.). Förändringar av vägbeskrivningsförfrågningarna sedan den 13 januari 2020 beträffande kollektivtrafik, att köra och gå jämförs sedan med den dagliga förfrågan i en graf. Även denna data är anonym och länkas inte samman med ett personligt Apple-ID. Apple sparar heller ingen platshistorik (Apple, n.d.).

Sammanfattningsvis har mobildata från olika aktörer världen över använts under pandemin COVID-19 som ett verktyg för att följa människans rörelsetrender (Folkhälsomyndigheten, 2022). Hur mobildata kan bidra till att bromsa spridningen av smittsamma sjukdomar kvarstår dock att besvara. I denna uppsats kommer mobildatans användningsområden under COVID-19-pandemin att kartläggas som ett möjligt stöd för att begränsa smittspridning i framtida pandemier.

1.2 Problemformulering

I Sverige finns det idag cirka 60 anmälningspliktiga sjukdomar enligt smittskyddslagen (SFS 2004:168) och smittskyddsförordningen (SFS 2004:255). Sedan 2020 finns även COVID-19 med på Folkhälsomyndighetens förteckning över anmälningspliktiga sjukdomar (Folkhälsomyndigheten, 2020a). Utan vaccin och med en pandemi där döds- och smittfall ökade för varje dag (WHO, n.d) blev det viktigt att hitta fler strategier för att stoppa dess framfart. Hålla avstånd och så kallad social distansering är restriktioner man i Sverige har använt sig av under en lång tid (Regeringskansliet, 2022a). Det är däremot svårt att veta i vilken utsträckning restriktionerna följs och enligt Harris (2020) krävs det att minst tre fjärdedelar av befolkning följer social distansering för att man ska kunna "flatten the curve".

Samtidigt som en pandemi sprids över världen pågår även en så kallad fjärde industriell revolution där big data spelar en viktig roll (SMER, 2019). Digitala verktyg har under COVID-19 kunnat användas för att bland annat övervaka smittspridningen, forska om viruset, utveckla diagnosmetoder, vacciner och behandlingar samt se till att invånarna kan koppla upp sig och vara trygga på nätet (Europeiska Kommissionen, n.d.). Även mobildata har varit till hjälp under pandemin för bland annat Folkhälsomyndigheten då datan gett viktiga insikter i arbetet med att bromsa smittspridningen (Telia, n.d.b). Eftersom ämnet är outforskat och metoderna för att följa rörelsemönster har börjat användas innan de utvärderats kan det vara svårt att veta vilka metoder som är de mest effektiva i att bromsa smittspridning.

Ytterligare en utmaning med att implementera analysmetoder som initialt baseras på individers rörelsedata är att bibehålla samhällets förtroende i hanteringen av personuppgifter (Jansen et al. 2021). Mansour, Kajatzi och Ghazawneh (2022) beskriver hur respondenterna i deras enkätstudie tror att mobildata och dess analyser kan bidra till att bromsa smittspridningen av COVID-19. Däremot är majoriteten av respondenterna inte övertygade om att deras data är helt anonymiserad. Jansen et al. (2021) belyser vikten av att myndigheter och verksamheter bör efterfölja särskilt framtagna principer i behandlingen av personlig data då det ännu inte finns något ramverk för tillämpning av mobildata. Andra utmaningar som uppkom med mobildata var att Folkhälsomyndigheten tidigare inte hade arbetet med denna typ av data vilket krävde framtagning av nya processer och utbildning för de involverade individerna (Ågren, Bjelkmar & Allison, 2021). Utmaningen med att hitta ett syfte för hur mobildatan kunde användas för att bromsa smittspridning beskriver också Ågren, Bjelkmar och Allison (2021) som en komplex uppgift för både Telia och Folkhälsomyndigheten.

Dolley (2018) understryker hur viktig big datas roll är för folkhälsan i världen och menar att det kan vara till stor hjälp att spåra individer via mobiltelefondata eftersom människor ofta rör sig snabbt över stora områden och kan hinna smitta många på kort tid. Mobildatan ger möjlighet att fatta snabba beslut om åtgärder baserat på statistik och exakta datapunkter (Telia, n.d.). Enligt Dolley (2018) är det även viktigt att tänka på att big data som används för att fatta beslut har hög kvalitet i form av egenskaper som aktualitet, noggrannhet, fullständighet och tillförlitlighet. Mobildatan har legat som grund för många viktiga beslut under pandemin exempelvis för att utvärdera i vilken grad social distansering följs (Willberg, Järv, Väisänen & Toivonen, 2021). Tidigare forskning tyder på att användningen av mobildata för att bromsa smittspridning fungerar (Mansour, Kajatzi & Ghazawneh, 2022). Däremot finns det begränsat med forskning beträffande hur man bäst använder sig av mobildatan. Därför är målet med studien att undersöka hur mobildatan använts för att bromsa smittspridningen. Även möjligheter och utmaningar att arbeta med mobildata kommer att undersökas.

1.3 Forskningsfråga

Mot bakgrund av det identifierade problemområdet ämnar vi att undersöka forskningsfrågan:

Hur kan mobildata användas för att bromsa spridningen av smittsamma sjukdomar samt vilka utmaningar och möjligheter finns med detta arbets sätt?

1.4 Syfte

Syftet med denna studie är att genom en kvalitativ datainsamling undersöka hur mobildata har använts för att begränsa smittspridningen av COVID-19 samt att belysa dess möjligheter och utmaningar. Därmed blir syftet i förlängningen att undersöka hur analysen av mobildata kan bidra till att bromsa spridningen av smittsamma sjukdomar.

1.5 Avgränsningar

Mobildata inkluderar olika typer av data. Vi kommer avgränsa oss till rörelsedata som baseras på mobildata. Eftersom COVID-19 är den enda pandemin i Sverige där Folkhälsomyndigheten använt sig av mobildata kommer vi att utgå från COVID-19-pandemin. Vi begränsar även den empiriska datainsamlingen till Sverige och Danmark.

2 Litteraturgenomgång

Litteraturgenomgången redogör för de mest centrala begreppen kopplade till forskningsfrågan. Litteraturgenomgången inleds med att definiera begreppet smittsamma sjukdomar och beskriva hur svensk lag ser på smittsamma sjukdomar. Sedan ges en bakgrund till begreppet big data vilket leder in på big data och dess koppling till folkhälsan. Därefter beskrivs utmaningar och risker med big data på folkhälsoområdet. Litteraturgenomgångens sista del redogör för mobildatans funktionalitet, användningsområden och går sedan in på dess begränsningar och andra länders användande av mobildata under COVID-19.

2.1 Smittsamma sjukdomar

För att hindra spridningen av smittsamma sjukdomar i Sverige finns en smittskyddslag (SFS 2004:168). Målet med smittskyddslagen (SFS 2004:168) är att tillgodose befolkningens behov av skydd mot spridning av smittsamma sjukdomar. I ett pressmeddelande från socialdepartementet beskrivs regeringens definition av en smittsam sjukdom som:

“alla sjukdomar som kan överföras till eller mellan människor och som kan innebära hot mot människors hälsa. I smittskyddslagen och smittskyddsförordningen ingår ett 60-tal sjukdomar” (Regeringskansliet, 2022b, n.p.)

Dessa 60 sjukdomar återfinns på Folkhälsomyndighetens förteckning över anmälningspliktiga sjukdomar (Folkhälsomyndigheten, 2020a). Allmänfarliga sjukdomar definieras som:

“smittsamma sjukdomar som kan vara livshotande, innebära långvarig sjukdom eller svårt lidande eller medföra andra allvarliga konsekvenser och där det finns möjlighet att förebygga smittspridning genom åtgärder som riktas till den smittade” (Regeringskansliet, 2022b, n.p.)

och samhällsfarliga sjukdomar definieras som:

“allmänfarliga sjukdomar som kan få en spridning i samhället som innebär en allvarlig störning eller överhängande risk för en allvarlig störning i viktiga samhällsfunktioner och som kräver extraordinära smittskyddsåtgärder” (Regeringskansliet, 2022b, n.p.)

Dessutom beskrivs det att alla sjukdomar som förekommer i smittskyddslagen och smittskyddsförordningen är anmälningspliktiga. När en läkare upptäcker en anmälningspliktig sjukdom bör den anmälas till regionens smittskyddsläkare samt till Folkhälsomyndigheten (Regeringskansliet, 2022b). I 1 kap. 7 § första stycket smittskyddslagen (SFS 2004:168) beskrivs även att:

“Folkhälsomyndigheten ansvarar för samordning av smittskyddet på nationell nivå och ska ta de initiativ som krävs för att upprätthålla ett effektivt smittskydd. Folkhälsomyndigheten ska följa och vidareutveckla smittskyddet. Folkhälsomyndigheten ska följa och analysera det epidemiologiska läget nationellt och internationellt.”

2.2 Big data

2.2.1 Begreppet big data

Favaretto, De Clercq, Schneble och Elger (2020) genomförde en studie med 39 forskare på ämnet big data för att reda ut begreppets betydelse. Resultatet av studien visade att det var svårt att hitta en entydig definition av big data. Anledningen enligt Favaretto et al. (2020) kan vara att forskarna hade svårt att erkänna begreppet som ett kulturellt skiftande fenomen och därför upplevde en viss osäkerhet kring termen. Många forskare beskrev dock termen genom att koppla den till processer såsom datainsamling och databehandling medan andra beskrev termen med de traditionella V:na. Även United States National Institute of Standards and Technology definierar begreppet big data utifrån fyra V:

“Big Data consists of extensive datasets—primarily in the characteristics of volume, variety, velocity, and/or variability—that require a scalable architecture for efficient storage, manipulation, and analysis.” (Chang & Grady, 2019, p. 6)

Arena och Pau (2020) förklarar att det som karakteriserar big data är volymen och komplexiteten på grund av heterogeniteten i den information som kan samlas in. De exemplifierar var big data används genom att nämna Internet of Things (IoT), smarta bilar, sociala nätverk och industrier. De förklarar även att de branscher där det just nu spenderas mest pengar på big data är inom bank, sjukvård, försäkring, värdepappers- och investeringstjänster samt telekommunikation. Tre av dessa branscher återfinns inom finanssektorn där det finns många användningsområden för big data-analys såsom bedrägeriupptäckt, riskhantering och kundtjänstoptimering. Definitionen av big data-analys refererar till processen att samla in och analysera big data för att få ut användbar information till verksamheten (Arena & Pau, 2020).

Söktermen “big data” har ökat kraftigt sedan 2011 (Google Trends, n.d.). Enligt Sarangi och Sharma (2019) finns det tre tydliga möjliggörande faktorer till varför big data blivit så populärt under de senaste åren. Den första är den explosionsartade ökningen av mängden data när man gått från analog till digital lagring av data. Samtidigt som kapaciteten för datalagring har ökat, har den även blivit markant billigare, exempelvis erbjöd Google år 2017 15 GB gratis lagringsutrymme jämfört med 1980 när 1 GB kostade \$438 000 (Sarangi & Sharma, 2019). Den andra faktorn är ökad beräkningskraft, en iPhone 5 har 2.7 gånger beräkningskraften av en 1985 Cray-2 supercomputer. Med det sagt har hög beräkningskraft blivit något som även “vanliga” människor har tillgång till. Den sista möjliggörande faktorn Sarangi och Sharma (2019) nämner är framväxten av molntjänster. Idag kan mindre företag hyra in sig för att få tillgång till lagringsutrymme istället för att själva behöva investera i dyr infrastruktur. Det gör att många företag har möjlighet att hantera, analysera och lagra stora mängder data, det vill säga “big data”.

2.2.2 Big Data och folkhälsa

Big data och folkhälsa är inte något nytt fenomen. Redan under kriget 1853 använde Florence Nightingale sin kunskap inom matematik och statistik för att hjälpa den brittiska armén att sänka dödligheten på sitt sjukhus (Polson & Scott, 2018). Av allmänheten var hon mest känd som sjuksköterska men hon tillbringade större delen av sitt liv med att analysera data som grund för beslutsfattande. Vissa beskriver henne som att hon utförde samma arbete

som en modern forskare inom big data, bara att hon hade mindre mängd data och fick räkna manuellt. Hennes sjukdomsklassificeringssystem används än idag och ligger till grund för all modern epidemiologi och medicinsk dataforskning (Polson & Scott, 2018).

Till följd av den fortsatta tillväxten i volym och variation av tillgänglig data menar Dolley (2018) att big data kommer vara en viktig del i framtidens folkhälsa. Big data har enligt Dolley (2018) möjligheten att förbättra många olika områden inom folkhälsan såsom sjukdomsövervakning, förutspå sjukdomsrisker, riktade behandlingar samt att förstå sjukdomar. Även Khoury och Ioannidis (2014) menar att big data kan ge bättre insikter gällande orsak och utfall av sjukdomar, bättre läkemedelsmål gällande precisionsmedicin samt förbättrad förutsägbarhet och förebyggande av sjukdomar. Khoury och Ioannidis (2014) ställer sig dock skeptiska till att lita blint på big data. Khoury och Ioannidis (2014) menar att det krävs en starkare epidemiologisk grund för big data då de anser att den idag baseras mycket på tillgänglighet. De hävdar också att det bör finnas ett sätt att integrera kunskap som är baserad på en iterativ process för att tolka vad vi vet och inte vet gällande vetenskapliga discipliner vilket kräver kunskapshantering, kunskapsyntes samt kunskapsöversättning. För att ta del av fördelarna med big data menar Khoury och Ioannidis (2014) att det är viktigt att man alltid har "hela bilden" i åtanke för att undvika fel.

Mooney och Pejaver (2017) kategoriserar big data för folkhälsa i fem olika kategorier:

- Genomisk/metabolomisk data vilket är medicinsk data, den ger en översikt över den kompletta uppsättningen genetiska instruktioner som tillhandahålls av DNA:t (Mooney & Pejaver, 2017).
- Geospatial data är tidsbaserad data som återger en specifik plats på jordens yta, exempelvis mobildata. Med hjälp av geospatial data kan man få insikter och dra slutsatser om en större mängd människors rörelsemönster (Mooney & Pejaver, 2017).
- Elektroniska hälsodatabaser är sekundärdata hämtad ifrån elektroniska databaser såsom sjukhusjournaler (Mooney & Pejaver, 2017).
- Personlig datainsamling är när man på egen hand samlar in hälsodata exempelvis via en Fitbit (Mooney & Pejaver, 2017).
- Sökdata är data som samlas in från den elektroniska världen, exempelvis sökdata från Google eller inlägg på sociala medier (Mooney & Pejaver, 2017).

Datainsamlingen från var och en av dessa källor utnyttjar nya teknologier för att samla in större mängder data än vad som tidigare varit möjligt. Trots det ger de olika dataformerna fundamentalt olika möjligheter inom forskningen och praktiken för folkhälsan enligt Mooney och Pejaver (2017). Även om Mooney och Pejaver (2017) tydligt kategoriserat olika källor för big data kan en given datamängd naturligtvis innehålla mer än en kategori. Exempelvis beskriver Mooney och Pejaver (2017) att sjukhusdatabaser kan inkludera både elektroniska journaler över en given patients besök men också resultaten från sekvensering av hela dennes genom. Dessa typer av sammanslagna datamängder kan vara nyckeln till att identifiera etiologiska länkar vilket hittills har förbryllat forskare, såsom frågor gällande arv och miljö. Dolley (2018) och Khoury och Ioannidis (2014) nämner olika områden för hur big data kan användas för folkhälsan. De olika områdena passar in i flera av kategorierna Mooney och Pejaver (2017) nämner. Den mobildata Telia (Telia, n.d., a), Google (Google, 2022) och Apple (Apple, n.d.) använt sig av under COVID-19 är geospatial data och den typen av data denna uppsats ämnar att fokusera på.

Under COVID-19 samarbetade Folkhälsomyndigheten med Telia vilket är första gången i historien då ett folkhälsoinstitut samarbetar med ett telekomföretag, en möjlighet som uppstått

tack vare att man kan samla in och analysera stora mängder data (Ågren, Bjelkmar & Allison, 2021). Både Khoury och Ioannidis (2014) och Ågren, Bjelkmar och Allison (2021) understryker däremot hur viktigt det är att gå från data till insikt, det vill säga att man gör datan meningsfull så att den blir användbar och möjlig att agera utifrån. Det är något Telia lyckats med eftersom Folkhälsomyndigheten har kunnat basera beslut gällande exempelvis Sveriges restriktioner på deras data (Ågren, Bjelkmar & Allison, 2021).

2.2.3 Utmaningar med big data inom folkhälsoområdet

Eftersom big data i kombination med folkhälsa är ett relativt nytt område finns det ett antal utmaningar och risker. Dolley (2018) beskriver olika risker men påpekar även att riskerna behöver utvärderas och undersökas då det i framtiden kan komma att finnas fler och större risker. Även Khoury och Ioannidis (2014) beskriver en del utmaningar med big data och folkhälsan. De är dock mer skeptiska redan från början och understryker vikten av att vara medveten om utmaningarna. De menar att även om big data ger många möjligheter måste man vara medveten om och se upp för dess risker.

Integritet är en utmaning gällande big data och risken att man ska identifiera specifika individer är något både Dolley (2018) och Mooney och Pejaver (2017) nämner som ett problem. I Sverige är det enligt lagen med kompletterande bestämmelser till EU:s dataskyddsförordning (SFS 2018:218) mot lagen att utan samtycke lagra information om enskilda individer som går att identifiera. Därför har exempelvis Telia anonymiserat sin data (Ågren, Bjelkmar & Allison, 2021). Trots detta saknar många tilltro till myndigheter och tror att analyser av mobildata baseras på identifierbara personer (Mansour, Kajatzi & Ghazawneh, 2022).

Dolley (2018) menar också att man måste vara försiktig när man hanterar stora mängder data för att undvika "overfitting", avsaknad i noggrannhet eller felaktiga resultat. Då man riskerar att skada de allmänna förtroendet för big data och dess positiva effekter på folkhälsan. Användandet av big data för riskprediktion kan rädda många liv från smittsamma sjukdomar men man ska samtidigt vara försiktig med att ha en övertro på dessa modeller (Khoury & Ioannidis, 2014). Khoury och Ioannidis (2014) menar att det finns en risk i att big data inte alltid ger rätt svar. Vidare förklarar de att big datas styrka ligger i att hitta samband snarare än i att bevisa om sambanden faktiskt har betydelse vilket kan ge upphov till "falska alarm". Big data analyser kan idag baseras på praktiska urval eller på den mest tillgängliga datan, när detta görs är det viktigt att ha i åtanke att forskningsnoggrannheten inte är starkare än dess svagaste länk (Khoury och Ioannidis, 2014). Även Dolley (2018) nämner datakvaliteten som en utmaning, att ignorera egenskaper som aktualitet, noggrannhet, fullständighet eller tillförlitlighet leder till forskningssvaghet.

Vidare utmaningar är att användandet av big data i sjukdomsprevention riskerar att endast fungera i rikare delar av världen där det finns teknik och pengar till att samla in, hantera och analysera data (Dolley, 2018). I och med att bara vissa länder har tillgång till data som kan hindra sjukdomar från att spridas riskerar ojämlikheten i hälsa i världen att öka.

Dolley (2018) beskriver faran med att big data inom folkhälsan kan hamna i fel persons händer. Exempelvis kan information som är tillgänglig digitalt riskera att hamna hos oreglerade kommersiella företag eller försäkringsbolag. Det finns också risk att personer hackar databaser för att tjäna pengar. Ett exempel är ett dataintrång i Finland där psykoterapiföretaget Vastaamo krävdes på 40 bitcoin för att hackarna inte skulle läcka deras

patientjournaler. När företaget sedan vägrade betala började journaler att läckas och istället krävdes patienterna på lösensummor för att få sin journal borttagen från nätet (Lindroos-Hovinheimo, 2020). Ett annat problem Dolley (2018) nämner kan vara att lagar ligger efter vilket kan göra att saker som borde vara olagliga fortfarande är lagliga på grund av att de inte hunnit regleras.

2.3 Mobildata

Mobildata är som big data ett brett begrepp och definieras som all typ av data som kan extraheras från en mobiltelefon (Urbaczewski & Lee, 2020). Urbaczewski och Lee (2020) förklarar att mobildata delvis kan vara signaldata från en mobiltelefon vilket också benämns mobiloperatörsdata (Jansen et al. 2021). Vidare förklarar Urbaczewski och Lee (2020) att mobildata också kan baseras på GPS-platsdata och Bluetooth-platsdata. Mobildata kan också extraheras av plattformar som Google och Facebook vilket även är datakällor som använts för att simulera rörelsetrender under COVID-19 (Edsberg Møllgaard, Lehmann & Alessandretti, 2021).

I denna uppsats har vi i huvudsak fokuserat på mobildata som har en direkt koppling till spårning av en individ eller en grups rörelse. Det är fallet med Telias data som har använts under COVID-19 för att följa upp rörelsemönster (Ågren, Bjelkmar & Allison, 2021).

2.3.1 Mobildatas funktionalitet

Historiskt sett har modeller för infektionssjukdomar antagit att populationer inom ett definierat område såsom en stad eller region är väl beblandade med varandra. Det vill säga att varje individ har lika stor sannolikhet att interagera med alla individer inom det definierade området (Wesolowski, Buckee, Engø-Monsen & Metcalf, 2016). När rörelsemönster beaktas på detta sätt för att hitta ett samband mellan exempelvis städer och populationens rörelse kvantifieras datan med relativt enkla modeller. Sådana förenklade modeller har ofta visat sig misslyckas att beskriva befolkningens beblandning och rörelsemönster i brist på detaljerad realtidsdata (Wesolowski et al., 2016).

Den största nordiska telekomoperatören Telia och Folkhälsomyndigheten inledde under början av pandemin COVID-19 ett samarbete. Förhoppningen var att kunna använda mobildata för att följa befolkningens rörelsemönster och med hjälp av den aggregerade datan kunna bromsa smittspridningen av COVID-19 (Ågren, Bjelkmar & Allison, 2021). Innan denna typ av rörelsedata fanns tillgänglig har myndigheter följt upp journaldata manuellt för att härleda individers resmönster (Wesolowski et al., 2016). Ett exempel som Wesolowski et al. (2016) presenterar är användningen av mässlingens säsongvariationer för att fånga barns rörelser till och från skolan inom städer.

Ågren, Bjelkmar och Allison (2021) beskriver att Telia Crowd Insights data baseras på aggregering och anonymisering av befintlig signaldata mellan mobiltelefoner och mobilmaster. Författarna bryter ner processen i tre steg. I första steget anonymiseras samt aggregeras datan. Anonymiseringen använder bland annat k-anonymitet, som syftar till att minst fem mobila enheter måste uppvisa samma rörelsebetende för att datasetet ska få användas i aggregeringen (Ågren, Bjelkmar & Allison, 2021). Telenor Norge har haft ett liknande samarbete med Folkhälsoinstitutet som är Norges motsvarighet till Sveriges

Folkhälsomyndighet. Telenor Norge har till skillnad från Telia Sverige baserat sin k-anonymitet på grupper av 20 personer (Mansour, Kajatzi & Ghazawneh, 2022).

Steg två i processen av Telias datainsamling består av extrapolering, vilket syftar till att anpassa Telias marknadsandel inom ett geografiskt område för att estimeras den okända datan och sedan kunna representera hela populationen. I steg tre som är det slutgiltiga steget avgörs den spatiala precisionen av platsen för mobilmaster kombinerat med tätheten av mobilmaster och antalet signaler från mobila enheter. Den högsta spatiala upplösningen Telia har är 500 x 500 meter och den längsta är 16 x 16 kilometer (Ågren, Bjelkmar & Allison, 2021). Telias dataleverans till Folkhälsomyndigheten gjordes med 48 timmars fördröjningsintervall för att följa lagen och undvika etiska konflikter. Även Telenor Norge har en liknande process för att aggregera och anonymisera rörelsedatan (Mansour, Kajatzi & Ghazawneh, 2022).

2.3.2 Användningsområden för mobildata

Innan COVID-19s utbrott undersöktes användningsområdena för hur mobildata skulle kunna förenkla processen att följa infektionssjukdomars rörelsemönster (Wesolowski et al., 2016). Tidigare forskning har identifierat möjligheten att med mobildata generalisera populationers rörelsemönster inom ett specifikt land vilket är relevant för att följa olika typer av virusutbrott (Wesolowski et al., 2016). Mobildata möjliggör en befolkningsuppskattning som tidigare har varit otillgänglig. Detta har gjort att man kan analysera rörelsetrender och förstå geografisk spridning av en smittsam sjukdom bättre (Oliver, Lepri, Sterly, Lambiotte, Deletaille, Nadai, Letouzé, Ali Salah, Benjamins, Cattuto, Colizza, Cordes, Fraiberger, Koebe, Lehmann, Murillo, Pentland, Pham, Pivetta, Saramäki, Scarpino, Tizzoni, Verhulst & Vinck, 2020). Både Telia Sverige och Telenor Norge har försett respektive lands myndigheter med denna typ av data under COVID-19. Datan har sedan använts för att analysera de olika populationernas rörelsetrender (Mansour, Kajatzi & Ghazawneh, 2022). Studier har även visat att mobildata innan COVID-19 har bidragit till ökad effektivitet i arbetet av att följa rörelsemönster hos grupper i syfte att bekämpa smittsamma sjukdomar såsom kolera, ebola, denguefeber, hiv, malaria (Dolley, 2018).

Tidigare forskning menar att finskalig och individuell rörelsedata är möjlig att följa upp med mobildata och kan användas för att definiera en mer konkret sjukdomsrisk (Wesolowski et al., 2016). I Sverige har man inte använt sig av någon individbaserad rörelsedata då detta tillvägagångssätt strider mot GDPR-lagstiftningen. Enligt GDPR ska anonymiserad data inte kunna åskådliggöra en individ i en större grupp, länka samman två dataset som berör samma individ eller härleda okänd information om en individ (Jansen et al., 2021).

Enligt Ågren, Bjelkmar och Allison (2021) har Telia Crowd Insights mobildata huvudsakligen använts för att kvantifiera och följa upp hur mottagliga Sveriges population har varit till att efterfölja Folkhälsomyndighetens reserestriktioner. Detta har gjorts genom att jämföra rörelsedata under liknande perioder 2019 då pandemin inte fanns vilket också kan benämnas som baslinjeperiod (Ågren, Bjelkmar & Allison, 2021). Genom att Telia tillgängliggjort rörelsedatan har Folkhälsomyndigheten kunnat förstå pandemins framfart för att skydda befolkningen vilket är något som Wesolowski et al. (2016) tidigt förutspådde som en möjlighet med mobildata. Oliver et al. (2020) belyser även hur mobildata och rörelsemönster kan fastställa vilka restriktioner samt åtgärder som har fungerat effektivt i arbetet att begränsa spridningen av COVID-19.

Utöver detta har Telias mobildata använts för att veckovis informera det svenska folket hur resandet i landet mellan regioner har varit. Veckorapporterna har också simulerat antalet tjänsteresor som utförts. Populationens anpassning till arbete i hemmet, rörelsetrender under högtider och resor inom Norden är händelser Folkhälsomyndigheten har offentliggjort med hjälp av mobildata (Ågren, Bjelkmar & Allison, 2021).

Google har även använt rörelsedata för liknande estimeringar som Telia gällande analyser av rörelsemönster. Enligt studien kan mobiloperatörsdata samt Googles data användas för att förutse antalet sjukhusinläggningar i Sverige på en regional nivå (Gerlee, Karlsson, Fritzell, Brezicka, Spreco, Timpka, Jöud, Lundh, 2021). Oliver et al. (2020) bekräftar i deras artikel att mobildata kan användas för att utföra prediktiva analyser och identifiera framtida smittspridningar genom att följa rörelsemönster.

2.3.3 Utmaningar och begränsningar med mobildata

En begränsning som Mansour, Kajatzi och Ghazawneh (2022) nämner beträffande användningen av mobildata för att förstå rörelsemönster under en pandemi är att datan saknar exakthet. Både Mansour, Kajatzi och Ghazawneh (2022) och Ågren, Bjelkmar och Allison (2021) bekräftar att exaktheten av mobildatan varierar beroende på hur tätt signalmasterna förhåller sig till varandra samt antalet signaler från mobila enheter. Tidigare forskning har även föreslagit att den exakta platsen från mobildata aggregeras till en bredare miljö vilket bidrar till att precisionen för rörelsemönster avtar (Wesolowski et al., 2016). En annan aspekt med Telias mobildata är att den fungerar som bäst inom ett tätbefolkat område där det finns mycket data att extrahera vilket skapar brister i att representera alla Sveriges städer och regioner med exakthet (Mansour, Kajatzi & Ghazawneh 2022).

Ytterligare ett hinder som Mansour, Kajatzi och Ghazawneh (2022) identifierade i sin enkätstudie var att majoriteten av respondenterna inte tror att organisationerna eller myndigheterna som hanterar datan kan garantera individuell anonymitet. Trots att respondenterna generellt trodde att mobildata skulle kunna vara fördelaktig för Folkhälsomyndighetens förståelse av rörelsemönster. Utmaningen att bibehålla allmänhetens förtroende vid användningen av mobildata för att forma rekommendationer eller policier är något Jansen et. al (2021) bekräftar. Oliver et al. (2020) förklarar även oron som finns om övervakningsverktyg som har använts under COVID-19 ska fortsätta användas när nödsituationen är över. Detta skulle i längden kunna kränka personlig integritet i demokratiska länder (Oliver et al., 2020).

Jansen et. al (2021) presenterar även en sammanställd checklista innehållande principer som rekommenderas att följas av myndigheter så att allmänheten kan behålla förtroendet för hanteringen och användningen av mobildata. Principerna ska kunna appliceras och efterföljas i framtida situationer som kräver brådskande åtgärder (Jansen et. al, 2021).

Kopplingen mellan rörelsemönster och smittspårning av en infektionssjukdom i realtid är även en utmaning som Wesolowski et al. (2016) nämner. En koppling mellan ett specifikt fall av en infektionssjukdom och ett specifikt rörelsemönster som kan länkas till en enskild individ kommer enligt Wesolowski et al. (2016) sannolikt aldrig att kunna hanteras på grund av mobilanvändares anonymitet. Jansen et. al (2021) belyser även att individuella data som samlas in av statistiska myndigheter för sammanställningar ska vara strikt konfidentiella och enbart användas för statistiska ändamål, oavsett om datan avser fysiska eller juridiska personer. Dock förklarar Wesolowski et al. (2016) att man med stor noggrannhet kommer

kunna att åskådliggöra specifika smittsamma infektionsutbrott med hjälp av mobildata men att det kommer att finnas ett bortfall i form av underrapportering eller förekomsten av asymtomatiska individer.

Mansour, Kajatzi och Ghazawneh (2022) belyser potentialen av att använda realtidsdata från mobiltelefoner för att kunna åskådliggöra en befolknings rörelsetrender. Författarna menar att denna typ av rörelsedata ska kunna hjälpa till med snabb upptäckt av virusutbrott och underlätta beredskapen samt åtgärder för folkhälsomyndigheter. Dock belyser även Mansour, Kajatzi och Ghazawneh (2022) utmaningen Telia har att leverera realtidsdata som är nyare än 48 timmar på grund av legala skäl i anonymiseringsprocessen.

Slutligen beskriver Oliver et al. (2020) att det finns liten politisk vilja samt investerade resurser i att stödja förberedelser för omedelbara åtgärder vid en framtida pandemi. Författarna förklarar att det är samspelet mellan länder, mobiloperatörer, myndigheter och teknikföretag som delvis begränsar möjligheten av att använda mobildata för att bekämpa smittsamma sjukdomar.

2.3.4 Andra länders användning av mobildata

Mobildata har under COVID-19 använts i flera olika länder i syfte att bromsa smittspridningen (Urbaczewski & Lee, 2020). Ågren, Bjelkmar och Allison (2021) fokuserar främst på hur Folkhälsomyndigheten i Sverige har använt sig av Telias data för att hitta generella rörelsemönster för den svenska befolkningen. Arbetet innefattar inga spårningar av individuella rörelser utan fokuserar på grupper och trender till skillnad från länder som har bevakat rörelsedata på individnivå (Urbaczewski & Lee, 2020). Till exempel har Israel använt sig av individbaserad mobildata för att identifiera personer som har varit i kontakt med andra individer som testat positivt för COVID-19 (Mansour, Kajatzi & Ghazawneh, 2022).

Under COVID-19 har mobildata även använts för kontaktspårningsapplikationer av myndigheter i omkring 50 länder (Urbaczewski & Lee, 2020). Applikationerna använder vanligtvis GPS-signal och Bluetooth för att spåra rörelse. Enligt Urbaczewski och Lee (2020) har vissa applikationer varit frivilliga att använda medan andra nationer har haft som krav att ladda ner applikationen om en person har umgåtts med en smittad individ. I Singapore bad premiärministern befolkningen att installera applikationen TraceTogether på sina privata mobiltelefoner för att kunna följa smittutbrott (Mansour, Kajatzi & Ghazawneh, 2022). Även Sydkorea krävde att befolkningen skulle ladda ner en applikation för bevakning av karantän och vilka personer man har umgåtts med (Urbaczewski & Lee, 2020). En utmaning när ett land tvingar befolkningen att dela sin personliga data utan samtycke är att förtroendet för landets myndigheter avtar. För att behålla förtroendet för officiell statistik behöver de ansvariga myndigheterna förhålla sig till professionella överväganden, yrkesetik samt vetenskapliga principer (Jansen et al., 2021).

I flera länder har Apple och Google använt sig av mobildata i arbetet att bromsa smittspridningen av COVID-19. De spårningsmetoder Apple och Google använder sig av baseras på att dela informationen mellan varje enskild smartphone-användare snarare än att datan förvaras av en statlig myndighet som Sverige har gjort (Urbaczewski & Lee, 2020). Företagen anser att denna typ av datainsamling hjälper till att varna personerna som har blivit utsatta för viruset samtidigt som man förhindrar potentiell kränkning av mänskliga rättigheter beträffande integritet (Urbaczewski & Lee, 2020).

I Danmark har man använt mobildata för att göra liknande analyser som har gjorts av Folkhälsomyndigheten och Telia i Sverige. Forskare har jämfört Facebooks och danska mobiloperatörers data i syfte i att följa upp hur effektiva reserestriktioner har varit för att minska spridningen av COVID-19 i Danmark, Spanien, Frankrike och Italien (Edsberg Møllgaard, Lehmann & Alessandretti, 2021). Författarna förklarar att de med hjälp av aggregerad mobildata kan se olika resetrender under arbetsdagar, helger respektive helgdagar. I perioder med strikta reserestriktioner minskade resandet under arbetsdagar drastiskt medan helgresorna ökade inom Danmark (Edsberg Møllgaard, Lehmann & Alessandretti, 2021). Författarna förklarar även att datan från danska mobiloperatörer var mer representativ än vad Facebooks data i Spanien, Frankrike och Italien var (Edsberg Møllgaard, Lehmann & Alessandretti, 2021). Ågren, Bjelkmar och Allison (2021) framhäver också att Telias data i Norden har kunnat representera hela populationer med hjälp av den aggregerade och anonymiserade mobildatan. Facebooks data är heller inte lika precis som mobildata från mobiloperatörer då datan inte tar hänsyn till resor inom områden. Exempel på detta är om individer har rest inom sitt bostadsområde så spårar inte Facebooks data den rörelsen förklarar Edsberg Møllgaard, Lehmann och Alessandretti (2021). Det har gjort att datan har varit mindre tillförlitlig för att analysera rörelsetrender i Frankrike, Italien och Spanien enligt Edsberg Møllgaard, Lehmann och Alessandretti (2021).

3 Metod

I metodavsnittet presenteras den metod som använts för undersökningen. Inledningsvis presenteras hur litteraturen samlats in, sedan valet av metod och sammanställandet av intervju- och frågeguide. Vidare presenteras urval, datainsamlingsmetoder samt analys. Avslutningsvis beskrivs etiska överväganden samt validitet och reliabilitet.

3.1 Litteraturgenomgång

Litteraturgenomgången är en viktig del av uppsatsen och utgör grunden för forskningsfrågan och metoden (Bryman & Bell, 2017; Rienecker & Stray Jørgensen, 2008). Litteraturen har studerats i syfte att utvärdera tidigare forskning på området, utreda relevanta begrepp och teorier, klargöra om det finns motsättningar i litteraturen samt om det finns obesvarade frågor inom området (Bryman & Bell, 2017).

För att hitta litteratur har vi främst använt oss utav Google Scholar och LUBsearch men även Google när vi använt mer aktuella källor från Folkhälsomyndigheten, WHO, Telia, Google, Apple, Regeringskansliet samt Europeiska kommissionen. De sökord vi huvudsakligen använt oss av enskilt och i kombination är följande begrepp:

- Smittsamma sjukdomar
- Smittskyddslagen
- COVID-19
- Big data
- Public health
- Mobility trends
- Mobility data
- Movement data
- Mobile data
- Telia
- Anonymised and aggregated data

Ambitionen har varit att utgå från att artiklarna ska vara peer-reviewed och välciterade för att säkerställa så hög kvalitet som möjligt. Peer-review innebär att andra forskare har granskat materialet samt gått i god för dess kvalitet vilket stärker nivån av vetenskaplighet (Rienecker & Stray Jørgensen, 2008). Eftersom ämnet är relativt nytt och aktuellt har få artiklar hunnit publiceras eller peer-reviewas vilket gjort att det inte varit möjligt att använda dessa källor i den utsträckning vi från början hade tänkt.

3.2 Val av metod

3.2.1 Kvalitativa studier

Enligt Jacobsen (2002) är en kvalitativ metod att föredra när man vill få mer klarhet i vad som ligger i ett begrepp eller fenomen. Med en kvalitativ metod kan man få fram hur människor

tolkar och förstår en given situation. Vidare är en kvalitativ metod fördelaktig för att få klarhet och en nyanserad bild av ett nytt ämne (Jacobsen, 2002). Målet med vår studie var att få en detaljerad och djupgående bild av hur mobildata kan användas för bromsad smittspridning samt att eventuellt kunna finna nya perspektiv på ämnet, därför lämpade sig en kvalitativ studie bäst (Jacobsen, 2002). Den kvalitativa metoden öppnar även upp för oförberedda händelser eller åsikter och är därför bra för att utveckla samt nyansera nya teorier (Jacobsen, 2002). Jacobsen (2002) beskriver också fördelen med att kunna analysera data efterhand den inkommer och kunna anpassa efterföljande datainsamling, vilket var målet att kunna göra under våra intervjuer. Metodvalet lägger även vikt vid detaljer vilket skapar unika nyanser hos varje respondent under intervjuerna (Jacobsen, 2002).

3.2.2 Semistrukturerade intervjuer

Då vår frågeställning syftar till att utreda hur mobildata kan bidra till att bromsa smittspridning är målet med studien snarare “discovery” än “checking”, därför passar en semistrukturerad intervju bra för vår empiriska datainsamling (Oates, 2006). Jacobsen (2002) förklarar att vissa menar att man vid en semistrukturerad intervju går miste om hela syftet med en kvalitativ undersökning. Jacobsen (2002) lyfter dock två motargument och belyser att för de första så kan det bli svårt och resurskrävande att analysera en helt öppen intervju. För det andra är i stort sett inga intervjuer helt öppna. Någon strukturering har alltid gjorts i förväg, om än omedveten. Vidare beskriver författaren att en förstrukturering inte behöver göra en intervju sluten utan snarare att enskilda aspekter sätts i fokus vilket varit vårt mål när vi valde semistrukturerad intervju som datainsamlingsmetod. Jacobsen (2002) menar också att en av styrkorna med kvalitativa studier är att man kan ändra upplägget på intervjuerna efterhand som de tar form vilket bidrog till vårt beslut att använda semistrukturerade intervjuer som empirisk undersökningsmetod. Då vi valt att intervjua personer med olika bakgrund kan vikten i intervjuerna läggas på olika delar varje gång samt att man kan ställa olika följdfrågor, exempelvis om en fråga behöver ytterligare förklaring. Dessutom rekommenderar Jacobsen (2002) denna typ av undersökning i de fall då ämnet inte är speciellt beforskat ännu vilket stämmer in på vårt fall där ämnet är relativt utforskat. Oates (2006) nämner även att en fördel med semistrukturerade intervjuer är att respondenten får möjlighet att själv belysa de områden de anser relevanta och att ta upp ämnen utanför intervjun. Det var relevant för oss då målet med de semistrukturerade intervjuerna var att respondenter med olika bakgrund skulle ge svar på vad de själva fann mest relevant kopplat till vår frågeställning.

3.3 Intervjuguide och frågeguide

3.3.1 Intervjuguide

Intervjuguiden (Appendix 1) har skapats baserat litteraturgenomgångens mest centrala områden (Oates, 2006). Intervjuguiden skickades även ut i förväg till alla som skulle intervjuas i syfte att respondenterna skulle få en inblick i vilka ämnen intervjun skulle beröra. Intervjuguiden översattes också till engelska inför intervjuerna med respondenterna från Danmark (Appendix 3).

3.3.2 Frågeguide

Frågeguiden (Appendix 2) är baserad på intervjuguiden och har varit central i våra intervjuer. Vi har dock varit öppna för att lägga tyngd vid olika områden beroende på respondenten, ändra ordningen på frågorna samt ställa följdfrågor under intervjuens gång (Oates, 2006). Eftersom frågeguiden såväl som intervjuguiden är förankrade i litteraturen har det underlättat vårt arbete i att ställa empirin i kontext mot litteraturen. Även frågeguiden har översatts till engelska i samma syfte som nämns i intervjuguiden (Appendix 4).

Enligt Kvale (1996) är det viktigt med en pilotintervju för att intervjuaren ska få självförtroende i att intervjua. Under en pilotintervju tränas intervjuaren men även frågorna testas. Dessutom ger en pilotintervju möjlighet att testa den tekniska utrustningen (Kvale, 1996). Vi kunde efter vår pilotintervju dra slutsatsen att vissa frågor var för långa och komplicerade. Därmed valde vi att revidera våra frågor och skicka dessa till respondenten som deltog i pilotintervjun. Respondenten instämde i att frågorna blivit bättre formulerade och var redo att användas. Pilotintervjun inkluderades inte i studien utan användes enbart i syfte att kvalitetssäkra vårt intervjuemanus, frågeguiden och inspelningsutrustningen.

3.4 Urval

Kriterierna för vårt urval av respondenter var att de behövde besitta en god kunskap inom ämnet och att deras arbetsroller hade en tydlig koppling till forskningsfrågan. Vidare var målet att det skulle vara personer från olika organisationer och med olika bakgrund för att få flera perspektiv samt en nyanserad bild på forskningsfrågan. Fokuset har till stor del legat på få kontakt med individer från Folkhälsomyndigheten eftersom de ansvarar för att skydda Sveriges medborgare från smittsamma sjukdomar enligt 1 kap. 7 § smittskyddlagen (SFS 2004:168). Telia är den största mobiloperatören inom Norden (Ågren, Bjelkmar & Allison, 2021) och de som samarbetat med Folkhälsomyndigheten under COVID-19, därför var det även viktigt för oss att etablera kontakt med ansvariga personer därifrån. Målsättningen har varit att intervjua åtminstone en individ från Folkhälsomyndigheten, Telia och en oberoende person med god kunskap inom ämnet. Detta för att kunna utreda olika perspektiv och att täcka området så omfattande som möjligt.

För att komma i kontakt lämpliga kandidater har vi använt oss av vårt befintliga kontaktnät, vilket resulterade i ett snöbollsurval (Jacobsen, 2002), att kontakta bekanta som i sin tur har hänvisat oss vidare till andra personer. I vissa fall har även de personerna hänvisat vidare igen tills vi i slutändan har kommit i kontakt med de vi intervjua. Urvalsmetoden har underlättat processen att komma i kontakt med rätt personer på ett effektivare sätt då vi inte har varit tvungna att aktivt söka upp personerna. Jacobsen (2002) beskriver även bekvämlighetsurval vilket liknar snöbollsurval men syftar istället till att intervjua personer som är bekanta. Denna urvalsmetod har inte använts då vårt kontaktnät saknar personer med tillräcklig kompetens inom vårt forskningsområde.

I huvudsak har vi initialt i tillfrågningsprocessen kontaktat lämpliga kandidater via mail eller LinkedIn. Några har direkt sagt ja medan andra har hänvisat oss vidare eftersom en del av de initialt tillfrågade har ansett att de själva inte besitter tillräcklig kunskap om ämnet. I vissa fall har de tillfrågade personerna rekommenderat andra lämpliga personer då de inte har haft tid. Vid ett tillfälle kunde inte en individ delta på grund av företagspolicies då individen hade bytt arbetsplats nyligen.

I nedanstående tabell visas respondenternas bakgrundsinformation samt i vilket appendix respektive intervju återfinns.

Tabell 3.4 Presentation av respondenter

Respondent	Ålder	Kön	Yrkesroll	Företag	Intervju-längd (minuter)	Appendix
R1	27	Man	PhD student	Technical University of Denmark	43	Appendix 5
R2	30	Man	Postdoctoral Researcher	Cornell University	44	Appendix 6
R3	43	Kvinna	Head of Product Strategy och New Product	Telia Company	42	Appendix 7
R4	44	Man	Head of Product (fd. Head of Data Insights)	Telia Company	46	Appendix 8
R5	30	Man	Utredare	Folkhälso-myndig-heten	45	Appendix 9

3.5 Datainsamling

Samtliga intervjuer har utförts på distans med videomöten via Microsoft Teams. Då vi har befunnit oss i Lund under examensarbetet och varit begränsade i att resa har videointervjuer skapat möjligheten att nå ut till flera respondenter med den efterfrågade profilen som beskrivs i tidigare avsnitt. Tidigare publikationer, långt före videomöten var förekommande förklarar även fördelarna med internetbaserade intervjuer såsom att man kan nå ut till flera personer samt att intervjuerna kan bli billigare då resekostnader elimineras (Oates, 2006).

Genomgående har intervjuerna inletts med en kortare introduktion av oss, vilket universitet vi kommer från, syftet med intervjun, vad materialet ska användas till och vår forskningsfråga för studien. Detta är något som Jacobsen (2002) förklarar ska fungera som en uppvärmning inför intervjun för båda parterna. Innan vi började ställa intervjufrågorna till respondenten frågade vi om individen hade frågor gällande intervjuguiden som hade skickats ut i förväg. Denna typ av öppenhet är något som Jacobsen (2002) beskriver som nödvändig. Vi informerade sedan respondenterna om att de gärna fick applicera sitt eget perspektiv och avslutade varje intervju med att fråga om deltagarna ville tillägga information för att

respondenterna skulle ha möjlighet att ta upp information de själva anser vara relevant i enlighet med vad Oates (2006) nämner om semistrukturerade intervjuer.

Datansamlingen har utförts genom att spela in intervjuerna och efteråt har ljudfilerna transkriberats manuellt. Ljudinspelningen har påbörjats efter att varje respondent godkänt det i samband med att de etiska förhållningssätten för intervjun har gått igenom. Bryman och Bell (2017) förklarar vikten av att spela in en intervju för att intervjuaren ska kunna fokusera på respondentens svar istället för att anteckna vad som sägs. Detta är något som Jacobsen (2002) bekräftar genom att belysa vikten av att ett bra samtal ofta kräver ögonkontakt vilket blir svårt om intervjuaren behöver anteckna. Anteckningarna riskerar även att inte återge ordagranna citat (Jacobsen, 2002). I efterhand har även transkriberingen av intervjun skickats ut till respondenten. Syftet med detta är respondenten ska få en möjlighet att bekräfta, lägga till, ändra eller ta bort delar av det transkriberade materialet (Oates, 2006).

3.6 Analys

Oates (2006) belyser vikten av att transkribera inspelade ljudfiler då det är lättare att analysera datan när den finns nedskrivet. Författaren förklarar även att det är först när intervjun kommer till liv igen vilket ger intervjuaren en chans att analysera och reflektera över datan (Oates, 2006). När en intervju är utförd och transkriberad efterföljer analys av datan. Då denna studie har baserats på en kvalitativ metod handlar analysen om att sammanställa intervjuunderlaget och utforska olika teman i intervjuerna (Oates, 2006).

Deduktiv analys baseras på existerande teorier som återfinns i litteraturen eller som en har utvecklat själv (Oates, 2006). Oates (2006) belyser också vikten av att inte förhålla sig till enbart en etablerad teori då man lätt kan bortse från andra teman i datan. I vår analys har vi utgått ifrån ett deduktivt tillvägagångssätt då målsättningen var att definiera teman baserat på litteraturgenomgången.

Inledningsvis läste vi igenom alla transkriberade intervjuer för att få en översiktlig bild av de olika teman som hade tagits upp. Därefter definierade vi teman som kunde appliceras på samtliga intervjuer. I nästa steg läste vi igenom transkriberingen igen och färgkodade de olika teman som behandlades i varje intervju. Efter att vi enskilt hade färgkodat allt intervjumaterial diskuterade vi oklarheter och lade även till teman som inte hade definierats tidigare. Slutligen blev de olika temana grunden för rubrikerna i empirin. Under analysarbetets gång har vissa fördefinierade teman slagits samman och andra har delats upp i underkategorier för att uppnå en tydligare datasammanställning.

3.7 Etiska överväganden

Bryman och Bell (2017) beskriver ett antal krav för att forskning ska vara etisk. Först förklarar författarna informationskravet som bygger på att den som intervjuas informeras om syftet. Sedan nämner de samtyckeskravet vilket innebär att den intervjuade ska vara medveten om att samtycket är frivilligt samt att det går att avbryta när som helst. Efter det beskriver författarna konfidentialitets- och anonymitetskravet vilket betyder att uppgifterna om de intervjuade ska behandlas konfidentiellt så att obehöriga inte får tillgång. Vidare nämns nyttjandekravet vilket beskrivs som att de insamlade uppgifterna bara får användas för

forskningsändamålet. Till sist nämns falska förespeglingar vilket betyder att den som intervjuar inte får ge någon falsk eller vilseledande information om intervjun. Även Oates (2006) beskriver en intervjudeltagares rättigheter vilka påminner mycket om de Bryman och Bell (2017) nämner. Oates (2006) förslag är som följer: rätten till att inte delta, rätten att ångra sig, rätten att ge ett informerat samtycke, rätten att vara anonym och rätten till konfidentialitet. Oates (2006) beskriver även skyldigheter man har som etisk forskare: inga onödiga inkräktningar, uppträda med integritet, följa professionella uppförandekoder, inget plagiat samt vara en etisk recensent. Dessa skyldigheter har även varit viktiga för oss att ha i åtanke under intervjuerna för att behandla de intervjuade på ett korrekt sätt.

För att säkerställa att intervjuerna genomförs på ett etiskt sätt har vi följt Bryman och Bell (2017) samt Oates (2006) krav för etisk forskning. Alla intervjudeltagare har fått information om vilka vi är och vad syftet med vår studie är. Vi har från början gjort intervjudeltagarna medvetna om deras rättigheter inför varje intervju samt fått ett informerat samtycke av varje intervjudeltagare. De intervjuade har också gjorts medvetna om att de när som helst kan avbryta intervjun i enlighet med Bryman och Bell (2017) samtyckeskrav och Oates (2006) rätten att ångra sig. Vi har även bett om samtycke för att spela in intervjuerna i syfte att underlätta vårt arbete för att bearbeta intervjuerna i efterhand. De intervjuades namn utelämnas då det är irrelevant för forskningen, däremot nämns mer generell information som respondenterna gett oss medgivande till. Viktigt att belysa är att den generella informationen inte ger möjlighet till att enkelt identifiera de intervjuade. Transkriberingen har överlämnats till de intervjuade då det är viktigt att säkerställa att båda parter har samma bild av intervjun och vad som sagts (Oates, 2006). Möjligheten att ändra eller ta bort delar har således givits till samtliga respondenter. Personuppgifterna har efter intervjun bevarats oåtkomligt för obehöriga i enlighet med Bryman och Bell (2017) konfidentialitets- och anonymitetskravet samt Oates (2006) rätten till konfidentialitet. Genom att följa dessa principer och tillvägagångssätt är meningen att datainsamlingen ska ha skett på ett etiskt vis och att de intervjuade ska känna sig trygga med sitt deltagande.

3.8 Validitet och reliabilitet

För att bedöma kvaliteten på den insamlade datan är det viktigt att kritiskt granska dess giltighet (validitet) och dess tillförlitlighet (reliabilitet), (Jacobsen, 2002). Både Jacobsen (2002) och Bryman och Bell (2017) nämner att validitet och reliabilitet är ett omdiskuterat ämne gällande kvalitativa studier. De menar att begreppen ursprungligen använts för kvantitativa studier och att många forskare ställer sig skeptiska till att applicera dem på kvalitativa studier. Jacobsen (2002) beskriver dock att användningen av begreppen validitet och reliabilitet för en kvalitativ studie inte handlar om att underkasta datan en kvantitativ logik utan snarare om att kritiskt granska den data som samlats in under en kvalitativ undersökning. För att anpassa kriterierna för validitet och reliabilitet till en bedömning av kvalitativ forskning skiljer Bryman och Bell (2017) på begreppen extern reliabilitet och intern reliabilitet samt intern validitet och extern validitet.

- Extern reliabilitet syftar till i vilken mån en undersökning går att upprepa. Enligt Bryman och Bell (2017) är detta relativt svårt att uppnå med en kvalitativ undersökning då det är svårt att "frysa" en social miljö. För att uppfylla kravet i den mån det går har vi inkluderat hur urvalet gått till, bakgrunden till urvalet, bakgrunden om respondenterna, hur kontakten med respondenterna gått till, hur intervjuerna genomförts,

vilka frågor som ställts samt inkluderat transkriberingarna i sin helhet i appendixet.

- Intern reliabilitet betyder att vid tillfällena där det finns fler än en forskare bör alla inom en forskningsgrupp vara överens om hur man ska tolka allt som syns och hörs (Bryman & Bell, 2017). För att säkerställa detta har båda forskarna deltagit vid alla intervjutillfällen. Dessutom har analysen skett i nära samarbete och samförstånd för att undvika att uttalanden har tolkats olika.
- Intern validitet avser att det finns en tydlig koppling mellan forskarnas observationer och de teoretiska idéerna de utvecklar, det vill säga att resultat och verklighet stämmer överens (Bryman & Bell, 2017). Intern validitet är vanligtvis en styrka i kvalitativa studier enligt Bryman och Bell (2017) eftersom man som forskare spenderar mycket tid i den sociala kontext som undersöks. För att säkerställa kravet om intern validitet är studien baserad på aktuell och relevant litteratur för forskningsområdet. Även den kvalitativa intervjuemetoden minskar risken för onyanserade bedömningar. Dessutom har vi baserat intervjufrågorna på litteraturen för att på så sätt kunna konfrontera respondenterna med de observationer som gjorts i litteraturen (Jacobsen, 2002).
- Extern validitet beskrivs som till vilken grad man kan generalisera forskningen, det vill säga hur enkelt det är att applicera studien på andra miljöer eller sammanhang (Bryman & Bell, 2017). Enligt Bryman och Bell (2017) är extern validitet ett problem vid kvalitativa studier på grund av hur studien är utformad. För att få en bredare bild av forskningsfrågan har personer från olika företag intervjuats men också oberoende forskare på ämnet. Trots det är urvalet begränsat då bara fem personer från två olika länder intervjuats. Svaren hade kunnat se annorlunda ut i en annan kontext och den externa validiteten kan ses som en brist i studien.

4 Empiriskt resultat

I empiriavsnittet presenteras resultaten vilka grundar sig i uttalanden från de intervjuer som genomförts. Inledningsvis presenteras upplägget av empirin. Därefter presenteras påståenden som är relevanta för att undersöka det som presenterats i litteraturgenomgången kopplat till vår forskningsfråga. Transkriberingen av intervjuerna återfinns i appendix 5-9.

4.1 Presentation av empiriskt resultat

Det empiriska resultatet presenteras utifrån tre teman som har delats in i huvudrubrikerna: användning och nytta med mobildata, utmaningar och risker samt möjligheter och framtid. Temana har genom deduktiv analys tagits fram efter en analys av intervjuguiden och vilka fokusområden som fanns. Forskningsfrågan har också tagits i beaktande vid identifiering av temana. Varje rubrik har även delats in i en del underrubriker som utgår ifrån olika teman. En presentation av informanternas bakgrundsinformation återfinns i tabell 3.4.

4.2 Användning och nytta med mobildata

4.2.1 Användningen av mobildata

Respondenterna från Telia (R3 och R4) var båda eniga om att transformera data till insikt i korthet innebär att ställa de rätta frågorna till en datakälla, de menar att en datakälla ofta har många svar men den kritiska biten är att definiera frågorna. När man har tagit reda på frågan och hur den ska ställas, då kan man få en insikt som är "actionable" menar R4. R5 förklarar att samtliga dataanalyser kräver att rådatan slås ut för att göra den begriplig.

"Data i sig svarar ju på ingenting, det är ju en möjliggörare så att gå från data till insikt så krävs det ju såklart en relevant datamängd men det krävs också en djup förståelse för problemställningen och sen inför det arbete som krävs för att transformera datan kanske sammanfoga med annan typ av data, visualisering, paketering av detta, förklaring av detta och i vårt fall även prissättning" - R4 (Appendix 8, 50-55)

R2 var en del av en grupp för expertmodellering av COVID-19 vid Technical University of Denmark och förklarar att de inte hade tillgång till rådatan från Danmarks telekomföretag under tiden han arbetade med samarbetet. Respondent 2 förklarar att telefonoperatörerna försåg gruppen med mobilitetsmatriser som baserades på aggregerad mobildata. Vidare berättade R2 att de genom matriserna kunde se hur många mobiltelefoner som hade rört sig från ett givet område till ett annat vid ett specifikt tidsintervall. I början av samarbetet vill han minnas att gruppen fick data från tre tidsintervall under en dag och det var så mobildatan samlades in. Respondent 2 beskriver även hur möten anordnades med telefonoperatörerna för att diskutera vilken mobildata som kunde användas för epidemiologisk modellering och hur datan skulle kunna användas på ett integritetsbevarande sätt. R5 förklarar även att Folkhälsomyndigheten inte har några stora databaser utan att de tar del av en produkt som har tagits fram av Telia som ett exempel.

“...vi har inte suttit med den absoluta rådatan där vi kan följa individer. Allting är ju anonymiserat och aggregerat på olika nivåer som vi har avtalat. Så i viss mån kanske det är att vi arbetar i gränslandet i vad som har varit big data men vi har inte några stora databaser som vi har anonymiserat.” - R5 (Appendix 9, 81-85)

R1 förklarar att mobildatan har använts för att förstå rörelsetrender genom att ta del av aggregerad och anonymiserad data från Danmarks telefonoperatörer. Vidare har man kunnat följa vilka master en mobiltelefon har varit närmast och vid vilken tidpunkt. Med den informationen kan en sammanställning göras av när och hur en individ har rört sig. R1 förtydligar dock att de aldrig har haft tillgång till individbaserad data och att syftet inte har varit att följa rörelser för enskilda individer. R1 förklarar att den sammanställda och anonyma mobildatan användes för att modellera hur populationer rörde sig.

R3 berättar att det inte är jätteintressant för världen att veta hur enbart Telias abonnenter rör sig. Vidare förklarar R3 att ett viktigt arbete för Telia har varit att extrapolera deras mobildata för att den ska kunna motsvara en anonymiserad population. R3 förklarar även att framtagningen av extrapoleringsmetoden har tagit lång tid att utveckla för att i det här fallet kunna representera Sveriges befolkning.

R4 förklarar även hur teknikutvecklingen beträffande big data har varit helt avgörande för att kunna hantera Telias mobildata. Tidigare krävdes tung specialistkompetens för att sköta servrar och för att hålla infrastrukturen i liv. Vidare förklarar R4 att datan är mer tillgänglig i form av molntjänster från företag som Google, Amazon och Microsoft utför det komplicerade uppgifterna kring hanteringen och ser till att man har rätt processorkraft och tillräckligt många servrar med tillgänglig lagring. Telia använder Google Cloud för att hantera stora mängder mobildata och kan ställa frågor till datan som om det vore ett litet dataset. R4 förklarar att man enbart behöver kunna Structured Query language (SQL) för att kunna bearbeta mobildatan. Vidare berättar R4 att Telia tar in flera miljarder datapunkter per dag genom att använda molntjänsten.

4.2.2 *Nyttan med mobildata*

På frågan om de intervjuade anser att mobildata har använts för att bromsa spridningen av COVID-19 (Appendix 2 och 3, fråga 8) gav respondenterna nedanstående svar:

“Well, in Denmark I would say yes.”- R1 (Appendix 5, 79)

R1 utvecklar sitt svar genom att förklara att mobildatan som telekomföretagen försåg forskningsgruppen med användes för att skapa en hemsida över Danmarks befolknings rörelsemönster. Vidare förklarar R1 att hemsidan visade den anonymiserade rörelsedatan på ett enkelt sätt för befolkningen och politikerna i landet. R1 belyser att gruppen hade som mål att presentera datan på ett lättillgängligt sätt. Webbplatsen innehåller interaktiva visualiseringar och R1 förklarar att man exempelvis kan få en överblick över hur populationen har rört sig mellan specifika kommuner. R1 berättar även att den danska regeringen har tagit del av rapporter som har baserats på gruppens hemsida. Vidare bekräftar R1 att de danska regeringen har använt sig av mobildatan men nämner att han inte vet specifikt hur den användes men är ganska säker på att rörelsetrender har legat till grund för deras policies.

R1 förklarar även att danska telefonoperatörer försåg forskare med aggregerad och anonymiserad mobildata för att forskarna skulle kunna analysera och använda informationen i infektionsspridningsmodeller. R1 berättar att deras metoder utgick ifrån maskininlärning för att prediktera och förstå datan djupare. Han har även publicerat en artikel som handlar om att dela upp rörelsetrender i veckodags-, helg- och helgdagsmönster. R1 förklarar att man kunde se hur folk ändrat sina vanor från veckodagsmönster till ett helgmönster i början av nedstängningen vilket han anser vara rimligt då befolkningen stannade hemma.

“Do I think so, I think in the beginning of the pandemic there was a lot of hope that one could do this. We certainly had a lot of hope.” - R2 (Appendix 6, 121-122)

R2 förklarar att det i början av pandemin fanns mycket hopp gällande vad man skulle kunna uppnå med mobildata. Vidare utvecklar R2 att man i det initiala arbetet tänkte på mobildatans möjligheter utan att ta hänsyn till integritet. Gruppen tänkte att man eventuellt skulle kunna följa upp vilka individer en person har varit i kontakt med. Vidare förklarar R2 att forskningsgruppen trodde att den sammantagna datan skulle vara användbar i modelleringen av en pandemi. R2 nämner slutligen att detta inte är möjligt om man inte har tillgång till GPS-koordinater, med mobiloperatörsdata baserat på signalmaster går det inte att utläsa vilka personer en individ har varit i kontakt med. Dock kom forskningsgruppen fram till att mobildatan kunde representera en komplett bild av hur grupper rör sig men att uppföljning av individuella rörelsemönster inte gick att utföra utan att inskränka på integriteten. Därför tror inte R2 att mobildata i sig självt har minskat smittspridningen.

R2 belyser styrkan i mobiloperatörsdatans fullständighet och nyttan av att kunna modellera rörelsemönster för en population. Vidare förklarar R2 att liknande datakällor från Facebook går att använda för modellering av rörelse på en befolkning men att fullständigheten är bristande i jämförelse med mobildata från telekomföretag då nästan alla har en mobiltelefon till skillnad från Facebookkonton. Danmark använde mobiloperatörsdata för att följa resetrender hos den danska befolkningen och genom detta kunde man avläsa att delar av populationen rörde sig till sina sommarhus under rådande nedstängning. R2 förklarar att mobildatan gav en analys av hur restriktioner efterföljdes men att beslut inte baserades på resultatet då smittspridningen inte påverkades märkbart av resandet.

“Ja, det tycker jag definitivt att det har gjort.” - R3 (Appendix 7, 87)

R3 förklarar att den stora frågeställningen som Telia hade var hur Sveriges population betett sig under olika restriktioner och rekommendationer. Vidare förklarar R3 att Telias mobildata har kunnat besvara om Folkhälsomyndigheten har behövt införa hårdare restriktioner för populationens rörelse inte förändrades av tidigare satta restriktioner.

R3 förklarar att Telia arbetar med retroaktiv data och därför har den använts som uppföljningsverktyg under COVID-19 för att följa upp hur populationer förflyttar sig. Ett exempel som R3 förklarar är att Sverige under påsk 2020 hade en restriktion att man inte skulle resa mer än två timmar från hemmet. Vidare förklarar R3 att Påsk är en stor rörelsehög där många reser längre sträckor än två timmar och därför var denna analys av mobildatan relevant. Resultaten Folkhälsomyndigheten kunde påvisa med mobildatan var att Sveriges befolkning anpassade sig till restriktionerna. R3 nämner att mobildatan även användes i Folkhälsomyndighetens algoritmer för att prediktera olika smittspridningsscenarion som man exempelvis ser på TV.

“Ja, det skulle jag säga i allra högsta grad och mycket på grund av samarbetet med Folkhälsomyndigheten i Sverige.” - R4 (Appendix 8, 134-135)

Respondent 4 förklarar att Telias mobildata har varit en del av beslutsunderlaget för vilka restriktioner och rekommendationer som behövde vidtas. Dels har datan använts i Sverige genom samarbetet med Folkhälsomyndigheten men även i Finland har datan använts på ett liknande sätt förklarar R4.

R4 förklarar även att det finns en hel uppsjö med forskningsartiklar om hur mobildata har kunnat användas. Dock förtydligar han att det inte finns speciellt mycket konkreta scenarion där mobildata faktiskt har använts under COVID-19. Därför anser R4 att Telia och Folkhälsomyndighetens samarbete skiljer sig från mängden. Telia i sig har inte styrt beslut utan bara försett Folkhälsomyndigheten med data som sedan har använts för att fatta beslut.

R4 berättar även att Telia har varit en del av ett samarbete med EU-kommissionen där deras anonyma mobildata med flera andra telekomoperatörernas data användes för att publicera en artikel rörande hur man kan omforma mobilitetsdata för att ge svar på frågan hur man bör forma åtgärder för att begränsa smittspridning.

“Ehmm, ja det kan vi absolut säga att det har varit en pusselbit i att utforma vårt arbete.” - R5 (Appendix 9, 104)

R5 berättar att rörelsedatan har utgjort en pusselbit för att bromsa spridningen av COVID-19. Dock belyser han att det är väldigt indirekt hur mobildata har bidragit till minskad spridning men att man med hjälp av datan har fått en förståelse för hur människor kan anpassa sig till restriktioner och rekommendationer. R5 förklarar att översikterna av rörelsedata som Folkhälsomyndigheten har publicerats för att informera allmänheten.

R5 förklarar att den viktigaste aspekten med mobildata för att kunna få insikter som är generaliserbara på en befolkning är att datan måste vara representativ. Vidare förklarar R5 att när Folkhälsomyndigheten har jämfört resor över kommungränser kontra regiongränser har syftet delvis varit att följa upp de restriktioner som syftade till att minska längre resor.

R5 förklarar även vikten av att jämföra datakällor för att kunna dra slutsatser gällande resemönster. Telias data har till exempel jämförts med Googles och Apples mobildata. Folkhälsomyndigheten har använt mobildatan genom att se följsamheten bland befolkningen gällande de åtgärder eller rekommendationer som har kommit från myndigheter förklarar R5. Respondent 5 återkommer till vikten av representativ data, hur generaliserbara fynden är och vilka fynd som kan kommuniceras väl i arbetet med att använda mobildata.

Folkhälsomyndigheten har även använt mobildata för att följa upp resande överlag genom att analysera i stort sett varje resa som är större än rörelsen från en telemast till en annan, förklarar R5. Folkhälsomyndigheten definierar det som totalt resemått. R5 förklarar att det finns svagheter då telemasterna inte är jämnt fördelade men att det initialt handlar om att få en god bild av datan man har tillgång till och därefter bedöma om det går att koppla till faktorer som man generellt ser i smittspridning vilket ofta är komplext.

Mätandet av följsamhet genom mobildata är något vi tidigare inte har haft någon baslinje för, förklarar R5. Mått som självrapporterad data för hur en individ har efterlevt rekommendationer är något som har använts innan mobildata vilket kan betraktas som partisk

data och därför inte går att basera en baslinje på. Därför menar R5 att en stor del av arbetet med mobildata har varit att identifiera en rimlig baslinje, exempelvis vilka förändringar vi kan förvänta oss i folks rörelse till följd av årstider.

4.2.3 Samarbetet mellan Telia och Folkhälsomyndigheten

R3 säger att det initiala målet för samarbetet mellan Telia och Folkhälsomyndigheten var att använda mobildata i syfte att hjälpa till att bromsa spridningen av COVID-19. R4 förklarar också att de enkla målet var att hjälpa till vilket Telia trodde sig kunna göra eftersom deras data redan vara produktifierad, i tjänst samt att allting kring etiska vägval och GDPR redan var utrett. Vidare säger han att Telia ville göra något bra för samhället och bidra till att mildra effekterna av pandemin. Både R3 och R4 berättar att det var många som kontaktade Folkhälsomyndigheten under pandemin men att anledningen till att de så smidigt kunde inleda ett samarbete var eftersom deras tjänst redan var färdig och hade visat sig fungera för liknande syften tidigare.

R3 menar att Folkhälsomyndigheten hade problem att kvantifiera sin data och se vad effekterna av rekommendationerna var. Telia ville bidra med en datadriven debatt och diskussion som baserades på rörelsedata istället för påståenden eller tyckanden.

“Så det var ju mycket reaktioner i media Stefan Löfven har sagt att vi inte ska träffas i grupp och jag såg att det var fullt på parkeringen på Gekås. Alltså det fanns ju massor med olika rapporter om olika saker där folk visst samlades fortfarande och så kan det ju vara men det är ju inte kvantitativa insikter utan det är ju väldigt här och nu vad jag kan se med mina två ögon och det här, potentialen i vår data är ju just den att du får en kvantitativ analys på en befolkning snarare än individer.” - R3 (Appendix 7, 161-167)

Vidare förklarar R3 att Folkhälsomyndigheten snabbt såg potentialen i Telias data men att de tillsammans fick arbeta för att paketera deras tjänst för att passa just Folkhälsomyndighetens frågeställning. Vidare säger R3 att för att skapa insikter måste man svara på de rätta frågorna. R4 menar att Folkhälsomyndigheten inte var vana att arbeta med denna typ av data så de initiala arbetet handlade mycket om hur de skulle kunna använda datan som en del i sina beslutsprocesser. R3 berättar vidare att det handlade mycket om att kunna paketera Telias tjänst så att Folkhälsomyndigheten kunde få svar på frågorna gällande rekommendationer och folkmassors rörelse, det vill säga grupper av människors rörelse. R4 förtydligar att han inte kan säga att just bidraget med mobildata till media ledde till konkret minskad smittspridning men att en datadriven dialog ofta bidrar till det.

“Det var ju ett otroligt intensivt samarbete och otroligt häftigt och hela teamet var ju engagerade och man kände ett purpose på ett sätt som jag tror att vi aldrig har gjort eller kanske aldrig igen kommer känna.” - R4 (Appendix 8, 217-219)

På frågan gällande om målen kring samarbetet förändrades över tid svarar R4 att målet att hjälpa till aldrig förändrades. Däremot kan man säga att målet utvecklades efter att Telia hade tagit fram många olika dashboards, visualiseringar och frågeställningar och insåg att det var en relevant simulering av data. Det blev sedan en av deras produkter som standardiserades och nu kallas Standard Dashboard eller Generic Dashboard. Han berättar vidare att det nu är ett kommersiellt upplägg till skillnad från i början av samarbetet när Telia arbetade pro bono med Folkhälsomyndigheten. Nu är de istället en kund och en stor skillnad blir att de kan ställa krav på service levels och support samt få en längre garanti. Han avslutar med att säga:

“Så målet blev väl om möjligt, så tidförvandlades vårt mål till att det är en viktig kund för oss och nu när pandemin förhoppningsvis inte kommer tillbaka, hoppas vi att det här datat fortsatt är viktigt för dem...” - R4 (Appendix 8, 240-242)

R5 som arbetar på Folkhälsomyndigheten förklarar deras syn på samarbetet.

“Jo men syftet eller arbetet kom ju till för att Telia som ni säkert redan har läst i Pär och Telias gemensamma artikel var ju att de hade en datakälla som vi såg kunde vara nyttig och öka vår förståelse av hur befolkningen rörde sig och hur de tog till sig diverse rekommendationer för att kunna följa det. Och eftersom de hade den datakällan var vi intresserade av att se hur vi kunde nyttja det. Och det är ju något vi har nyttjat i kombination med annan rörelsedata men även data som på olika sätt följt efterlevnad av rekommendationer.” - R5 (Appendix 9, 132-137)

Vidare berättar han att målen med att använda datan från Folkhälsomyndighetens sida var att förstå följsamhet, rekommendationer och befolkningens rörelser allt igenom. Han nämner också att samarbetet fortgår även om den värsta delen av pandemin förhoppningsvis är över om än med lite mer otydliga mål.

Respondenterna från Danmark har ingen insyn i samarbetet mellan Telia och Folkhälsomyndigheten men drar däremot paralleller till hur de arbetat i Danmark. R1 förklarar att alla telekomföretag i Danmark gick med på att dela sin data gratis med den danska regeringen så att forskare kunde använda datan för att försöka modellera smittspridningen. Efter ett tag började dock vissa företag kräva betalning vilket ledde till att forskarna inte längre fick någon data. Ingen av dem kan däremot svara på vad målen med samarbetet i Danmark var men R2 säger:

“I know that I made a model that could take in the cell phone data and implement it into a model. I did communicate to the people who were actually making the models how it should be done and I hope that was used eventually. But I know that the mobility pattern was tracked often, how many people were moving around. Although the exact goals I do not know about.” - R2 (Appendix 6, 199-202)

4.2.4 Alternativa metoder

I Danmark användes en smittspårningsapp under pandemin. R1 berättar att det var en anonym app som använde Bluetoothsignaler och kände av vilka man varit i nära kontakt med. Vidare förklarar R1 att om man blev sjuk skrev man in det i appen och då fick alla som man varit i närheten av en notis att även de behövde testa sig. Han påpekar dock att den var väldigt bra i början när inte så många hade blivit smittade av COVID-19 men senare blev det ett problem när för många var sjuka.

“wherever you walk it would just say “oh you have been close to one”, “oh you have been close to one”, “oh you have been close to one”. Then basically every time you had to go to the centre to get a covid-test and the when you got home the next day you would get a new notification that would say “oh you have been close to one” because in the queue to get the test you were always close to people who got one.” - R1(Appendix 5, 115-119)

R3 nämner att det var mycket diskussion på marknaden gällande smittspårningsappar men samarbetet mellan Telia och Folkhälsomyndigheten fokuserade enbart på Telias data. R4 säger också att många företag ville utveckla en smittspårningsapp med Telia men att det inte var något de gick vidare med. R3 förklarar vidare att smittspårningsappar handlar om att datan är samtyckesbaserad och att man tittar på varje individ vilket inte är något Telia Crowd Insights arbetar med. Det är det också svårare att få ett representativt dataset när man behöver samtycke från varje individ.

R3 och R4 berättar att Telia arbetat med Ericsson, Chalmers och Sahlgrenska där man använde samma typ av data för att förutspå sjukhusinläggningar. Ericsson byggde AI-modeller på Telias data samt annan data för att kunna förutspå vårdbehovet under pandemin. Sahlgrenska använde denna modell i sin planering kring resurser, hur många bäddar som behövdes, hur många intagna det skulle vara och det i sin tur ledde till effektiv vårdplanering.

4.3 Utmaningar och risker

4.3.1 Tekniska utmaningar

På frågan om vilken som är den största utmaningen i arbetet med mobildata under COVID-19 svarar R1 att det är en teknisk utmaning. Han förklarar att mobildatan är representativ och att han i sitt arbete fått förarbetad data av telekomföretagen. Problemet i Danmark var att de fick data av många olika företag och företagen förarbetade sin data på olika sätt. Han ger ett exempel på när folk gör längre resor stannar man ofta för att kissa, vilket vissa företags data såg som en slutdestination medan andra förstod att det var ett stopp för en längre resa och därför inte inkluderade stoppet. Vidare förklarar han att det var svårt att kombinera dessa dataset när de såg så olika ut och att han hade önskat en standardisering av hur man förarbetar dataset.

Ett annat problem R1 nämner är att datan inte är så specifik, den visar bara vilken som är den närmaste telemasten till en mobiltelefon vilket kan vara flera kilometer ifrån. Vidare säger han att beroende på vad du vill använda mobildatan till är det viktigt att vara medveten om denna felaktighet - att datan inte är så granulär. Även R3 förklarar att man inte kan vara så granulär när man arbetar med anonym data vilket gör att man inte kan vara exakt i sina insikter.

“vi kan ju titta på beteenden det ser ut som folk åker till jobbet i den här resan, de har ett beteende som vi kan anta är att man pendlar till en arbetsplats, men vi vet ju inte vad de gör vid sin destination. Det är en anonym analys, vi har ingen aning om man är ute och shoppar eller om pendlar...” - R3 (Appendix 7, 210-213)

Hon förklarar vidare att dessa analyser handlar om att anta och förlita sig på statistik vilket oftast stämmer men det finns en viss osäkerhet. Hon exemplifierar genom att berätta om en “jobba hemma analys” Telia gjorde. Där kunde man tydligt se *“hur rörelsemönstrena ändrades från att ha ett jobba hemma beteende till att börja jobba igen och vi såg en väldigt tydlig korrelation med rekommendationer och större bolag som öppnade upp igen efter pandemin”* R3 (Appendix 7, 217-219). Vidare berättar R3 att de har ett väldigt stort dataset och följer trender på hur folk rör sig vilket gör att insikterna blir tydliga och risken för fel väldigt låg.

Även R5 menar att det är en ny metod och att det finns en stor variation i hur människor resor och betar sig under året vilket gör att man inte ska låta de tolkningar man gör gå för långt - varken för defensivt eller för offensivt. För att minimera risken har Folkhälsomyndigheten ofta jämfört olika källor och bedömt olika källor som olika tillförlitliga vilket gjort att man inte alltid kan dra slutsatser, *“men som sagt måste man vara försiktig med sina slutsatser och vilket håll man anser att det lutar emot och både väga epidemiologiska principer kring smittspridning gentemot det man ser i datan.”* - R5 (Appendix 9, 167-169).

4.3.2 Utmaningar med datakvalitet

På frågan om hur man presenterar data av så hög kvalitet som möjligt svarar både R1 och R2 att de inte riktigt vet eftersom de fått datan förarbetad av telekomföretagen. R1 och R2 fick heller ingen insikt i den processen på grund av integritetsskäl.

R4 från Telia inleder med att berätta att ha så hög kvalitet som möjligt på deras data är kritiskt och därför arbetar de mycket med det. Bland annat har de olika samarbeten exempelvis med Statistiska centralbyrån för att producera bra och jämförbar statistik. De har ansvariga personer men även en organisationsförmåga kring det. Samtidigt nämner R4 att inga datakällor är perfekta men att de alltid försöker vara öppna gentemot sina kunder och kontinuerligt arbetar med transparens och att förbättra kvaliteten. R3 berättar mer specifikt hur de arbetat med datakvalitet innan och under pandemin men poängterar också hur viktigt öppenheten är för Telia. Innan pandemin levererade de oftast månatliga rapporter över rörelsedata och då hade de en dataanalytiker som manuellt satt och kvalitetssäkrade all data innan den skickades iväg. Detta blev mycket svårare under pandemin när deras dashboards skulle uppdateras dagligen samtidigt som hundratals kunder ville följa deras insikter.

“Vi fick ju skala upp vår datatjänst, det vill säga vi var ju tvungna att bygga automatik kring datakvalitet och även då vi byggde en dashboard där vi kunde ratea dagarna där vi kunde säga om det var bra kvalitet för det datasetet vi har det är ju inte till för att följa mobiler utan det är ju till för att upprätthålla ett nätverk för att vi ska kunna ringa.” - R3 (Appendix 7, 237-240)

Vidare förklarar R3 att när uppdateringar och liknande görs i nätverket blir deras data sämre. Efter ett tag insåg man dock hur viktigt deras arbete var för samhället och de fick själva avstyra uppdateringar under “viktiga” dagar såsom påskhelgen eller midsommar. Hon poängterar dock att datan kanske bara är sämre under en dag och Folkhälsomyndigheten tar inte beslut på en dags data utan baserar beslut på trender som snarare sträcker sig veckovis vilket gör att felmarginalen inte blir så stor vid en dålig dag. Sammanfattningsvis påpekar hon hur viktigt det är med transparens mot kunden eftersom data aldrig är 100 % och att de därför bedömt hur bra kvalitet det är på datan i sina dashboards för kunden att se.

R5 från Folkhälsomyndigheten menar att en viktig del i deras arbete med kvalitetssäkrad data är att jämföra olika datakällor och se att fynden överensstämmer. Han menar även att en svårighet är att datan kommer retroaktivt och att det därför blir svårt att se om en rekommendation följs direkt eller inte. En annan viktig del han nämner är att inte dra förenklade slutsatser av datan man får.

4.3.3 Etiska utmaningar

“Det är lätt för någon att säga att vi har anonym data. Okej, på vilket sätt. Har ni testat att se om det går att återidentifiera någon. För om det går att återidentifiera någon, då är det inte anonym data.” - R4 (Appendix 8, 311-313)

R1, R2, R3 och R4 är eniga om att en viktig del inom arbetet med mobildata är att företagen eller forskningsgrupperna är transparenta i hur datan används och i vilket syfte. R1 förklarar att en fördel med forskning i jämförelse med privata företag är att forskare ofta publicerar sitt arbete för allmänheten vilket skapar full insyn i forskningen. Vidare förtydligar R2 att det även är viktigt med en öppenhet i hur och var datan lagras. R2 förklarar att han personligen fick ett större förtroende för myndigheter efter att ha varit en del av arbetet med rörelsemönster baserat på mobildata.

“ So I was very impressed how serious the authorities took the privacy aspect. I remember in Denmark that separate systems were set up to handle sensitive data and in the beginning we had to use really involved ways of accessing data since it had to be very secure but they actually managed to set up a separate and nice system in the end which I think worked.” - R2 (Appendix 6, 235-238)

R2 förklarar att det är viktigt i arbetet med mobildata att följa upp att personlig integritet inte kränks. Vidare förklarar han att integriteten leder till begränsningar i vad man kan göra med datan men förtydligar att det mest relevanta var rörelsedatans fullständighet. R2 utvecklar även att datan de arbetade med var väldigt aggregerad och därmed inte integritetskränkande.

R1 förklarar utmaningen för den danska regeringen att fatta beslut under pandemin. Ofta uttalade sig experter hur regeringen skulle agera medan den faktiska regeringen gjorde något annat vilket han tror skapade en misstro hos befolkningen. En annan utmaning som R1 beskriver är att bibehålla mobildatans integritet. Trots att Danmark inte spårade resdata om mindre än 10 personer förflyttade sig förklarar R1 att det skulle vara lättare att identifiera vilka som förflyttade sig på en plats där det bor väldigt få människor. R2 förklarar dock att trösklarna för att inte spåra rörelsedata om ett mindre antal personer förflyttade sig var en styrka i att göra datan anonym. Vidare utvecklar R2 att sättet som hälsomyndigheter hanterade mobildatan var riskfritt enligt honom. I länder som har använt mobildata baserat på GPS-kordinater tror R4 att personlig integritet har varit en större utmaning. R1 förklarar även att det förmodligen går att återidentifiera data så länge det finns tillräckligt mycket information som kan kopplas till den datakällan. Vidare förtydligar R1 att återidentifiering av anonym data är komplex och att han inte är orolig för att sin personliga mobildata.

R3 hävdar att dataforskare inte skulle påstå att det existerar anonym data då data aldrig kan vara helt anonym. Dock är R4 ändå säker på att Telias dataset är helt anonymt. R3 och R4 förklarar att Telias data kan klassificeras som anonym och aggregerad genom att den är legalsäkrad samt uppfyller GDPR-lagstiftningen. Vidare har Telia tagit extern hjälp i syfte att försöka återidentifiera den anonyma mobildatan vilket har misslyckats. R4 förklarar även att Telia lägger in safeguards i sin data genom att aggregera stora grupper vilket säkerställer att det aldrig går att få tillbaka.

R4 förklarar när det gäller etiska avväganden att trots att mobildatan är anonym så tillhandahålls den aldrig i realtid då Telia anser att det skapat för stora sidorisker. Vidare förklarar R4 att syftet med deras verktyg inte är att förstå var folksamlingar är i realtid utan

deras data är fördröjd minst 24 timmar och ofta två dygn. R3 förklarar även att det har lagts stort fokus på hur Telia bibehåller sina kunders integritet. Vidare förklarar R3 att det finns många verksamheter som använder applikationer som är baserade på användarens GPS-platsdata utan att individen riktigt förstår att dennes data samlas in. Telia vill kunna stå för och presentera data på rätt sätt. Vidare förklarar R3 att Telia gör avkall på granularitet vilket gör att man inte kan följa rörelsemönster lika precist som andra typer av datainsamling men datan är anonym och aggregerad. R4 förklarar att Telia ständigt arbetar med att informera hur mobildatan används och fungerar. Utbildningen i datan återges i media, på Telias hemsida och sociala medier. Vidare förklarar R4 att han tror många verksamheter hade kunnat gynnas av den anonyma datan istället för att hantera persondata som är svårare på grund av GDPR.

En annan etisk utmaning som R3 nämner med användning av mobildata är om den skulle behöva samtycke. Vidare förklarar R3 att samtycket är tvådelat, i vissa fall vet inte individer vad de ger samtycke till vilket kan skapa motsättningar. R4 förtydligar dock att Telias data är fri från personuppgifter då den är anonymiserad och aggregerad.

R5 hänvisade till Folkhälsomyndighetens hemsida när vi ställde frågor angående hantering av persondata. R5 förklarade dock att de enbart hade tillgång till den aggregerade och anonymiserade datan.

4.3.4 Övriga utmaningar

R2 nämner att han anser att en av de största utmaningarna var att veta hur man skulle använda datan på bäst sätt. Han nämner också en proportionalitetsprincip vilket handlar om att väga nyttan av datan mot hur inkräktande den är på människor integritet, att hitta balansen där nämner han som en stor utmaning.

R4 menar att den största utmaningen var att många av deras kunder inte arbetat med denna typ av data och insikter tidigare. Under pandemin när det dessutom var ett pressat läge kunde det bli ganska utmanande att försöka lära sig att jobba med något nytt och okänt.

R1 förklarar att extrapolera data ofta är svårt men att många länder nu har kunskap på området. Vidare utvecklar R1 att länder har hanterat pandemin på olika sätt vilket har skapat data som kan nyttjas i framtida pandemier. Den enda utmaningen som R1 ser är att vissa länders datamängder kan vara otillräcklig. Sverige och Danmark har haft förhållandevis komplett data för att följa rörelsemönster hos befolkningen men han nämner att exempelvis länder i Afrika kan ha otillräckliga dataset.

4.4 Möjligheter och framtid

4.4.1 Mobildatans möjligheter

R4 och R2 förklarar att arbetet med mobildata under COVID-19 har skapat en vana att arbeta med big data. Vidare förklarar R4 att under denna pandemi har just mobildata varit relevant men att det har skapat en förståelse för hur man kan använda andra datamängder också. Med den ökade förståelsen för big data tror R4 att man i framtiden kan involvera mer AI i modelleringen istället för rena statistiska modeller. R4 förklarar att framtiden handlar om att

precisera epidemiologiska modeller. Tidigare har dessa modeller varit grovt generaliserade och man har exempelvis undersökt historisk data, det vill säga hur rörelsen mellan regioner var för några år sedan. R1 förklarar även att traditionella modeller inte har baserats på stora mängder data. Vidare utvecklar R1 att under pandemin har mer fokus lagts på att använda faktisk data i modellering och han tror att framtida modeller kommer kunna baseras på mobildata på ett mer effektivt sätt.

“Det vore fantastiskt om man kan använda detta för att skapa early warnings system.” - R4 (Appendix 8, 370-371)

Vidare förklarar R4 att problemet med COVID-19 var att när man upptäckte smittan var det redan försent. R4 har förhoppningen att i framtiden kunna kombinera mobildata och data från tidig RNA-sekvensering för att kunna förutse ett smittoutbrott. R4 förklarar att det skulle kunna leda till en ökad produktivitet i arbetet med att begränsa smittspridning. Dock belyser R4 och R3 att det är kritiskt att arbeta med sådana lösningar trots att det inte är pandemi, beredskapen behöver byggas upp. Även R1 och R2 hoppas att världen kommer vara bättre förberedd inför en liknande pandemi. Vidare förklarar R1 att det har investerats mycket pengar i att utveckla modeller och att det nu finns historisk data som kan användas.

“jag tror det kan finnas en nytta av att följa den här typen av data och även att försöka skapa sig en typ av baslinje som man kan lättare förstå variationerna utifrån.” - R5 (Appendix 9, 213-214)

R5 förklarar att svårigheten är att identifiera vilka typer av rörelse som kan vara associerade med ökad risk för smittspridning. Vidare förklarar R5 att mobildata definitivt kan tillföra värde i arbetet med att skapa en baslinje.

“Det finns inget direkt samband där mer rörelser, mer resor innebär mer exponering för smitta.” - R5 (Appendix 9, 230-231)

Vidare förklarar R5 att man vill kartlägga rörelser som skulle kunna öka smittspridning och även identifiera rörelser som inte gör det. R5 berättar att Folkhälsomyndigheten har arbetat med publikationer angående att försöka skapa en baslinje och kommer att fortsätta göra det. Slutligen menar R5 att mobildata i framtiden kommer vara användbart men att arbetet är väldigt spekulativt då det är osäkert när eller om en ny pandemi skulle bryta ut. Utifrån det förklarar R5 att det är svårt att veta om det behövs en kontinuerlig beredskap och i så fall hur.

R1 förklarar även hur mobildata har möjliggjort billigare forskning genom den tillgängliga datan som kan beskriva rörelser. Vidare berättar R1 att man tidigare behövde samla ett antal människor för att kunna utföra liknande forskning. Smartphones och smartklockor samlar in stora mängder olika datatyper förklarar R1. Vidare förklarar R1 att man kan kombinera dessa datatyper för att prediktera hälsa som ett exempel.

R5 menar att det hade varit positivt om täckningen av mobilmaster blev högre i norrlands inland då det hade ökat representativiteten där. R1 ställer sig lite mer skeptiskt till det och menar att de data man får redan är väldigt korrekta men att det såklart beror på vad man har för syfte. Han nämner däremot också just norrland som ett ställe med dålig masttäckning men förklarar att i en liten stad kanske de flesta ändå träffar varandra och då är en telefonmast tillräcklig medan i en större stad rör man sig i mer specifika områden och därför kan det också behövas fler master.

4.4.2 Mobildata med högre precision

R2 är säker på att korrektheten för mobildatan kommer att förbättras men poängterar att han hoppas folk tänker mer på integritet i framtiden och att data är mer begränsad. Däremot anser han att sättet de jobbade med data var på en bra nivå integritetsmässigt.

På frågan om mer precis data hade kunnat bidra mer till att bromsa smittspridningen inleder R2 med att svara att den mest precisa datan han kan tänka på är GPS-data. Han fortsätter förklara att ifall de haft tid att komma på hur man kan modellera den datan hade det varit givande att använda i deras modeller. Han funderar däremot över hur stor fördel det faktiskt skulle ge. Problemet han förklarar är att man möter väldigt många individer varje dag och om man blir smittad eller inte beror på hur länge man möts men även vilken typ av möte det är. Går man bara förbi någon på gatan blir man nog inte smittad men är man på en fest och dricker ur någons glas är risken för smitta större. Detta är något som rörelsedata inte kan svara på vilket komplicerar bilden och han avslutar med att säga att det nog inte hade varit så givande med jätteprens data ändå. R5 tror inte heller att mer detaljerad data hade varit bättre utifrån sättet de arbetar men påpekar också att de enbart har arbetat med anonymiserad och aggregerad data. Vidare förklarar han att de är mer intresserade av trender snarare än en viss individs exakta GPS-koordinater.

R3 drar paralleller till hur det sett ut i Kina och menar att det inte är ett sätt man vill behandla sin befolkning på. För att få exakt data i Sverige på alla människor hade staten behövt lägga sig i och kräva det. Hon nämner också att de inte har den typen av data, utan deras noggrannhet är som bäst på +/- 200 meter. R3 säger å andra sidan att ifall man håller koll på hela befolkningen så att de inte träffar någon så smittar de ju inte någon, så det hade hjälpt men påpekar igen att det inte är ett samhälle man vill ha. Även R4 drar paralleller till Kina men säger också att det är mindre en fråga om data och mer en fråga om vilka åtgärder man är villig att införa på bekostnad av den personliga integriteten. R4 anser inte heller att det hade gjort någon skillnad om datan inte var anonymiserad och aggregerad. Även R1 håller med om detta.

“I think theoretically you could, yes. But I do not think we have the tools ready to make that good predictions because also like if you do not have anonymised data I would still aggregate it since it makes it a lot easier to work with. Of course you could look at more granular levels but you need to make some kind of assumptions that is more like a, how the population moves as a whole.” - R1 (Appendix 5, 264-268)

Han tillägger dock att vid specifika utbrott exempelvis på en skola hade man eventuellt kunnat få fram lite mer information om man hade ickeanonym data, men det gäller bara mer specifika händelser.

I framtiden tror dock R4 att man med hjälp av med 5G kommer kunna se mer precist var folk befinner sig. Vidare berättar han att det då kan bli så att man behöver medgivande för att använda datan då det blir svårt att anonymisera när datan är så exakt vilket kommer leda till andra arbetsprocesser. Han tror att de tekniska möjligheterna såväl som marknadens mognad kommer att spela stor roll i framtiden och i detta scenario.

4.4.3 Fortsatt arbete med mobildata

R1 berättar att deras hemsida fortfarande ligger uppe men att den inte uppdateras längre eftersom de inte längre får gratis data från telekomföretagen. Facebook ger också ut gratis data, något de använt sig lite av men han påpekar att den datan inte är lika korrekt som datan från telekomföretagen. Han nämner dock att regeringen ibland hör av sig om vissa uppdateringar på deras sida och då gör han det, numera måste han dock uppdatera hemsidan manuellt förut gjordes det automatiskt.

R3, R4 och R5 berättar alla att samarbetet mellan Telia och Folkhälsomyndigheten fortgår. R3 berättar vidare att de använder datan för att titta på samhällsrörelser och förklarar att även om hon inte vet om något tydligt syfte i dagsläget så vet Folkhälsomyndigheten i alla fall om att man kan arbeta med mobildata nu, något de inte visste innan COVID-19. R4 hoppas att datan de ger fortsätter vara viktig för dem och att de är en viktig kund hos Telia.

R2 ser även fler möjligheter med mobildata än att spåra smittsamma sjukdomar i framtiden. Han nämner att det gjorts försök att följa personens rörelse i områden som drabbats av katastrofer något han anser är ett viktigare arbete än folks rörelse under nedstängningar. Han förklarar uttalandet genom att säga att katastrofer runt om i världen är mer förekommande än stora pandemier. Vidare berättar R2 att han hoppas att det forskas mer på området rörelsedata och vad det kan användas till nu när man under pandemin har sett vilken potential det har.

R3 tror också på nya och förbättrade möjligheter med mobildata. Exempelvis tror hon att 5G har en stor potential gällande positionering men påpekar att det centrala är transparensen kring positionering. Hon berättar bland annat om olika studier för dynamiska vägtullar.

“Ehm jag betalar ju mycket hellre skatt beroende på hur mycket jag kör på vägen, än betalar någon schablon om jag inte kör alls, det känns ju inte rättvist. Så då kanske jag skulle vara villig att dela min location när jag kör på vägen för att jag skulle kunna möjliggöra så att jag skulle kunna få rättvis skatt.” -R3 (Appendix 7, 338-341)

R3 tror att om man är tydlig med vad datan används till och människor kan få ett värde tillbaka för att dela sin position kommer många vara villiga att göra det vilket kan leda till flera framtida användningsområden.

5 Diskussion

I denna del diskuteras empirin genom att ställas mot litteraturgenomgången. Materialet analyseras och diskuteras med fokus på att identifiera skillnader och likheter samt urskilja nya perspektiv som framkom under intervjuerna. Slutligen presenteras metoddiskussionen där en reflektion över datainsamlingen görs.

5.1 Resultatdiskussion

5.1.1 Användning och nytta med mobildata

5.1.1.1 Användning av mobildata

Respondenterna R3, R4 och R5 är eniga om att en kritisk del i att transformera data till värdefulla insikter handlar om att göra rådatan begriplig eller att definiera en eller flera frågor som kan besvaras av datakällan. Litteraturen bekräftar även detta då definitionen av dataanalys refererar till processen att samla och analysera big data för att få ut användbar information (Arena & Pau, 2020). Vidare förtydligar R3 och R4 att datamängder ger många svar men att det inte handlar om att få ut så många svar som möjligt av en datamängd utan det kritiska är att definiera frågor som faktiskt är tillräckligt relevanta för att besvaras. Vikten av att datan är meningsfull så att den blir användbar och möjlig att agera utifrån är även något som litteraturen bekräftar (Khoury & Ioannidis, 2014; Ågren, Bjelkmar & Allison, 2021). R4 konkluderar att om man ställt rätt fråga till datamängden och vet vad man ska göra av den, är datan användbar och möjlig att agera ifrån.

R4 belyser att teknikutvecklingen har varit avgörande för att kunna använda mobildata i den omfattning som den har gjorts under COVID-19. Detta är även något som R1 instämmer med och förklarar att traditionella modelleringar inte har baserats på stora mängder data. Sarangi och Sharma (2019) förklarar att kapaciteten för datalagring har ökat samtidigt som den har blivit betydligt billigare. Detta skulle kunna bekräfta respondenternas syn på att mobildatan är mer tillgänglig än tidigare på grund av teknikutvecklingen. Dolley (2018) förtydligar respondenternas resonemang att tillväxten samt variationen av big data kommer utgöra en viktig del för framtidens folkhälsa. Sarangi och Sharma (2019) förklarar även att företag nu kan hyra in sig för lagringsutrymme istället för att investera i dyr infrastruktur. Detta har varit en möjliggörare för Telias datamängd genom att använda tjänsten Google Cloud enligt R4.

Rörelsedata som har baserats på data från mobiloperatörer är den typ av data som har använts av myndigheter i både Danmark och Sverige enligt samtliga respondenter. Samtliga respondenter är eniga att mobildata har använts under COVID-19 för att följa upp grupper eller en befolknings rörelsemönster genom att följa vilka signalmaster en mobiltelefon har varit i anslutning till. Vidare förklarar respondenterna att man inte har följt upp rörelsemönster på individnivå utan att man har fokuserat på grupper rörelser mellan regioner, kommuner, städer, länder eller specifika platser. I både Danmark och Sverige har detta uppnåtts genom att använda mobiloperatörersdata som är aggregerad och anonymiserad med flera försiktighetsåtgärder för att inte inskränka på individers integritet. I litteraturen förklarar Mooney och Pejaver (2017) att geospatial data är tidsbaserad data som återger en specifik

plats och att man med hjälp av denna datatyp kan dra slutsatser om en grupp människors rörelsemönster.

Mobiloperatörsdatan som har använts under COVID-19 för att primärt följa en befolknings rörelse beskrivs som väldigt komplett data då den kan representera en hel population. Fullständigheten av mobiloperatörsdata är något respondenterna vid flera tillfällen belyser. R1, R3 och R5 förklarar att man genom extrapolering av mobildata kan representera en hel befolknings rörelse trots att man inte har komplett mobildata på en population. Respondenterna förklarar att Google, Apple och Facebook också erbjuder mobildata som kan användas för analyser av rörelsemönster på liknande sätt som mobiloperatörsdata. Dock förklarar en del av respondenterna att datamängden mobiloperatörer kan ta fram är betydligt mer komplett än annan mobildata. Litteraturen belyser även att datan från mobiloperatörer har varit mer representativ än Facebooks data (Edsberg Møllgaard, Lehmann & Alessandretti, 2021). Ågren, Bjelkmar och Allison (2021) framhäver även att Telias data i Norden har kunnat representera hela populationer med hjälp av den aggregerade och anonymiserade mobildatan vilket framhäver fördelen med mobiloperatörsdata.

5.1.1.2 Nyttan med mobildata

R2 var inte enig med de andra respondenterna beträffande frågan om mobildata har bidragit till minskad smittspridning av COVID-19 (fråga 8, Appendix 4). Vidare utvecklade R2 att hans forskningsgrupp hade mycket hopp om hur mobildatan skulle användas för att effektivt bromsa COVID-19 men att stora delar av det initiala arbetet med mobildatan förbisåg personlig integritet. Dock förklarar R2 att det inte går att följa individuella rörelsemönster utan att vara integritetskränkande och därmed anser han att mobildata i sig självt har inte minskat smittspridning av COVID-19. Slutligen förklarar även R2 att danska myndigheter fick en analys av populationens rörelse i form av rörelsemönster men att inget samband mellan ökat antal resor under nedstängning och ökad smittspridning kunde göras. Enligt R2 resonemang skulle man kunna konkludera att mobildata inte har använts för att minska smittspridning i någon större utsträckning. Ett annat resonemang skulle även kunna vara att mobildata indirekt har minskat smittspridning av COVID-19 genom att delvis informera och följa upp rörelser i samband med exempelvis restriktioner. R5 menar att mobildatan har varit en del av arbetet att begränsa smittan men att det har skett på ett indirekt sätt genom att öka förståelsen för hur människor anpassar sig till restriktioner. Även R3 och R4 förklarar hur Telias mobildata har använts i media för att utbilda och informera allmänheten. Vidare menar R3 och R4 en visualisering baserad på data av var folkmassor har befunnit sig också kan vara ett indirekt sätt att minska smittspridning genom att dementera rykten och informera om situationen.

I kontrast till detta var de resterande fyra respondenter överens om att mobildata har bidragit till att minska smittspridningen av COVID-19 (fråga 8, Appendix 2). R1 förklarar att mobildatan har använts i Danmark för att visualisera hur populationen rör sig inom landet på en hemsida. Syftet var att åskådliggöra rörelsedatan på ett enkelt sätt så att allmänheten och danska myndigheter skulle kunna ta del av analyserna. Denna typ av data för att visualisera rörelsemönster kan användas för att förstå geografisk spridning av en smittsam sjukdom bättre och metoden har tidigare varit otillgänglig i brist på mobildata (Oliver et al., 2020). Genom att kombinera Oliver et al. (2020) och R1s resonemang kan man påvisa att rörelsedata har bidragit till att informera politiker och befolkning kring rörelsen under COVID-19 samt hur restriktioner har efterlevts. Även R5 förklarar att Folkhälsomyndigheten har publicerat översikter på rörelsedata veckovis för att informera allmänheten om befolkningens rörelsetrender.

Då Telia under en längre tid har arbetat med att samla in och analysera stora mängder rörelsedata inledde de ett samarbete med Folkhälsomyndigheten under pandemin (Ågren, Bjelkmar & Allison, 2021). R3 och R4 bekräftar detta och är båda eniga om att Telias mobildata har kunnat besvara frågeställningen hur Sveriges population har förflyttat sig under olika former av restriktioner och rekommendationer. Telias data har delvis legat till grund för Folkhälsomyndighetens beslutsfattande under COVID-19 förklarar R3 och R4. R3, R4 och R5 är även eniga om att Telias data används som ett uppföljningsverktyg för att följa upp hur en population har förflyttat sig. Tidigare forskning stärker även resonemanget med att mobildata kan generalisera populationens rörelsemönster inom ett specifikt land för att följa virusutbrott som COVID-19 (Wesolowski et al., 2016). Även litteraturen belyste att Telias rörelsedata har använts för att kvantifiera och följa upp Sveriges populations mottaglighet att följa Folkhälsomyndighetens restriktioner (Ågren, Bjelkmar och Allison, 2021). Tidigare studier bekräftar också att mobildata har ökat effektiviteten i arbetet att följa rörelsemönster för att bekämpa smittsamma sjukdomar såsom kolera, ebola, denguefeber, hiv och malaria (Dolley, 2018).

Oliver et al. (2020) förklarar att mobildata och rörelsemönster kan fastställa vilka restriktioner och åtgärder som fungerar effektivt för att minska smittspridning. Enligt Ågren, Bjelkmar och Allison (2021) var Telias och Folkhälsomyndighetens förhoppning att använda mobildata för att följa befolkningens rörelsemönster och kunna bromsa smittan. R3, R4 och R5 är eniga om att detta är ett mål som har uppnåtts genom att man under pandemin har kunnat förstå följsamhet, restriktioner och rekommendationer. Detta är något Telia och Folkhälsomyndigheten fortfarande arbetar med även fast den värsta delen av pandemin förhoppningsvis är över. Dock har Folkhälsomyndigheten under pandemin arbetat mot att förstå och definiera en baslinje, det vill säga vilka förändringar i rörelse man kan förvänta sig av en population till följd av årstider som ett exempel. Tidigare studier förklarar även vikten av att förstå en pandemis framfart för att kunna skydda en befolkning (Wesolowski et al., 2016) vilket Telia och Folkhälsomyndighetens samarbete har uppnått med hjälp av rörelsedatan enligt respondenterna.

Under COVID-19 har Telias mobildata även varit en del av ett samarbete för att prediktera sjukhusinläggningar. R3 och R4 berättar att en AI-modell som baserades på samma data Telia försåg Folkhälsomyndigheten med användes för att förutspå vårdbehovet på Sahlgrenska sjukhuset. Även en studie som är baserad på Googles mobildata har påvisat att det är möjligt att förutspå vårdbehovet i Sverige på regional nivå med mobildata (Gerlee et al., 2021). Således förklarar respondenten och litteraturen hur mobildata indirekt har bromsat spridningen av COVID-19 genom att möjliggöra planering av vården i förebyggande syfte och på så sätt kunna hjälpa fler individer. Oliver et al. (2020) bekräftar även att rörelsedata kan användas för att utföra prediktiva analyser och identifiera framtida smittspridningar.

5.1.2 Utmaningar och risker

5.1.2.1 Tekniska utmaningar

Mansour, Kajatzi och Ghazawneh (2022) och Ågren, Bjelkmar och Allison (2021) menar att en begränsning med mobildatan är att den saknar exakthet eftersom exaktheten varierar beroende på hur tätt signalmasterna förhåller sig till varandra samt antalet signaler från mobila enheter. Dessutom fungerar datan bäst inom ett tätbefolkat område där det finns mycket data att extrahera vilket skapar brister i att representera alla Sveriges städer och regioner. Detta är något även R1 nämner som en utmaning med mobildatan och att det är viktigt att vara

medveten om denna felaktighet. R3 förklarar att man inte kan vara så granulär när man arbetar med anonym data vilket gör att man inte kan vara så exakt i sina insikter.

Khoury och Ioannidis (2014) ställer sig skeptiska till att lita blint på big data och menar att det krävs en starkare epidemiologisk grund för big data då de anser att den idag baseras på datas tillgänglighet. Vidare förklarar de att forskningsnoggrannheten aldrig är starkare än dess svagaste länk samt att det är viktigt att ha "hela bilden" i åtanke för att undvika fel. Samtliga respondenter är eniga kring relevansen av att kombinera olika datamängder för att kunna stärka datans trovärdighet. R5 utvecklar att man måste vara försiktig att dra slutsatser baserat på data och att man både behöver väga datakällor mot varandra samtidigt som de behöver jämföras med epidemiologiska principer kring smittspridning.

Khoury och Ioannidis (2014) menar att det finns en risk att big data inte alltid ger rätt svar eftersom big datas styrka ligger i att hitta samband snarare än i att bevisa om sambanden faktiskt har betydelse. Detta nämner även R3:

“vi kan ju titta på beteenden det ser ut som folk åker till jobbet i den här resan, de har ett beteende som vi kan anta är att man pendlar till en arbetsplats, men vi vet ju inte vad de gör vid sin destination. Det är en anonym analys, vi har ingen aning om man är ute och shoppar eller om pendlar...” - R3 (Appendix 7, 210-213)

Dessa analyser handlar om att anta och förlita sig på statistik vilket oftast stämmer men det finns en viss osäkerhet säger R3. Även R5 förklarar svårigheten med att få rätt bild av datan och berättar vidare att lösningen kan vara att koppla rörelsedatan till faktorer som man generellt ser i smittspridning.

Från intervjuerna framkom även en utmaning som inte nämns i litteraturgenomgången där R1 berättar att de i Danmark tog emot data från flera olika telekomföretag. Dock var datan förarbetad vilket gjorde att de olika företagen hade förarbetat den på olika sätt. Detta blev en utmaning i Danmark att kombinera de olika dataseten. I Sverige var detta inget problem eftersom det enbart var Telia som försedde Folkhälsomyndigheten med data.

5.1.2.2 Utmaningar med datakvalitet

Dolley (2018) nämner datakvaliteten som en utmaning, att ignorera egenskaper som aktualitet, noggrannhet, fullständighet eller tillförlitlighet leder till forskningssvaghet då man kan riskera att skada det allmänna förtroendet för big data och dess positiva effekter på folkhälsan. R4 nämner att hög datakvalitet är en av de viktigaste delarna i Telias arbete. För att säkerställa detta gör de även samarbeten där de jämför sin data med exempelvis statistik. R4 säger också att inga datakällor är perfekta men att Telia kontinuerligt arbetar med öppenhet och att förbättra kvaliteten.

5.1.2.3 Etiska utmaningar

Wesolowski et al. (2016) nämner att kopplingen mellan rörelsemönster och smittspårning av en infektionssjukdom i realtid är en utmaning vilket även Mansour, Kajatzi och Ghazawneh (2022) bekräftar på grund av anonymiseringsprocessen. Enligt R4 är riskerna för stora med realtidsdata och därför är Telias data alltid fördröjd åtminstone 24 timmar. Även R3 nämner att detta är på grund av att det är viktigt för Telia att bibehålla sina kunders integritet. Vidare förklarar R3 att det finns många verksamheter som använder applikationer som är baserade på användarens GPS-platsdata utan att individen förstår att dennes data samlas in vilket inte är etiskt korrekt. Telia gör hellre avkall på granularitet och realtid för att rörelsedatan ska förbli anonym. Eftersom datan inkommer retroaktivt till Folkhälsomyndigheten anser R5 att det kan

vara svårt att följa upp om en rekommendation följs direkt eller inte vilket gör att det är viktigt att inte dra förenklade slutsatser av den analyserade datan. Litteratur och respondenterna belyser svårigheten att behandla realtidsdata när man värnar om personlig integritet.

Både Dolley (2018) och Mooney och Pejaver (2017) nämner risken med att identifiera specifika individer när man använder big data. Även Wesolowski et al. (2016) menar att en koppling mellan ett specifikt fall av en infektionssjukdom och ett specifikt rörelsemönster som kan länkas till en enskild individ sannolikt aldrig kommer att kunna hanteras på grund av att man vill bevara mobilanvändares anonymitet. R2 är skeptisk till om individuell rörelsedata faktiskt skulle bidra till mer specifik data. Problemet som han förklarar är att man möter många personer varje dag och om man blir smittad eller inte beror på vilken typ av möte det är. R5 tror inte heller att mer detaljerad data hade varit att föredra utifrån sättet Folkhälsomyndigheten arbetar men påpekar också att de bara arbetat med anonymiserad och aggregerad data. R1 nämner att vid specifika utbrott exempelvis på en skola hade man eventuellt kunnat få fram lite mer information om man hade ickeanonym data. Litteraturen belyser begränsningen med att följa upp rörelsedata på individnivå då det inskränker på den personliga integriteten. Däremot menar respondenterna att det inte är något problem i dagsläget eftersom de följer trender snarare än individer.

Mansour, Kajatzi och Ghazawneh (2022), Jansen et. al (2021) samt Oliver et al. (2020) nämner alla att allmänheten saknar ett visst förtroende för hur deras personliga data hanteras. Trots att R1, R2, R3 och R4 är övertygade om att deras data är anonym så understryker de alla vikten av att arbeta transparent med hur och i vilket syfte de använder data för att skapa ett större förtroende hos befolkningen. Jansen et al. (2021) förklarar också att förtroendet hos ett lands myndighet avtar ifall de tvingar människor att dela sin personliga data. R1, R3 och R4 håller med om detta men förtydligar att det inte är den typen av data de arbetar med och de alla anser att det inte är ett sätt man vill behandla sin befolkning på. R2 nämner en proportionalitetsprincip vilket handlar om att väga nyttan av datan mot hur inkräktande den är på människors integritet. En annan lösning kan vara att som Apple och Google ha spårningsmetoder som baseras på att dela informationen mellan varje enskild smartphone-användare snarare än att datan förvaras av en statlig myndighet. Denna typ av datainsamling hjälper till att varna personer som har blivit utsatta för viruset samtidigt som man förhindrar potentiell kränkning av mänskliga rättigheter beträffande integritet (Urbaczewski & Lee, 2020).

Dolley (2018) nämner risken med att big data inom folkhälsan kan hamna i fel händer. Det är dock inget R3 och R4 är oroliga för eftersom deras data är anonymiserad och aggregerad vilket inte klassificeras som persondata. Dessutom har Telia tagit extern hjälp för att försöka återidentifiera den anonyma mobildatan vilket har misslyckats. Både R1 och R2 nämner också att de inte är oroliga över att deras data ska hamna i fel händer baserat på deras erfarenheter med hur Danmarks myndigheter har arbetat med mobildata.

5.1.2.4 Övriga utmaningar

Oliver et al. (2020) beskriver att det finns liten politisk vilja och få investerade resurser i att stödja förberedelser för omedelbara åtgärder vid en framtida pandemi. R1 bekräftar detta eftersom hemsidan han arbetat med fortfarande ligger uppe men inte längre uppdateras då de slutat få gratis data från telekomföretagen. I kontrast berättar R3, R4 och R5 att samarbetet mellan Telia och Folkhälsomyndigheten fortgår. R3 är osäker på syftet i dagsläget men poängterar att Folkhälsomyndigheten nu vet att man kan arbeta med mobildata, något de inte

visste innan COVID-19. R4 hoppas att datan de ger fortsätter vara viktig för dem och att de är en viktig kund hos Telia. Precis som litteraturen säger har alltså beredskapen minskat men det kommer troligtvis inte vara svårt att ta vid om en ny pandemi skulle inträffa eftersom samarbetena finns kvar.

Dolley (2018) nämner att en risk med användandet av big data i sjukdomsprevention är att det enbart fungerar i rikare länder i världen vilket kan leda till att ojämlikheten ökar. Detta nämner även R1 som en utmaning då vissa länder har otillräcklig data eftersom det blir svårare att dra slutsatser då. Detta kan bli ett problem exempelvis i Afrika där infrastrukturen inte finns på plats i samma utsträckning som i Norden.

5.1.3 Möjligheter och framtid

Flera av respondenterna belyser att man under COVID-19 har byggt upp en vana att arbeta med den tillgängliga mobildatan. En framtidsmöjlighet majoriteten av respondenterna nämner är att precisera de epidemiologiska modellerna ytterligare genom att integrera mer mobildata och att förstå en befolknings rörelsetrender bättre. R5 förtydligar att man innan COVID-19 inte har haft någon data på befolkningens rörelse och att Folkhälsomyndigheten behöver fortsätta definiera en baslinjeperiod. Detta för att med större säkerhet kunna bevisa ett samband mellan rörelsemönster och smittspridning. R5 förklarar att det exempelvis inte finns något direkt samband som säger att fler resor innebär större exponering för smitta. Ågren, Bjelkmar och Allison (2021) förklarar att mobildatan under COVID-19 har jämförts med perioder innan pandemin. Baserat på litteraturen och respondenterna kommer man i framtiden att kunna skapa en mer konkret baslinje som kan användas för att förstå en befolknings rörelsemönster bättre. Även Wesolowski et al. (2016) menar att mobildata har potential att precisera epidemiologiska modeller vilket förstärker respondenternas framtidstro.

R4 förklarar att en förhoppning är att kombinera mobildata med andra datakällor för att skapa varningssystem i förebyggande syfte och berättar att man under COVID-19 upptäckte smittan för sent. Mansour, Kajatzi och Ghazawneh (2022) belyser också potentialen att använda realtidsdata från mobiltelefoner för att upptäcka virusutbrott mer effektivt och underlätta beredskapen. R5 förklarar att mobildata kommer användas i framtiden men att arbetet framåt är spekulativt då det är osäkert om och när en ny pandemi skulle bryta ut. Detta beskriver komplexiteten om att det behövs en kontinuerlig beredskap och hur den i sådant fall ska utformas.

R4 förklarar att Telia hade ett samarbete med Sahlgrenska sjukhuset som lyckades prediktera antalet sjukhusinläggningar genom att använda mobildata i en AI-modellering. Litteraturen nämner att en liknande modell baserad på Googles mobildata har påvisat att det är möjligt att förutspå sjukhusinläggningar på regional nivå (Gerlee et al., 2021). Genom att kombinera resonemangen av Telias samarbete med Sahlgrenska sjukhuset och litteraturen skulle en framtidsmöjlighet kunna vara att använda sig av prediktion för sjukhusinläggning på en regional nivå i Sverige.

R1 förklarar att framtidsmöjligheterna för att arbeta mer med AI i modeller är god då man under pandemin använt sig av dem i syfte att förstå rörelsemönster. Vidare menar R1 att forskningen kring rörelsemönster har blivit billigare att utföra på grund av tillgängligheten av mobildata. I litteraturen återfinns att infrastrukturen kring datalagring har blivit billigare (Sarangi & Sharma, 2019) vilket stärker R1s uttalande. Genom att använda mer AI i

epidemiologisk modellering samtidigt som forskningen blir billigare att utföra skapas bättre förutsättningar att förstå hur man kan motverka spridning av smittsamma sjukdomar.

Utöver litteraturen är majoriteten av respondenterna eniga om att precisionen av mobildata kommer förbättras i framtiden. R3 och R4 förklarar att man med 5G kommer kunna se mer precist var folk befinner sig. Dock skapar detta utmaningar då rörelsedatan kommer behöva medgivande för att kunna användas. Den tekniska utvecklingen samt marknadens mognad menar R4 att kommer vara kritiskt för användningen av denna data.

I empirin framkom även nya användningsområden för mobildata som inte återfanns i litteraturen. R2 nämner möjligheten med att följa människors rörelse i katastrofdrabbade områden. Han förklarar att det redan gjorts försök kring detta men att han ser stor potential på detta område. R3 tror att ett framtida användningsområde kan vara att människor delar med sig av data för att få värde tillbaka. Exempelvis skulle det kunna vara dynamiska vägtullar där man genom att dela sin position får betala skatt utefter hur mycket bil man kör.

5.2 Metoddiskussion

Nackdelen med en kvalitativ metod är att den kan vara resurskrävande enligt Jacobsen (2002). Det har vi märkt av i form av att det inte alltid är så lätt att få tag på de man vill intervjua då många har högre positioner och kan vara mycket upptagna vilket gjort att det i vissa fall tagit tid att boka intervjun. Dessutom var det mycket data som samlades in via intervjuerna. Även Oates (2006) nämner att en nackdel med kvalitativ dataanalys kan vara att man blir överväldigad av mängden data från få intervjuer samt att det blir svårt att identifiera teman. Vi valde därför en deduktiv analys då vi först identifierade relevanta teman och sedan kategoriserade datan från intervjuerna. Det tog dock mycket tid och det var inte alltid lätt att kategorisera datan då ett visst stycke exempelvis kunde passa in under flera teman.

En kvantitativ metod hade troligen varit enklare att sammanställa men studien berör användandet av mobildata under en pandemi vilket fortfarande är förhållandevis outforskade områden. En kvantitativ studie hade blivit svår att genomföra då metoden kräver mycket kunskap hos en större mängd respondenter. Därför blev en kvalitativ studie där vi intervjuade experter på området mer passande för att få svar på vår forskningsfråga. Dessutom är en kvalitativ metod bra när man själv inte har alltför stor kunskap inom området man undersöker då det kan bli svårt att göra en kvantitativ studie och formulera exakta frågor (Jacobsen, 2002). Jacobsen (2002) nämner också att det blir ett färre antal respondenter i en kvalitativ studie med begränsade resurser vilket kan ses som en begränsning i vår studie. För att ändå ge olika perspektiv på forskningsfrågan har vi intervjuat personer med olika bakgrunder, titlar och arbetsplatser.

Valet att ha alla intervjuer på distans blev enkelt då det hade varit tidskrävande att resa runt i Sverige och Danmark vilket hade krävts om vi genomfört intervjuerna på plats. Enligt Jacobsen (2002) har människor lättare att öppna upp sig när intervjun utförs ansikte mot ansikte eftersom respondenten får lättare att uppfatta subtila signaler och minspel vilka går förlorade på distans. Jacobsen (2002) menar även att respondenten är mindre benägen tala sanning i en sådan situation. Även om det därför i viss mån kan ses som en nackdel att ha distansintervjuer anser vi att efter pandemin har många blivit så pass bekväma med det att det inte bör ha påverkat intervjuerna negativt i någon större utsträckning.

Ett etiskt dilemma vi upplevde med intervjuerna var att vi inte ville ta upp för mycket tid hos respondenterna. Många av respondenterna är väldigt upptagna personer på höga positioner och det upplevdes därför svårt att kräva för mycket tid av dem. Vi var även oroliga över ifall de skulle ha tid att delta om intervjuerna blev för tidskrävande. Samtidigt var det viktigt att få tillräckligt utvecklade svar med olika perspektiv på våra frågor eftersom målet med de semistrukturerade intervjuer var att få nyanserade svar. Samtliga intervjuer blev cirka 45 minuter långa och vi upplevde att frågorna besvarades även om ännu längre intervjuer hade kunnat ge mer nyans på vissa frågor men vi ansåg oss ha tillräckligt med material för att kunna dra välinformerade slutsatser.

6 Slutsats

Syftet med studien var att genom en kvalitativ datainsamling undersöka hur mobildata har använts för att begränsa smittspridningen av COVID-19 samt att belysa dess möjligheter och utmaningar. För att uppnå syftet var målet att besvara följande forskningsfråga: *“Hur kan mobildata användas för att bromsa spridningen av smittsamma sjukdomar samt vilka utmaningar och möjligheter finns med detta arbetssätt?”*.

Genom analys av empiri och litteratur kan vi dra slutsatsen att mobildata, primärt från mobiloperatörer i Sverige och Danmark har varit en viktig del i kampen mot COVID-19. Vårt resultat indikerar att mobildata indirekt har använts för att bromsa smittspridningen. Den indirekta och huvudsakliga funktionen som respondenterna samt litteraturen belyser att mobildata har haft, syftar till att fatta beslut om restriktioner och rekommendationer baserat på en befolknings rörelsemönster. Detta innefattar att representera en befolknings rörelsemönster retroaktivt under specifika tidsintervall och områden för att öka förståelsen för hur en population efterlever restriktioner. Vidare har mobildata även använts indirekt för att bromsa smittspridning av COVID-19 genom att sprida kunskap till allmänheten och myndigheter via publicering av rörelserapporter. Slutligen har man kunnat prediktera sjukhusinläggningar baserat på mobildata. För att lyckas med det ovannämnda har följande faktorer varit kritiska; ställa rätt frågor, använda data av hög datakvalitet, kombinera datakällor samt att datamängden är anonymiserad och aggregerad.

Eftersom arbetet med mobildata har organiserats under en pressad situation och är relativt nytt har en del utmaningar uppkommit i användandet. De mest kritiska momentet för att arbeta med mobildata handlar om det etiska perspektivet då processen för att otillgängliggöra möjligheten att återidentifiera mobildata är komplex och har tagit lång tid att utveckla. Trots att mobildatan som har använts för att bromsa smittspridning bevisligen har varit anonym krävs en transparens och öppenhet i användandet för att bibehålla allmänhetens förtroende. Vidare identifierades även följande tekniska utmaningar med mobildata:

- Bristen på exakthet och varierande datakvalitet
- Mobildata kan inte ensamt användas för att beskriva en befolknings rörelsebetenden
- Datakällor kan vara förarbetade på olika sätt

Mobildata har möjliggjort datainsamling för att förstå en befolknings rörelse och skapa en baslinje för hur folk rör sig till följd av parametrar som årstider, högtider, helger och semestrar. Detta är något som tidigare inte har kunnat utföras vilket indirekt har skapat möjligheten att bromsa smittspridningen genom att följa en populations rörelsemönster. Det kommer även vara kritiskt att fortsätta utveckla en populations baslinje för att i framtiden få en bättre förståelse beträffande rörelsebetenden. Vidare har följande möjligheter med mobildata identifierats:

- Prediktera sjukhusinläggningar mer omfattande
- Precisera epidemiologisk modellering
- Vara en del i utvecklandet av ett varningssystem för framtida smittsamma sjukdomar och andra katastrofdrabbade områden
- Bli mer preciserad och individbaserad i samband med 5G där användaren själv får välja om den vill dela sin mobildata för att få värde tillbaka

Avslutningsvis kan det konstateras att mobildata varit en del i att bromsa smittspridningen under COVID-19 och med stor sannolikhet kommer att vara det vid en framtida pandemi eller

smittoutbrott. Det finns även möjlighet att använda mobildatan för andra ännu relativt utforskade områden men det är däremot viktigt att ha dess utmaningar i åtanke.

6.1 Förslag till vidare forskning

Användningen av mobildata för att analysera rörelsemönster under en pandemi är relativt utforskade område. Vi uppmuntrar till fortsatt forskning på området. Ämnen som kan vara intressant att forska vidare på är bland annat hur mobildata kan användas i andra syften än för att bromsa smittspridning. Vidare hade det varit intressant att se vad som händer vid nedrustningen av pandemin på detta område och hur arbetet ser ut framöver. Även att följa upp hur myndigheter och mobiloperatörers samarbete fortskrider är ett relevant område att utforska. Detta för att följa upp de definierade möjligheterna med mobildata som uppsatsen berör och utforska om några användningsområden har utvecklats.

Appendix 1 – Intervjuguide

Kategori	Beskrivning
Introduktion- och bakgrundsfrågor	Introduktion- och bakgrundsfrågor om personen och dennes arbetsroll. Ska ge kontext och förståelse för vem personen är. Det är även bra att börja med “enklare” frågor för att få en mer avslappnad stämning under intervjun.
Användning och nytta av mobildata	Frågorna ska ge förståelse för hur man använt mobildata under COVID-19 samt vad mobildatan har bidragit till.
Förväntningar på samarbetet, kontra utfall	Frågorna ska belysa förväntat resultat mot faktiskt resultat.
Risker och utmaningar	Frågorna ska ge svar på hur de intervjuade ser på olika risker och utmaningar kopplade till användningsområdet.
Möjligheter och framtid	Frågorna ska identifiera olika framtida möjligheter kopplade till användningsområdet.

Appendix 2 – Frågeguide

Indragna frågor är följdfrågor som ställts efter behov och har fungerat som stöd under intervjuerna.

Ämne/Område	Fråga
Introduktion- och bakgrundsfrågor	1. Hur gammal är du?
	2. Vilket kön identifierar du dig som?
	3. Vilket företag jobbar du på och kan du kort beskriva din yrkesroll?
	4. Hur länge har du arbetat med mobildata kopplat till COVID-19?
Användning och nytta av mobildata	5. Vikten av att transformera data till insikt är något som ofta nämns av Telia Crowd insights. Hur skulle du kortfattat beskriva processen från data till insikt. <ul style="list-style-type: none"> • Vad blir outputen av processen • Vad klassificeras som en insikt?
	6. Big data har använts länge inom folkhälsan. På vilket sätt har möjligheten att samordna data och kunna analysera datan med moderna verktyg förenklats arbetet?
	7. Hur har arbetssättet sett ut innan man hade tillgång till mobildata? <ul style="list-style-type: none"> • Har man tidigare studerat rörelsemönster vid smittsamma sjukdomar utan tillgång till mobildata och hur har man utfört dessa rapporter?
	8. Anser du att mobiloperatörsdata har använts under COVID-19 för att bromsa smittspridningen? <ul style="list-style-type: none"> • Om ja, på vilket sätt har den bidragit till bromsad smittspridning?
	9. Vilka slags beslut har man kunnat ta baserat på Telias data? <ul style="list-style-type: none"> • Har dataanalysen legat till stöd för att forma restriktioner, rekommendationer och policies? • Om ja, exemplifiera gärna. • Googles data har använts för att förutspå sjukhusinläggningar är det något ni gjort också? • Kommer mobildata användas för att återgå till det normala samhället i ett postcovidsamhälle?

	<p>10. Många länder implementerade kontaktspåringsapplikationer under COVID-19. Fanns det tankar på att Telia och Folkhälsomyndigheten skulle utforma en?</p> <ul style="list-style-type: none"> • Om ja, varför blev det inte så? • Om nej, fanns det något konkret som gjorde att en sådan plan inte var aktuell?
Förväntningar på samarbetet, kontra utfall	<p>11. Vad fanns det för mål initialt gällande samarbetet mellan Telia och Folkhälsomyndigheten?</p> <ul style="list-style-type: none"> • Vad ville man uppnå med samarbetet?
	<p>12. Har de uppsatta målen beträffande rörelsemönster baserat på mobildata uppnåtts?</p>
	<p>13. Tillkom eller ändrades mål under samarbetets gång?</p>
Risker och utmaningar	<p>14. Vilka är de största utmaningarna för att kunna använda mobildata under COVID-19 anser du?</p>
	<p>15. Hur har kvalitetssäkringsprocessen gått till för att presentera data med så hög kvalite som möjligt?</p> <ul style="list-style-type: none"> • Bortfall av dataset, hantering av dataset som inte kan representera en grupp korrekt. • Overfitting är ett modelleringsfel och kan beskrivas som ett resultat som är baserat på inkorrekt data
	<p>16. Det finns en misstro i samhället gällande hur verksamheter och myndigheter hanterar personlig data. Är det något ni aktivt arbetar med för att bibehålla samhällets förtroende i ert arbete med mobildata?</p>
	<p>17. Finns det brister med mobildatans anonymitet idag?</p> <ul style="list-style-type: none"> • Finns det några hot som skulle kunna påverka eller avslöja den anonyma datan?
Möjligheter och framtid	<p>18. Om man spekulativt inte behövde anonymisera datan som i mindre demokratiska länder, hade man då mer effektivt kunna bromsa smittspridningen med datan?</p>
	<p>19. Tror du exakthet och precision av mobildata kommer förbättras i framtiden?</p> <ul style="list-style-type: none"> • Kommer tätheten av signalmaster öka? • Tätbefolkade områden i jämförelse med mindre städer? Kan datakvaliteten i småstäder komma att bli lika bra som i storstäder? • Undvika fördröjningen av dataleverans (24 h), kan idag realtidsuppdateras till Folkhälsomyndigheten men strider mot lagen, kan detta komma att ändras?

	20. Vad anser du att det finns för framtida användningsområden med mobildata i syfte att bromsa spridning av smittsamma sjukdomar?
--	--

Appendix 3 - Intervjuguide (engelska)

Category	Description
Introduction and background questions	Introduction and background questions about the respondent and the working role of the respondent. Should provide context and understanding of who the person is. It is also good to start with "simpler" questions to get a more relaxed atmosphere during the interview.
The use and contributions of mobile data	The questions should give an understanding of how mobile data has been used during COVID-19 and what mobile data has contributed with.
Expectations of the collaboration between Telia and Folkhälsomyndigheten versus outcomes	The questions should highlight the expected results compared with how the collaboration turned out.
Risks and challenges	The questions should provide an answer on how the respondents view different risks and challenges connected with the area of use regarding mobile data.
Possibilities and future	The questions should identify different future possibilities connected to the area of use.

Appendix 4 – Frågeguide (engelska)

Indragna frågor är följdfrågor som ställts efter behov och har fungerat som stöd under intervjuerna.

Ämne/Område	Fråga
Introduction and background questions	1. How old are you?
	2. Which gender do you identify yourself as?
	3. Which company or institution do you work for, could you describe your working role briefly?
	4. How long have you worked with mobile data connected to COVID-19?
The use and contributions of mobile data	5. Transforming data into insight is often mentioned by Telia Crowd Insights. Could you describe with your own words how mobile data goes from input to output? <ul style="list-style-type: none"> • What would be the output of the process of aggregating mobile data? • What would be considered as an insight?
	6. Big data has been used within public health for a long time. How do you think that the ability to coordinate data and be able to analyze data with modern tools has improved the work with data?
	7. What did the working method look like before you had access to mobile data? <ul style="list-style-type: none"> • Have you previously studied movement patterns in infectious diseases without access to mobile data and how have you performed these reports?
	8. Do you think mobile data has been used during COVID-19 to decrease the transmission of the infection spread? <ul style="list-style-type: none"> • If yes, in what way has mobile data contributed to decreasing the virus to spread? • If no, why not?
	9. Which kind of decisions have been made based on mobile data?

	<ul style="list-style-type: none"> • Has the data analysis been supportive of shaping restrictions, recommendations and guidelines? • If yes, please exemplify • Google's data has been used to predict hospital admissions, is that something you did as well? • Does mobile data play a role when going back to normal life in a postcovid world?
<p>Expectations of the collaboration between Telia and Folkhälsomyndigheten versus outcomes</p>	<p>10. Many countries implemented contact tracing applications during COVID-19. Do you know if there were any thoughts about implementing one in Denmark?</p> <ul style="list-style-type: none"> • If yes, why did it not happen? • If not, was there anything concrete that made such a plan out of the question?
	<p>11. What were the initial goals regarding collaboration between Telia and Folkhälsomyndigheten?</p> <ul style="list-style-type: none"> • What did you/they want to achieve with the collaboration?
	<p>12. Have the set goals regarding movement patterns based on mobile data been achieved?</p>
	<p>13. Were goals added or changed during the collaboration?</p>
<p>Risks and challenges</p>	<p>14. What do you think are the biggest challenges to be able to work with mobile data during COVID-19?</p>
	<p>15. How has the quality assurance process been to present mobile data with as high quality as possible?</p> <ul style="list-style-type: none"> • Loss of datasets, handling of datasets that cannot represent a group correctly • Overfitting is a modeling error and can be described as a result based on incorrect data
	<p>16. There is a mistrust in society regarding how organizations and authorities handle personal data. How do you think the mistrust could be decreased?</p> <ul style="list-style-type: none"> • Are there any key factors for maintaining a society's trust?

	<p>17. Do flaws exist with mobile data's anonymity today?</p> <ul style="list-style-type: none"> • Are there potential threats that could affect or reveal the anonymised data • If not, is anonymised and aggregated data 100% safe?
<p>Possibilities and future</p>	<p>18. If one did not speculatively need to anonymise data as in less democratic countries, could one then have been able to slow down the spread of infection more effectively with data?</p>
	<p>19. Do you think accuracy and precision of mobile data will improve in the future?</p> <ul style="list-style-type: none"> • Will the density of signal masts increase? • Densely populated areas compared to smaller cities. Can the data quality be equally good? • Avoid delay in dataset delivery time? Can be done today but against the GDPR-law, can that change?
	<p>20. What do you think there are for future uses with mobile data in order to slow down the spread of infectious diseases?</p>

1 Appendix 5 – Transkribering intervju 1

2 SL = Simone Lewenhaupt

3 NG = Nils Gunnarson

4 R1 = Respondent 1

NG

5 Let us begin with some introduction and background questions then.

6 SL

7 Yes, so first question, how old are you?

8 R1

9 I am 27.

10 SL

11 Okay and what gender do you identify yourself as?

12 R1

13 As a male.

14 SL

15 Yes and which company or institution do you work for and could you describe your role briefly.

17 R1

18 Yeah, I work for Technical University of Denmark and my role is as a PHD student so basically aspiring researcher I guess.

20 SL

21 Perfect, how long have you worked with mobile data connected to COVID-19.

22 R1

23 Since the beginning, basically. I guess COVID started in march right. I think I began working may, I think, so may 2020.

25 NG

26 Thank you, and now we will go into the specific mobile data questions and the use and contribution of mobile data. So, transforming data into insight is often mentioned by Telia Crowd Insights. Could you describe with your own words how mobile data goes from input to output?

30 R1

31 So, it depends on, the way that we use mobile data is often that telecom companies, I have used it for mobility reasons, how people move around right, and the way it goes is that the telephone company does not have your GPS signal they just know which cell tower you are connected to, so the nearest cell tower that is how of course how they transmit information. So the way we use it is that then they have like these traces of which cell tower and what time you are connected to and then you can infer like how do you move and when do you move, when do you make stops and so on so you can basically get an approximate trace of where you have been moving around. And what we then use the data for is then to not to look at individuals, we do not have access to individual data we just get like collections of mobility throughout, we use it in Denmark right. So in comunes/municipalities so we can see how many move from one municipality to another municipality in different time frames. Then we used that data to then in however we want to model how COVID-spread or where the people move and so on.

44 NG
45 Yes, thank you, we are very happy with that answer. So the next question, big data has been
46 used within public health for a long time. How do you think that the ability to coordinate data
47 and be able to analyse data with modern tools has improved the work with data?
48 R1
49 Can you repeat the start again?
50 NG
51 Yes, big data has been used within public health for a long time. How do you think that the
52 ability to coordinate data and be able to analyse data with modern tools has improved the
53 work with data?
54 R1
55 I think it improved a lot specifically that you can... beforehand often when you did a study,
56 you had to like collect an amount of people and you did a specific study on these people. The
57 difference now is that you have these smartphones, wearables, whatever and they collect a lot
58 of different data and it is not so specific but it also means that you can combine different data
59 types since you have different data types for given individuals. So it has helped to like figure
60 out what more like casual or just correlations between different things from whatever you
61 want to predict so if you want to predict like health or could be cancer or could be something
62 else so instead of having only one type you have a lot of different types that you can combine
63 which makes it a lot more would not say easy. But yeah it makes it a lot more easy to start and
64 try things out but beforehand it was like you really need to commit to a study and it cost a lot
65 of money but now it is a bit more approachable, also for the smaller people.
66 NG
67 Yes that is true, thank you. So the next question what did the working method look like before
68 you had access to mobile data? Maybe you were sort of answering that question on the previ-
69 ous question but.
70 R1
71 Yes so the thing is I did not, in my full research life I have always worked with it also when I
72 did my bachelor thesis as you do now, I also worked with mobile phone data. So I do not re-
73 ally know how it was before I have always used phone data.
74 NG
75 Hahah, yes that is a good answer. One of our maybe most important questions for this project:
76 Do you think mobile data has been used during COVID-19 to decrease the transmission of the
77 infection spread?
78 R1
79 Well, in Denmark I would say yes. I do not know the exact details but you can say one thing
80 we did is that you have all these different telecom companies which agreed to give data. Of
81 course censored in the way that you only get like collections of people. By combining them
82 we for example, in our group we made a website which kind of officialised all this data, one
83 of the things is that. The way it decreases is by letting politicians who control the country
84 know what is going on. So we wanted to make an easy and approachable way for them to go
85 in and look it up. I can send you a link afterwards if you want to look. But we made like a
86 website with visualisations that were interactive so you can kind of go and see like if you
87 were in this commune on this day where did people move and so on. So we kind of wanted to
88 make all this approachable for everyone and we know that, we also did reports which we sent
89 to the Danish government so we know that they used our website for example. I do not know
90 specifically what they used it for but I am quite sure that they took a lot of their policies and
91 what they did is based upon what they saw on this data and also from what we told them. So
92 yes I believed it helped but I do not know exactly how they used it if that makes sense.
93 NG

94 Yes that makes sense and we are really interested to look into your website as well so if you
95 could send the link that would be very interesting for us. But that leads us into one of the next
96 questions which we already touched on and I think you answered it but I will ask it anyways.
97 Which kind of decisions have been made based on mobile data? But I assume that your an-
98 swer on that question is that you are not really sure.
99 R1
100 No I am not sure on the exact decisions no. I could hypothesise but I do not think that it
101 would help that much no.
102 NG
103 Good okay, and the last question for this segment is: Many countries implemented contact
104 tracing applications during COVID-19. Do you know if there were any thoughts about imple-
105 menting one in Denmark?
106 R1
107 Yes we also, we had one an anonymous one. So, I know Sune was also an advisor on that one,
108 he did not make it but he was an advisor. Basically it was using bluetooth signals and collect-
109 ing bluetooth, like seeing where, who... basically just like taring and so on like everyone is an
110 individual and everything is anonymised and you go around and then it says like you have
111 been close to this bluetooth signal and then you could go in on the app if you got COVID and
112 say I am sick. Then the app would check who you had been close to with your bluetooth sig-
113 nal. So yeah we had one and it was good in the start when so few people were sick but the
114 higher the infections the worse these approach gets because basically wherever you walk it
115 would just say "oh you have been close to one", "oh you have been close to one", "oh you
116 have been close to one". Then basically every time you had to go to the centre to get a covid-
117 test and the when you got home the next day you would get a new notification that would say
118 "oh you have been close to one" because in the queue to get the test you were always close to
119 people who got one. This made it like yeah fuck that, but in the start it was probably worth it.
120 NG
121 Yeah, thank you that was helpful in understanding we have not heard about that in Denmark.
122 Okay now we are going into the third segment with expectations between the collaboration
123 between Telia and the public health agency in Sweden so now we will see how you approach
124 these questions as we mentioned before.
125 SL
126 Yes but you can apply them to telecom companies in Denmark and the politicians in Denmark
127 since it is probably easier for you. So do you know what were the initial goals regarding col-
128 laboration between Telecom companies and yeah public health agencies.
129 R1
130 Well basically it is just because like who knows where you are, google knows, maybe apple
131 knows but besides that nobody really knows except telecom companies. Of course they do not
132 have exact locations. So the collaboration was to, at least actually I think all the telecom com-
133 panies in Denmark agreed to give some form of estimate about where people were during the
134 day to the government. So that researchers then could analyse and try to use the information
135 to model infectious spread. So some people used like traditional epidemiological modelling to
136 see like where people go and so on. We personally took the approach to use more machine
137 learning, data driven methods, to predict stuff or to try to understand the data more. Me per-
138 sonally I made a paper regarding this how we basically split up these mobility patterns like
139 this network how movement around in Denmark. Into like a working pattern, weekend pattern
140 and leisures and holiday pattern. Then we could see over time how these patterns changed so
141 you could see like in the start people pretended to go from a more working pattern into week-
142 end pattern during lockdown which makes sense since you are staying at home. But then
143 through time with different outbreaks you could basically see what happened and this is how

144 we, or not we the government tried to collaborate with the telecom companies and they sup-
145 plied all this information for free, at least in the start upon the end of 2020 I think some com-
146 panies began saying we could not get more data.
147 SL
148 Okay yeah and next question is have the set goals regarding movement patterns based on mo-
149 bile data been achieved? If you know if you can answer this.
150 R1
151 I do not know, I actually do not know what the exact goals were. I have not seen why they did
152 that. As I said I came in two months after the start so I just got the benefits not really the ques-
153 tions so much so I do not know.
154 SL
155 Haha, no that is fine, do you know if there were goals added or changed during the collabora-
156 tion?
157 R1
158 Again, I do not know.
159 SL
160 Haha that is fine. So now we will move on tho the next section about risks and challenges.
161 R1
162 Mm
163 NG
164 So, what do you think are the biggest challenges to be able to work with mobile data during
165 COVID-19?
166 R1
167 One of the bigger challenges is a technical one. That is that mobile phone data is representa-
168 tive, that means that mobile phone companies will not give you individuals so they will have
169 to preprocess the data before we get it. The problem is that when you have a lot of different
170 telecom companies, every company preprocess that data in different ways, that means for ex-
171 ample you can imagine that if you have this data of where you are like which cell tower you
172 are connected to, converting that into actually like locations like everytime you did a stop. Be-
173 cause I do not want to know if you went from northern Sweden, you went from Göteneim or
174 what is it called in northern Sweden and then went to Gothenburg then I do not want to know
175 the full trip I just want to know you started up there in Umeå or whatever and then you went
176 to Gothenburg. But how do you measure what is a stop and what is a location? That is a non-
177 trivial task and different companies did this in trivial ways and non-trivial ways. So some data
178 sets were a lot better than others and then when you try to combine data sets they kind of like
179 it does not work that well always. So one of the challenges was which we would have liked to
180 basically make like a more standard way to preprocess data which would have been much eas-
181 ier for us researchers to use the data. Can you repeat the question again I feel like there was
182 something more I wanted to say but I do not remember?
183 NG
184 Yes, what do you think are the biggest challenges to be able to work with mobile data during
185 COVID-19?
186 R1
187 Yes I would say that is probably that haha.
188 NG
189 Yes I can understand that that is a big challenge when trying to compile all the data.
190 R1
191 Yes.
192 NG

193 How has the quality assurance process been to present mobile data with as high quality as
194 possible?
195 R1
196 It kind of goes back to the other thing again which means that it is, we do not know how com-
197 mon it is to preprocess the data so we can only take what they have and take their word for it.
198 Since it is like you know their algorithm and their way of doing it and we are not allowed to
199 do it because then we could backtrack maybe. So I would say that is one of the bigger chal-
200 lenges again.
201 NG
202 Yes because you are not specifically involved in like the lost data sets from the telecom com-
203 panies or overfitting or things like that.
204 R1
205 No, so for example also you could see that for some of the companies you never really saw
206 long trips. Because let us assume that you went to Gothenburg, then probably along the way
207 you would take a pee break or something. Then some companies would take that as a stop
208 where other companies would not take it as a stop and it would be more like “oh we know
209 you did not stay you were on a long trip”. So in some of the data sets you never saw long trav-
210 els or distances since you basically did it but it just thought that you took a lot of small ones
211 which kind is of like noise in your data.
212 NG
213 Okay, moving on to the next question there is a mistrust in society regarding how organisa-
214 tions and authorities handle personal data. How do you think the mistrust could be de-
215 creased?
216 R1
217 I feel like actually authorities, especially like the way that we got it like researchers, one of
218 the good things about research compared to private companies is that we actually publish the
219 stuff that we work on so it is basically free for everyone to know. I think it is not a trual solu-
220 tion to how to decrease the mistrust other than just inform what it is used to the problem is
221 that you can use all this information to if it comes to the wrong hands to just gain money right
222 and we want to avoid. Just like inform people and then also make things public especially
223 when you do research at universities it is public so then I think yeah you will understand bet-
224 ter but it is also all the other stuff that I think people will mistrust behind closed doors. For a
225 lot of the decisions in Denmark it was really hard from the government side, often experts
226 said something but they did something different I think that can lead to a little bit of mistrust
227 and so on. But I think as long as it is transparent then I think it will be more trustworthy.
228 NG
229 Yes, and next question which sort of connects with the question we just discussed but do
230 flaws exist with mobile datas anonymity today?
231 R1
232 Depends on what you mean with flaw but one flaw with mobile data from telecom companies
233 is that it is not very specific, as I said it is only like what is your nearest cell tower, and a cell
234 tower can be several kilometres apart right. So depending on what you want to use the data
235 for you need to be aware that there is this bias towards it, it is not very granular or how you
236 put it. Another flaw of course is to retain privacy, you can not, if you have some location
237 where, in Denmark I think it was less than ten people so if less than ten people took a trip in
238 some given time frame that it had data from then it would just pound as a zero even though
239 there actually was trips. That was also to say like if there are so few people living in some
240 place in the middle of nowhere then it would be very easy to backtrack who it actually was.
241 But it depends on what you see as a flaw because this is more things that you should be aware
242 of when you are using it as a datasource.

243 NG
244 Would you say that anonymised and aggregated data is a 100% safe in the regard of back-
245 tracking as we were discussing before, to track it down to a specific person or group is say so?
246 R1
247 I would say as long as you, with all data if you have enough outside information you could
248 probably backtrack something. Let us assume that I both have your telecom information and I
249 also knew something else about you as a individual I could probably do it but since we have
250 so much data and we removed trips and X amount of people I think it would be very hard. I
251 would say it is very truthful and you should not be afraid of like your personal information
252 getting leaked.
253 NG
254 I agree, and now we will discuss the last topic about possibilities and the future regarding mo-
255 bile data.
256 SL
257 Yeah so, if one did not speculatively need to anonymise data as in less democratic countries,
258 could one then have been able to slow down the spread of infection more effectively with
259 data?
260 R1
261 Hmmm, I think theoretically you could, yes. But I do not think we have the tools ready to
262 make that good predictions because also like if you do not have anonymised data I would still
263 aggregate it since it makes it a lot easier to work with. Of course you could look at more gran-
264 ular levels but you need to make some kind of assumptions that is more like a, how the popu-
265 lation moves as a whole. So in theory yes, but I do not think that it would have helped that
266 much. It might have been only if there were some specific outbreaks like you know that in
267 Denmark we had some outbreaks at schools, then you could have a little more information if
268 you had individuals but is it more of a special occasion.
269 SL
270 Yes. Do you think accuracy and precision of mobile data will improve in the future?
271 R1
272 Well it is already pretty accurate if you just have the right information, depends on who you
273 are. So of course telecom companies only know the granularity of their antennas, but if you
274 are google or something that uses it then you know specifically where you are, so I do not
275 think it will improve that much but it depends on who you are.
276 NG
277 Yes, but in regards of like for a telecom company with density of signal masts, I know in
278 Sweden for example that the data is less accurate, or there is a smaller amount of data sets in
279 smaller cities. Do you think that could be improved by adding more signal masts?
280 R1
281 Yes a 100%, so I think in Denmark we are a pretty small country, kind of dense, when in
282 Sweden you have the problem that you have a lot of non-populated areas where do not know
283 that much but then again for spreading of diseases you are probably not going to spread the
284 disease when you are out in the middle of nowhere so I do not think it matters that much. Of
285 course you might have some small societies around the country where you could improve a
286 little bit but I honestly think if you are in a small society everybody knows everyone and, now
287 I am going to make a lot of assumptions, I am sorry if that is wrong, but I think it is more
288 like... If you live in a big city you actually meet everyone but in a small society you meet
289 everyone. So if you only have one antenna in one city it is probably a good enough measure.
290 NG
291 Yes, a valid point.
292 SL

293 So the last question, what do you think there are for future uses with mobile data in order to
294 slow down the spread of infectious diseases?
295 R1
296 I think one of the good things now compared to traditional modelling like logical modelling, it
297 is more like a random thing, like a statistics space model but now you actually have actual
298 data. COVID is the first instance where we had all attention and this focus on trying to use ac-
299 tual data to make models and I think from now on models will become a lot better where we
300 can use mobile phone data in more efficient ways.
301 NG
302 Yes, that is a good answer, and do you think like these models will be ready for a country to
303 have and work with proactively so they can launch it if a pandemic like COVID hits again,
304 that they will be more like prepared?
305 R1
306 I hope, I would believe that if it happens again we would be better, I am not sure if we have it
307 right now like in this instance but I think that there has been a lot of money invested into like
308 developing models and now we actually have some historical data that we can use. Because
309 often it is very hard to extrapolate right but now we actually have some that from what hap-
310 pened and one of the good things is that a lot of countries did a lot of different things. If you
311 take Sweden compared to Denmark like Sweden was very relaxed in the start and was like
312 “whatever we will see what happens” when Denmark had full lockdown from the start and we
313 had so many countries that did it in different ways which we can kind of like from that infer
314 which methods did which things. The only problem you could encounter is like insufficient
315 data where I think in Denmark we had pretty good data coverage whereas in Sweden you had
316 maybe not as good but maybe good enough since you still have data from public cities it is
317 still good coverage just like out in the middle of nowhere. But if you take like countries like
318 in Africa for example it would probably be a lot harder since they do not have the infrastruc-
319 ture to support this type of approach. I do not know but I would assume so.
320 NG
321 Yes, that is a good answer. The website you launched is that still in use in a post COVID
322 world, that is another thing we understood while doing research that it is a problem that many
323 are doing research and progress but then the connection between establishing guidelines or
324 models that will be used in the future is often forgotten or they do not invest in the actual
325 models from a state perspective.
326 R1
327 From our website point of view, we do not get data from the telecom companies anymore
328 since we had a deal where they gave it for free basically and now you have to pay for it and
329 the government universities do not do that anymore. Facebook also had something called fa-
330 cebook data for good, where they basically gave out for free tracking information about peo-
331 ple around the world, same as telecom data but it is not as, not everyone does not have face-
332 book and I also think that it is more generation based. We could still update it but we stopped
333 updating it. But sometimes the government comes like we have this can you please update it
334 and then I do it, back in the days we had it automatic then it updated all the time. So it is still
335 there but we are not updating it anymore so it depends on how you look at it.
336 NG
337 That wraps the interview up actually we have asked all the questions and we are really happy
338 with your responses as well.
339 R1
340 Yes, I am happy to hear.
341 NG

342 And thank you very much for taking your time and as we said before we started we will tran-
343 scribe the interview and we will send it over to you if you want to look through it and confirm
344 that we can use it in our report we would be really happy.
345 R1
346 Yes, I will do so, no worries.
347 SL
348 Perfect.
349 NG
350 Thank you very much.
351 SL
352 Thank you.
353 NG
354 Wish you a nice end of the week.
355 R1
356 Thank you and good luck with your thesis and everything.
357 NG
358 Thank you.
359 SL
360 Thank you, bye.
361 R1
362 Bye.
363 NG
364 Bye.

1 Appendix 6 – Transkribering intervju 2

2 SL = Simone Lewenhaupt

3 NG = Nils Gunnarson

4 R2 = Respondent 2

5

6 NG

7 With that said, it is time to begin the interview with some background and introduction ques-
8 tions.

9 SL

10 Yes, that is right. How old are you?

11 R2

12 I am 30 years old.

13 SL

14 Which gender do you identify yourself as?

15 R2

16 Male.

17 SL

18 Which company or institution do you work for and can you describe your role briefly?

19 R2

20 Right now I am working for Cornell University in the US. I am a postdoctoral researcher,
21 which is a researcher employed with a temporary contract following one's phd. At the time
22 that I was working with COVID in Denmark I was employed as a postdoctoral researcher at
23 the Danish Technology University and I was a part of what's called the expert modeling group
24 of COVID-19 in Denmark at that time. I participated in that between April 1st and August
25 15th, so for four months and after that I started with my US contract and I did not work with
26 this anymore.

27 SL

28 Aha, okay. Then you already answered our question regarding how long you have worked
29 with COVID-19.

30 R2

31 Yes and I can tell you with this specific thing about the cellphone data. All the models were
32 developed on a rolling basis. Apparently the Danish authorities found out that they had no
33 models when the outbreak happened and then they started out making simple models and
34 added complexity on top of that. And my role was to to... So I was interested in how telecom
35 data and how models can interact with each other. So my role was to be one step ahead of the
36 actual models and think about how telecom data could be implemented in the models. But I
37 left the group so early so I am not really sure exactly what was done after I left. So I can tell
38 you about my thoughts about these things but I honestly don't know how exactly the telecom
39 data ended up being used in the Danish models.

40 SL

41 That is totally fine, we are just interested to hear your opinion on the different questions.

42 R2

43 Okay, good.

44 NG
45 Lets go over the use and contributions of mobile data. First question. Transforming data into
46 insight is often mentioned by Telia Crowd Insights, could you describe with your own words
47 how mobile data goes from input to output?
48 R2
49 I am not sure what you mean by the question how mobile data goes from input to output?
50 NG
51 I mean from the mobile signal being collected to how you are using it for modeling like a sce-
52 nario of, for example, mobility patterns.
53 R2
54 Sure, in Denmark, people working on epidemiological models to my knowledge or during the
55 time I was a part of this collaboration did not have the access to the raw data at all. So I don't
56 know how this data is collected. What we did have access to was mobility matrices I would
57 say. So we had some levels of aggregation, in our case there was either by... It is a little bit
58 more complicated like this: you can think of zip codes or municipalities. Some raw level of
59 aggregation and what we had access to was how many cell phones moved from one area or
60 box or whatever to another on a given day or a given time window. And I think there might
61 have been three windows per day, something like that. So how was the actual data collected, I
62 don't know. We were in touch with people from the companies that were tasked by the com-
63 panies to collaborate with the authorities I guess and I think I participated in a meeting like
64 this once but the other people did it several times. And one would discuss back and forth what
65 kind of stuff we would like to put in our models and then we discussed how this could be
66 done in a privacy preserving way. And the companies would handle everything and actually
67 do the collection of the data. We would just give them matrices.
68 NG
69 Okay, yes yes. Intrestering! Moving on the next question. Big data has been used within pub-
70 lic health for a long time. How do you think the ability to coordinate data and to be able to an-
71 alyse data with modern tools has improved the work with data?
72 R2
73 How has it improved work with data? So you mean how it has improved public health?
74 NG
75 Or how the modern tools have improved the work with data so simplifying the processes.
76 R2
77 Aha, so you are asking about the development of time and the tools people can actually use to
78 handle the data.
79 NG
80 Yes, exactly!
81 R2
82 Oh, okay. It is a good question! I have not been working with this for so long. I mean, I would
83 expect that one has really changed is not the tools more like the scalability data is collected
84 and saved time. To be honest, in my case the tools that I have been using have been around
85 for a long time. So I think it is more processing power and hard drive capacity.
86 NG
87 Yeah, but it is interesting that you mentioned that you used the samt tools that have been
88 around for a long time.
89 R2
90 At least that is my impression, I usually program in Python and other people use SaaS and R
91 and we use standard packages. It would be interesting to hear other peoples take on it but as I
92 mentioned I am 30 years old and graduated from my phd five years ago so not much has hap-
93 pened during my time.

94 NG
95 Good answer and what did the working method look like before you had access to mobile
96 data? This is sort of a similar question since you have just worked with mobile data for a lim-
97 ited time. But do you have a take on that?
98 R2
99 So epidemiological models were made before mobile data and after?
100 NG
101 Yes, exactly
102 R2
103 I think... So what we used telecom data for or our use of telecom data was in the form of
104 these mobility matrices. How many people moved from one municipality to another and you
105 could have other data sources that told you more or less the same thing. We had similar data
106 from Facebook where they tracked or people let them track their location and they would
107 know how many people of their users at least walked from one municipality or one region. So
108 at least in my case what the telecom data did was that it provided an incredibly complete data
109 set, everyone has a cellphone as opposed to Facebook accounts. So with the telecom data you
110 would like... The dataset of movement was incredibly complete. I don't think the people I
111 talked with in the group had never seen anything like this before. So the potential is huge. An-
112 other thing one of course has to consider is how to use that kind of data in a way that respects
113 the privacy of the users and that puts some limitations on what you want to do with the data. I
114 would stress that completeness is the most important thing. If you were willing to go down to
115 a list aggregate level you could perhaps get some other interesting effects. But completeness
116 is the most important thing.
117 NG
118 Yes, we totally agree. Do you think mobile data has been used during COVID-19 to decrease
119 the transmission of infectious spread?
120 R2
121 Do I think so, I think in the beginning of the pandemic there was a lot of hope that one could
122 do this. We certainly had a lot of hope. But we quickly found that... So in the beginning, what
123 we were thinking was if cell phone data without thinking of privacy you might think that we
124 can know who talked with each other and who meets up and that it would be really good to
125 know when modelling a disease. First of all, you can not redo this because you do not have
126 access to the GPS-signals. I think Norway did use GPS-location but if we just have access to
127 the cell tower that the cell phone is connected to then you won't know who people are talking
128 to, you would have hundreds of people connected to one cell tower. So in the end what we
129 could use it for was really as I said, complete data of who moves where and then that could
130 have been great if we knew exactly who moves where. So this person lives in this municipal-
131 ity, works in another municipality but that's not even the scalability we had access to the data
132 in. We did not know who moved, we just knew how many people moved. So that made the
133 data less applicable for the modelling. So in the end, did cellphone data itself decrease the
134 number of infected. Not so sure or that was not how we were thinking of it when I was leav-
135 ing. But at least I know that some people had ideas that could have made it extremely useful
136 even when preserving privacy. And that I think is why the group asked to have access to the
137 data for a while but I am not sure whether that was implemented. There were some studies in
138 the US and from other places that published results that were really promising but I think
139 most of them had access to GPS-location and not just the aggregated data that we had access
140 to. Instead of using telecom data people were also using data from smart watches. Apparently
141 some watches had GPS or other things in them. My point is that non telecom data could give
142 some us information as well. But then again. How many people own a smartwatch. So the
143 strength with mobile data is the completeness. But I guess it is also nice to have several

144 different companies that can administer data independently of the authorities so you do not
145 end up with a lot of raw data on some government computer.

146 NG

147 Yes, that is true. The next question might be hard for you to answer but I am going to ask it
148 anyway. Which kind of decisions have been made based on mobile data?

149 R2

150 Yes, so that is a big question. So the way you make epidemiological models, you can make
151 one for one population and that is usually what one would do in the first place. Imagine think-
152 ing of the whole Swedish population as particulars on urn. Then if you want to have more de-
153 tail you can split the screen up to different municipalities or regions. You can make one
154 model of one urn for one municipality and then have the particles move between the urns and
155 you need information about how these particles move between the urns and that's where the
156 telecom data usually would come in. So I know that the Danish models moved to this kind of
157 multi urn framework and it is called a metapopulation model. So in that way data if it was
158 used in those models the data would be extremely important for the modeling effort. You
159 could use other sorts of data but the model would not be as accurate. So in that way, every de-
160 cision that was based on the models would be partially based on the telecom data. I remember
161 that we were looking a lot at how people were moving around during holidays in the coun-
162 try. So one worry was, if there will be a lockdown, will all the people... So if there is a lock-
163 down one of the reasons why you make a lockdown is to stop people from spreading the dis-
164 ease to each other and one worry was that people living in the city will move to their cabins or
165 summerhouses out of the country and thereby actually spread the from disease centres to less
166 disease impacted areas. So the telecom data of course gave us an analysis if this were happen-
167 ing. We actually saw that people were moving from the centre to the outskirts of the country
168 during the lockdown and the holidays. And if we would have seen severe spikes in infection
169 after that the telecom data might have been influencing mitigation efforts to avoid that situa-
170 tion again. But I do not think that spyke arrows in the area even if people were doing this kind
171 of travelling. So in this case the telecom data might influence the absence of a decision.

172 NG

173 Thank you for the answer and one last question regarding this topic. Many countries imple-
174 mented contact tracing applications during COVID-19. Do you know if there were any
175 thoughts regarding implementing one in Denmark?

176 R2

177 There was a huge contact tracing in Denmark. Both based on interviews, who did you have
178 contact with. So if you got infected in Denmark you had to call a hotline and they would tell
179 you what to do, like isolate and they would also ask you who you were in touch with and then
180 they would call these people if they were close contact. That was one way, the other was that
181 the Danish Health Authorities took an initiative to develop an app which used this Google,
182 Apple bluetooth contact tracing. So the telecom data was not involved in the contact tracing
183 but the cell phone was. It is actually called Smittestopp, at least Apple's app.

184 NG

185 Thank you, I will write it down.

186 SL

187 Now moving on the next section with expectations of the collaboration between telecom com-
188 panies and the public health agencies. So I am not sure if this question might be hard to an-
189 swer but do you know what the initial goals regarding the collaboration between telecom
190 companies and public health agencies in Denmark were?

191 R2

192 I did not have anything to do with that but I expect that there were some expectations both to
193 look at the mobility of the Danes, how might the disease spread and incorporate the Danes into
194 the models.
195 SL
196 The next question is if the goals have been achieved? Which also might be hard to answer.
197 R2
198 I know that I made a model that could take in the cell phone data and implement it into a
199 model. I did communicate to the people who were actually making the models how it should
200 be done and I hope that was used eventually. But I know that the mobility pattern was tracked
201 often, how many people were moving around. Although the exact goals I do not know about.
202 SL
203 Do you know if there were any goals added or changed during your time working with
204 COVID-19
205 R2
206 No, not exactly but I do know that many people started to think alot about privacy and it was
207 extremely important to do this in a proper manner.
208 NG
209 Interesting. Now we will go through risks and challenges with mobile data. What do you
210 think are the biggest challenges to be able to work with mobile data during COVID-19?
211 R2
212 The biggest challenge was to find out exactly what is the best use of the data. There is a prin-
213 ciple of proportionality that I am sure they use in Sweden as well. So if you get access to
214 some data it has to be important enough compared to how intrusive it is to people. So finding
215 the balance there and to find how it could be most useful. But I don't know, the very big chal-
216 lenge is to be sure that privacy is okay.
217 NG
218 Moving on to the next question, how has the quality assurance process been to present mobile
219 data with as high quality as possible?
220 R2
221 That I do not know, that was the private companies that took care of.
222 NG
223 Another more wide question, there is a mistrust in society regarding how organizations and
224 authorities handle personal data. How do you think the mistrust can be decreased? Are there
225 any key factors for maintaining a society's trust regarding personal data?
226 R2
227 Ehhm, I am not sure if there is a general mistrust, I know that groups are thinking a lot about
228 this but in the general population I am not sure, perhaps there is. My instinct says openness is
229 important and that you should be completely open about what the data is used for and why it
230 is used in that way and how the data is kept. Personally, being a part of this effort strength-
231 ened my trust in the authorities a lot. So things were going very fast in the beginning and you
232 might have expected that people were not careful or not as careful as they should have been.
233 Even if this was the case, I am not sure whether it was, I am sure that such cases were handled
234 very quickly afterwards. So I was very impressed how serious the authorities took the privacy
235 aspect. I remember in Denmark that separate systems were set up to handle sensitive data and
236 in the beginning we had to use really involved ways of accessing data since it had to be very
237 secure but they actually managed to set up a separate and nice system in the end which I think
238 worked. I did not use this system but was acting really quickly and thinking about privacy all
239 the time.
240 NG

241 Yes, interesting and this sort of leads us into the next question. Do you think flaws exist with
242 mobile data's anonymity or do you think that the Danish government made these systems so
243 secure that the threat is close to zero or that data could be accessed by others?
244 R2
245 I can tell you that all we had access to was this very very aggregated data, not even tracking
246 individuals was made, just the flow between municipalities. So this extremely aggregated data
247 I can not see any... And I know there would have been thresholds to insure even if only 30
248 persons or something were in one municipality and were moving then those counts would
249 have been masked or not included to keep the anonymity. So in the data that public health au-
250 thorities had access to, I am pretty confident that there is not any risk there. But I know that
251 other countries based their efforts in which I said collecting GPS-signals instead of more ag-
252 gregated counts and in those countries it might have been a problem.
253 NG
254 Interesting and now we are moving on to the last topic which is possibilities and the future of
255 mobile data.
256 SL
257 Yes. If one speculatively does not need to anonymise data as in some less democratic coun-
258 tries do you think we could have been slowing down the infection spread more effectively
259 with that kind of data?
260 R2
261 It depends on what kind of data you have access to. Lets say, the most precise data I could
262 think of would be the complete GPS-location from all individuals in a country. Would that
263 have been good to put into the models, I am sure. So I am sure that if you had time to think
264 out how to do the models with that kind of data you might be able to make the models more
265 precise and accurate. The question is how much of an edge that would give you. So the prob-
266 lem is that you are in touch with so many people during one day, especially if there is not a
267 lockdown. The important thing is not just who and how long you are around some people or
268 for how long you are around with that person. It is also what you are doing while you are
269 around that individual. So if you are at a party and you are playing beer pong. Even if you are
270 just around that person for 30 seconds, if you are drinking from a glass that seven people are
271 drinking from then that behavior will be so irresponsible since that contact would be really
272 important even if it was only 30 seconds long. So I think there are a lot of things like this that
273 complicate the picture a lot. So I do not think it would be worth it to collect extremely precise
274 location data. That is my feeling.
275 SL
276 Yes, good point and do you think accuracy and precision of mobile data will improve in the
277 future?
278 R2
279 Yes, I am sure. I also hope that people will be thinking more about privacy and access to the
280 data will be limited more than it is now. But that is of a general level, I think that the data the
281 public health authorities had was reasonable. People were thinking about privacy there so
282 don't take my statement to mean that the health authorities should have had less data access.
283 That is not what I mean.
284 SL
285 No, of course. Last question: what do you think there is for future use of mobile data in order
286 to slow down the process of infectious diseases?
287 R2
288 Yes, that is a good question. I know that there are many efforts to track movement of people
289 in areas affected by catastrophes. And my feeling is that that will be a more important use in-
290 stead of tracking a whole population during lockdown. Because the risk of a pandemic this

291 big won't happen as often as major catastrophes. I hope and expect that more research will be
292 done in the coming years where people will think hard about what data was available during
293 this pandemic and then figure out how to use this data in the future. So as I said people were
294 figuring out how to use the data while the pandemic was evolving and hopefully the
295 world will be more prepared next time.

296 NG

297 Yes, indeed. But good answers and thank you. That was it for the interview. Do you have any
298 questions for us or something that you want to discuss further?

299 R2

300 No, but I wish you a lot of luck and I hope that you have fun while writing your thesis.

301 NG

302 Thank you and once again thank you for your time and we will send over the transcript when
303 it is transcribed.

1 Appendix 7 – Transkribering intervju 3

2 SL = Simone Lewenhaupt

3 NG = Nils Gunnarson

4 R3 = Respondent 3

5

6 SL

7 Toppen, då börjar jag med att fråga hur gammal du är?

8 R3

9 43 år.

10 SL

11 Yes, och vilket kön identifierar du dig som?

12 R3

13 Kvinnligt.

14 SL

15 Yes, vilket företag jobbar du på och vill du kort beskriva din yrkesroll?

16 R3

17 Mm, jag jobbar på Telia Company och där jobbar jag i en... På en avdelning som heter
18 Division X där vi tittar mer på framtida lösningar som inte är Telias core idag. Där sitter jag
19 som head of product strategy och new product, vilket betyder att jag tittar på nya produkter
20 som inte ännu har skalats upp och som vi inte ens har bestämt om vi ska skala upp eller inte
21 utan möjligheter att testa grejer som kanske ligger lite längre fram och sen förstås även så är
22 jag ansvarig för produkt strategin på Division X.

23 SL

24 Mm, eh toppen och hur länge har du arbetat med mobildata kopplat till COVID-19?

25 R3

26 Sen januari 2019. (Hon menade avdelning och inte COVID-19 men då har hon arbetat med
27 det sen start drar vi som slutsats.)

28 SL

29 Mm så typ sen starten då helt enkelt.

30 R3

31 Jaa typ sen starten, jag började mm eller ja, jag har varit med nästan sen början men inte allra
32 första början men ganska tidigt ja.

33 SL

34 Yes, men då ska vi gå in på lite mer användning och nytta av mobildata.

35 NG

36 Eehm yes, vikten av att transformera data till insikt är något som ofta benämns av Telia. Hur
37 skulle du kortfattat beskriva processen från data till insikt?

38 R3

39 Ehm, jo men dels så handlar det om att se till att den datan som man samlar in, att insikten
40 som man får av den är representativ så när man pratar om mobildata så har ju vi data på hur
41 Telias abonnenter rör sig ehm men då anonymt. Men, men det är ju inte jätteintressant för
42 världen att veta hur Telias abonnenter rör sig utan då måste vi ju se till att extrapolera den
43 datan för att få den att motsvara en population, i det här fallet då Sverige population. Och det

44 är ju själva, det är ju ett jobb som har tagit lång tid att utveckla, hur ska vi extrapolera för att
45 få det att vara representativt för befolkningen och där är ju förstås vår marknadsandel och
46 också att vi har me konsumentrelaterade varumärken såsom Halebop gör att vi kan göra ett
47 bra jobb där men, men det är ju de är ju en viktig del men när det gäller att skapa insikter
48 måste man ju veta vilket problem man ska lösa. Det är ju, data, ofta när man jobbar med
49 datatjänster finns det ju väldigt mycket svar i datan men frågan är ju vad frågan verkligen är,
50 vilken frågan är som är tillräckligt viktig att lösa och är det så att det här datasetet ensamt kan
51 lösa den frågan eller är det så att man måste koppla den här datan med andra datakällor till
52 exempel för att komma till problemlösningen. Så det är viktigt att veta, för att få ett värde
53 utav data så måste man veta vilken problemställning man har.

54 NG

55 Aa, intressant, ehm och nästa fråga då. Big data har använts länge inom folkhälsan. På vilket
56 sätt har möjligheten att samordna data och kunna analysera datan med moderna verktyg
57 förenklat erat arbete?

58 R3

59 Alltså vi har ju aldrig haft som syfte att kunna minska spridning av olika sjukdomar när vi
60 tittar på mobildata. Men det finns ju naturligtvis en koppling mellan hur folk rör sig och smitt
61 ehh och spridning av smittsamma sjukdomar. Ehm och då handlar det ju mycket om att kunna
62 använda den här typen av data i modeller som kan förutspå hur en smittspridning kommer att
63 se ut. Där tror jag framförallt att från våran del, det är ju som sagt folkhälsomyndigheten som
64 bygger den typen av modeller och där vi ju att mobildata kan spela en viktig del även för att
65 ett mobildataset som vårt är ju väldigt vänligt för modellering för att när väl en population
66 som har stabiliserat sig om man tittar på hur det såg ut innan pandemin så har ju vi ett
67 beteende som är väldigt inrutat och vi är ju vanedjur vi gör väldigt mycket samma sak. Om
68 man tittar på en måndag i november och sen tittar man på en måndag tre veckor senare så ser
69 det ganska likt ut. Sen är det ju förstås att vi har säsongsskillnader och andra saker som
70 påverkar vårt beteende men en vanlig pendlarpndag den ser ganska lik ut om allting annat är
71 lika. Det handlar ju om att vi ska kunna leverera våran data i strukturerade former så att den
72 kan användas i den typen av modellering.

73 NG

74 Mm, intressant och nästa fråga som kanske är lite mer åt Folkhälsomyndigheten då men vi ska
75 se i alla fall. Hur har arbetssättet sett ut inna man hade tillgång till mobildata just då i ehh..

76 R3

77 Ja det har jag nog svårt att svara på, det vet inte jag.

78 SL & NG

79 Eh ja haha.

80 R3

81 Man tittar ju ehh nej, det, man har väl inte kvantifierat mobil.. Eh rörelser på det sättet som
82 mobildata kan göra. Men vilken typ av annan data dem använder det kan jag inte svara på.

83 NG

84 Neej, men då hoppar vi till nästa fråga. Eh, anser du att mobiloperatörsdata har använts under
85 COVID-19 för att bromsa smittspridningen.

86 R3

87 Ja, det tycker jag definitivt att det har gjort. Ehm det är ju, den stora frågeställningen som vi
88 hade var ju framförallt hur beter sig folket efter olika typer av rekommendationer och
89 restriktioner. Behöver vi, behöver folkhälsomyndigheten som myndighet sätta i starkare
90 restriktioner för att populationens rörelse inte har förändrats. Den typ av frågor kan ju vi svara
91 på. Så definitivt.

92 NG

- 93 Aa men kanon och det leder oss in lite på nästa fråga också. Vilka slags beslut har man kunnat
94 ta eller beslutat om baserat på Telias mobildata.
- 95 R3
- 96 Vi tittar ju på historisk data så vi kan ju titta efter hur, folkhälsomyndigheten ger en
97 rekommendation gällande rörelser, det kan vara till exempel inför påsk 2020 så hade vi en
98 restriktion kring att man inte skulle resa mer än 2 timmar ifrån hemmet. Påsk är ju en stor
99 rörelsehög där många reser längre än så och där kunde vi ju då efter analysera hur svenskarna
100 reste över påsk just. Då kunde dem se om svenskarna hade lyssnat på rekommendationerna
101 och det resultatet av det var ju väldigt, väldigt visande på att ja svenskarna ändra verkligen sitt
102 beteende efter att vi då hade fått folkhälsomyndighetens rekommendationer på oss att inte
103 resa.
- 104 NG
- 105 Äj men intressant eh och nästa fråga då. Många länder
106 implementerade kontaktspåringsapplikationer under COVID-19. Var det någonting som
107 fanns på kartan för Telia och Folkhälsomyndighetens samarbete att utforma en sån
108 kontaktspåringsapplikation?
- 109 R3
- 110 Nej, det tittade vi inte på tillsammans, vi hade ju mycket, alltså det var ju mycket diskussion i
111 på marknaden om de här apparna och hur de skulle fungera och sådär ehh och mellan oss och
112 folkhälsomyndigheten så handlade det om just Crowd insights som är en anonym tjänst där vi
113 tittar på anonym data. Den kan ju användas för den gjorde olika inte tillsammans bara med
114 folkhälsomyndigheten men vi gjorde ju också ett samarbete med Ericsson och Sahlgrenska
115 där man använde samma typ av data för att förutspå hur mycket inläggningar man ska ha på,
116 man kommer att ha på sjukhusen. Så det finns många use case förstås men den här datan är ju
117 inte consent baserad utan det är en anonym tjänst det vill säga att det behöver inte finnas
118 någon opt out för det finns inga individer i det här datasetet som vi analyserar. Det gör ju
119 också att man måste ehm, man måste svara på andra frågor för att kontaktspårning då måste
120 du titta på individer och det är ju alltså inte det som Crowd insights gör. Då måste man titta på
121 individ och då måste man ha tillåtelse att göra det från individen och då är det väldigt svårt att
122 få ett representativt dataset.
- 123 NG
- 124 Mm, aa men jätteutförligt och bra svar och intressant det där med sjukhusinläggningar också.
- 125 R3
- 126 Mm, det kan ni, det finns en artikel om det också det kan ni läsa mer om om ni vill.
- 127 NG
- 128 Yes.
- 129 SL
- 130 Aa men det är alltså Telia har också alltså jobbat med att förutspå sjukhusinläggningar och
131 sånt?
- 132 R3
- 133 Vi har jobbat i ett projekt som har gjort det så vi jobbade tillsammans bland annat då med
134 Ericsson och Sahlgrenska sjukhuset för att titta på, det var ju också som ett innovationsprojekt
135 men som fick en väldigt god, ett gott resultat. Men sen har Telia inte tittat på
136 smittspåringsappar av olika slag.
- 137 SL
- 138 Mm, yes.
- 139 NG
- 140 Men då går vi över lite på förväntningar på samarbetet mellan Folkhälsomyndigheten och
141 Telia och lite kontra hur utfallet blev.
- 142 SL

143 Ehh ja precis, så vad fanns det för mål initialt gällande samarbetet mellan
144 Folkhälsomyndigheten och Telia.
145 R3
146 Jo målet var ju ganska, eh det var väldigt enkelt. Det var ju kan vi använda mobildata för att
147 hjälpa till att bromsa smittspridningen av COVID-19.
148 SL
149 Ehh toppen och vad har det målet uppnåtts skulle du säga?
150 R3
151 Absolut det har det gjort och det var ju mycket på grund av just att vi hade en färdig tjänst
152 som baseras på mobilitetsdata där vi redan hade spenderat aa men nästan två år på att
153 verkligen göra det här på rätt sätt och se till att vi har en robust anonymiseringsprocess och att
154 vi har, vi vet att vi är GDPR-compliant. Vi hade också det här väl dokumenterat det var ju när
155 vi hamnade där i slutet av april, i början av ap... maj, i slutet av maj, början av april 2020. Vi
156 var ju inte de enda som kontaktade Folkhälsomyndigheten och ville hjälpa till utan det var ju,
157 de hade ju en uppsjö av olika bolag och start-ups och stora bolag som ville hjälpa till. Men vi
158 kunde väl snabbt visa på att vi kan se en skillnad i rörelser hos det svenska folket. Dem hade
159 ju i sin tur ett problem med att kvantifiera, vad är effekterna av rekommendationerna som vi
160 har satt på när det gäller att folk inte ska träffa andra, att vi inte ska umgås med andra
161 människor. Så det var ju mycket reaktioner i media som sa att aa men ehmm uterestaurangerna
162 är fulla på medborgarplatsen. Stefan Löfven har sagt att vi inte ska träffas i grupp och jag såg
163 att det var fullt på parkeringen på Gekås. Alltså det fanns ju massor med olika rapporter om
164 olika saker där folk visst samlades fortfarande och så kan det ju vara men det är ju inte
165 kvantitativa insikter utan det är ju väldigt här och nu vad jag kan se med mina två ögon och
166 det här, potentialen i vår data är ju just den att du får en kvantitativ analys på en befolkning
167 snarare än individer. Ehm och det var ju därför Folkhälsomyndigheten väldigt snabbt såg
168 potentialen i vår data nog för att den inte var paketerad för att dem skulle kunna ta till sig
169 insikterna på en gång utan det var ju ett jobb som behövdes göras där tillsammans för att vi
170 skulle kunna paketera vår tjänst så att den passade just deras frågeställning. Det handlar ju
171 om precis det jag sa i början om man ska göra insikter så handlar det om att man ska kunna
172 göra någonting som, man måste svara på de rätta frågorna. Ehm och det var ju precis det som
173 vi lyckades göra då men, men i botten låg ju verkligen att vi hade en tjänst i bakgrunden som
174 redan var utvecklad, som inte slängdes ihop över en helg för att vi såg att det kanske finns
175 potential utan den var ju redan färdig att börja användas. Den skulle justeras liksom paketeras.
176 SL
177 Aa och då skulle jag också kolla bara var det några mål som tillkom eller ändrades under
178 samarbetets gång?
179 R3
180 Nej men vi hade ju en väldigt bred ingång, vi visste inte riktigt på vilka sätt som
181 mobilitetsdatan skulle kunna användas och Crowd insights finns ju i olika typer av former så
182 vi, det undersöktes ju på ganska bred front hur vi kunde hjälpa till. Det handlade både om att
183 vi tittade på effekten av rekommendationer, det som dem använde väldigt, väldigt aktivt och
184 fortfarande använder idag. Sen var det ju även att vi testade på hur vi skulle kunna använda
185 mobilitetsdatan i deras algoritmer där de tittade, där de försökte prediktera de här olika
186 scenarierna som man ser till exempel på TV. Där var det väl så att de har dem ju möjlighet att
187 göra fortfarande, just då så var det inte aktuellt för att man kan inte lägga till ett helt nytt
188 dataset i en sån algoritm i mitten av en pandemi, det blir väldigt svårt om man inte har gjort
189 det. Nu är det ju frågan om dem har det till nästa gång, det får du ju fråga dem om, det har jag
190 ingen aning om. Men det var väl det, det såg dem att det man har liksom inte möjlighet att
191 bara addera ett nytt dataset in i en algortim som ska ge aa men dem litar ju väldigt mycket på
192 dem här olika scenarierna som dem modellerar så jag vet att dem tittade på det men där får du

193 fråga dem exakt hur utfallet blev men framförallt så handlade det om att kunna paketera
194 Crowd insights så att de lätt kunde få svaren på frågorna angående just rekommendationer och
195 folkmassors rörelse det vill säga grupper av människors rörelse. Så det lyckades vi väldigt,
196 väldigt bra med under projektet.

197 NG

198 Aa men kanon jättebra svar. Nu tänkte jag att vi ska gå över lite på risker och utmaningar med
199 mobildata då. Vilka är de största utmaningarna för att kunna använda mobildata under
200 COVID-19 anser du?

201 R3

202 Jo men, det är väl ett par och från början måste man ju också, riskerna om man väljer att
203 använda consentdata till exempel, jo men den är ju samma som riskerna alltid är när man
204 använder consent data, det vill säga att den människan som ger consent inte vet vad den ger
205 consent till. Att den kanske är blind och döv och den kanske inte har förstått vad consent
206 innebär och det är ju alltid en risk och då använder du en data som egentligen inte är
207 consented. När det gäller anonym data som vi använder oss av så är ju det integritets, den
208 delen är ju täckt men fortfarande då har du ju ett annat problem att du kan inte vara lika
209 granulär. Du kan inte vara lika noggrann i dina insikter, du måste svara på andra typer av frågor
210 och sen har vi ju också, vi kan ju titta på beteenden det ser ut som folk åker till jobbet i den
211 här resan, de har ett beteende som vi kan anta är att man pendlar till en arbetsplats, men vi vet
212 ju inte vad de gör vid sin destination. Det är en anonym analys, vi har ingen aning om man är
213 ute och shoppar eller om pendlar eller om man... så det är klart att det finns också, det är
214 alltid så när man gör den här typen av analyser att man måste anta och det är statistik det
215 betyder att jag antar, det finns alltid en viss osäkerhet i det men i det här fallen var det ju inte.
216 Vi gjorde en del mer specifika analyser kring så här jobba hemma analys till exempel, vi
217 följde hur, hur rörelsemönstrena ändrades från att ha ett jobba hemma beteende till att börja
218 jobba igen och vi såg en väldigt tydlig korrelation med rekommendationer och större bolag
219 som öppnade upp igen efter pandemin. Så vi kunde ju följa den och det handlar ju också om
220 att vi har ett så väldigt stort dataset, vi har ju liksom i våra marknader så har vi 16 miljarder
221 datapunkter varje dag som analyseras så vi har ju så väldigt mycket data att titta på så därför
222 är ju den risken. Alltså vi tittar ju på trender hur folk rör sig och så länge vi håller oss på den
223 nivån så blir insikterna väldigt tydliga och vi kan lita på att dem motsvarar verkligheten.

224 NG

225 Aa men jättebra, tack sen en fråga då hur har kvalitetsäkringsprocessen gått till för att
226 presentera data med så hög kvalitet som möjligt?

227 R3

228 Ja, hmm en jättebra fråga för att innan pandemin så jobbade vi också med kunder och den här
229 typen av insikter men då hade vi ju. (Teams bröts). ...så levererade vi också insikter när det
230 gäller hur folk rör sig. Då var det ju månatliga rapporter egentligen där vi hade manuell, alltså
231 en dataanalytiker som satt och kvalitetssäkrade datan innan vi skickade iväg den. Det funkade
232 inte när man är i mitten på en pandemi och vi har ju en fördröjning på ungefär 36-48 timmar
233 när det gäller anonymiseringsprocessen och så där vi tittar ju alltid på data som är minst så
234 gammal. Men behovet var ju här att vi skulle uppdatera i våra dashboards dagligen och då
235 funkar det inte, och då vi dessutom hade hundratals kunder som var jätteintresserade av att
236 följa den här, dem här insikterna så kunde vi inte ha folk som sitter och gör det här manuellt.
237 Vi fick ju skala upp vår datatjänst, det vill säga vi var ju tvungna att bygga automatik kring
238 datakvalitet och även då vi byggde en dashboard där vi kunde ratea dagarna där vi kunde säga
239 om det var bra kvalitet för det datasetet vi har det är ju inte till för att följa mobiler utan det är
240 ju till för att upprätthålla ett nätverk för att vi ska kunna ringa. Så det finns ju du vet, det är
241 uppgraderingar som ska göras här och där och det var ju inte så att vi hade prioritet att säga nå
242 men ni kan inte uppdatera nu för det är påsk och vi måste ha de här insikterna. Dem hade

243 kunnat liksom stänga av bara för att liksom vadå det här ligger i våran road map vi ska
244 uppgradera den här, dem här servrarna ska ner nu eller vad det kan vara. Där ändrade ju det
245 sig förstås eftersom vi, vi var ju viktiga för samhället just då så då fick vi ju förstås säga till
246 om att nej Telia vi ska vänta tre veckor med att uppgradera de här servrarna för att vi måste ha
247 bra kvalitet på data under till exempel en helg som kristi himmelfärd eller påsk eller
248 midsommar. Så men, men vi automatiserade de här processerna för att då också kunna vara
249 transparenta med våra kunder kring när datat inte är lika noggrant, när vi då vet att här finns
250 det större felmarginaler så. Sen är det ju den här typen av data man, det är ju en trend man
251 tittar efter och jag vet att Folkhälsomyndigheten ja men dem skulle aldrig basera någonting på
252 mindre än en trend som håller sig i tre dagar så dem tittar liksom, dem tittar framförallt
253 veckovis på den här datan. Så att en dålig dag det spelar liksom ingen roll för att man tittar
254 liksom på hur det ser ut över flera dagar. Så att jätteviktigt med datakvalitet och vi känner ju
255 också att det är jätteviktigt att vi är så transparenta som vi kan kring kvaliteten på datan för att
256 data är aldrig 100%, det lär man sig ganska snabbt när man jobbar inom den här inom det
257 datafield, att det är frågan om hur mycket kan vi lita på datan och det är viktigt att vi är
258 transparenta med våra kunder så just nu så har vi den informationen i alla våra dashboards och
259 även i våra datatjänster att kunden vet dem dagar som är sämre kvalitet.

260 NG

261 Aa men jätteintressant och det här kanske vi har berört lite lätt men jag ställer frågan så ser vi.
262 Det finns en misstro i samhället gällande hur verksamheter och myndigheter hanterar
263 personlig data. Är det något ni aktivt arbetar med för att bibehålla samhällets förtroende i ert
264 arbete med mobildata?

265 R3

266 Absolut, jag skulle säga att det är någonting som är absolut viktigast för Telia, Telia hade
267 aldrig låtit oss sälja den här tjänsten om vi inte kunde stå upp och väldigt raka i ryggen prata
268 kring integritet och hur vi bibehåller våra kunders integritet och det är ju precis det är ju
269 liksom prio ett i allt vi gör. När det gäller den här typen av insikter, när vi satte igång för två
270 år sen då var det väldigt många, det finns väldigt många sådana spelare på marknaden
271 fortfarande som använder liksom app-data med GPS:er som trackar locations utan att
272 människan riktigt vet att man är trackad och så vidare. Ganska så här shady data collection. Vi
273 vill ju stå upp och säga att vi gör det här på rätt sätt, och det gör vi verkligen. Vi får ju göra
274 avkall då på till exempel granualiet vi kan inte se lika noggrant, våra rutor är större på våra
275 kartor än vad dem kan vara när man använder andra typer av datainsamling. Men vi kan också
276 stå upp och säga att vi har ett anonymt dataset. Det finns ingen opt out för det finns inga
277 individer i så det finns ingenting att opta out från så att det. Det här är någonting som är
278 jätteviktigt för Telia och också att vi kan vara så transparenta som möjligt och kunna berätta
279 på en hög nivå hur det faktiskt fungerar.

280 NG

281 Så då är det det aktiva arbetet handlar då om så stor transparens som möjligt men även att aa
282 undervisa eller förklara hur era tjänster används då till folket.

283 R3

284 Absolut, vi ska kunna ha förklarbara och transparenta tjänster och sen jobbar vi ju konstant
285 med hur vi ska kunna med ny teknologi och privacy preserving techniques kunna få ännu
286 mera insikter vi är ju inte nöjda där vi är idag naturligtvis men vi gör det aldrig på bekostnad
287 av våra kunders integritet.

288 NG

289 Aa men jättebra svar, nästa fråga då. Finns det brister med mobildatans anonymitet idag.
290 Finns det några hot som skulle kunna påverka eller avslöja den anonyma datan?

291 R3 08:40

292 När det gäller våran tjänst så har vi ju gjort såna attemppts, liksom att försöka programmera
293 bakåt och sådär, men där har dem ju misslyckats. Jag ser ju att jag menar, hur långt är ett
294 snöre, det är väl klart att någonting som du har insikter i aldrig kan vara helt anonymt men
295 GDPR säger ju “med rimliga åtgärder” och där är jag ju helt säker på att vårt dataset är helt
296 anonymt. Sen så du vet om du pratar med en data scientist så kommer dem säga att det finns
297 ingenting som är helt anonymt om det inte är helt noll, det är ingenting och då får du ju inga
298 insikter ur det heller så det är ju liksom en skala definitivt. Men vi har lagt väldigt mycket
299 fokus på just den delen av våran tjänst så jag känner mig superbekvämt med hur den opererar
300 idag.
301 NG
302 Jättebra svar.
303 SL
304 Yes då är det sista delen då med lite möjligheter och framtid. Så om man spekulativt inte hade
305 behövt anonymisera datan som i mindre demokratiska länder, hade man då mer effektivt
306 kunna bromsa smittspridningen med datan tror du?
307 R3
308 Svår fråga, alltså inte med sådana metoder som vi vill, vi vill ju inte ha ett stängt samhälle
309 men om man tittar på vad som hände i Kina så var det ju så att då vet dem var du är
310 någonstans och om du har husarrest blir du ju bestraffad om du bryter mot den. Eehm det
311 hade man ju asså rent, vi hade ju inte kunnat gjort de på befolkningen i alla fall efter som som
312 sagt vi har inte data på alla. Men asså om man skulle, det här handlar ju om staten i så fall
313 skulle kräva, asså det här kommer ju aldrig att hända, ni ser ju att det här är väldigt mycket
314 om och om och om. Men om man tittar på en sån stat som Kina dem gjorde ju precis det, så
315 det är klart att om man vill behandla sin befolkning på det sättet så det är klart, det handlar ju
316 om att om du inte träffar någon så smittar du ingen. Det är ju ganska enkelt om man kan sära
317 på alla kontaktytor så kommer du att bromsa smittspridning. Sen om det är etiskt rätta sättet
318 att göra det, aa men det har vi säkert alla tre en annan åsikt kring. Men sen är det ju så här,
319 mobildata har ju, våran mobildata har inte, vi vet ju inte exakt var folk befinner sig, vi vet
320 vilken mast dem är, eller ett visst imsi som inte är ett riktigt imsi utan det är en anonym
321 nyckel, vi vet var den är någonstans, vilken mast. Men det gör ju att jag menar vi har ju en
322 accuracy som bäst på +- 200 meter eller någonting sånt så att det är ju även om vi hade haft
323 helt whitelistade data, om vi såg allting hade vi fortfarande inte kunnat säga exakt var folk är
324 någonstans för det är inte det, det här datasetet gör inte det.
325 SL
326 Nej. Ehh okej och tror du att exakthet och precision av mobildata kommer förbättras i
327 framtiden?
328 R3
329 Absolut.
330 SL
331 Mm.
332 R3
333 Det tror jag, det kan jag nästan sia in i framtiden, jo det kommer den att göra. Det finns redan
334 om man tittar på nya teknologier som kommer när det gäller 5G och andra så finns det andra
335 möjligheter till positionering och jag tror ju att så länge vi är väldigt transparenta kring vad vi,
336 vad är nyttan med att vi positionerar dig. Så tror jag att det är ju människans eget val, det
337 finns ju olika typer av studier på hur man till exempel kan använda positionering för att göra
338 dynamiska vägtullar. Ehm jag betalar ju mycket hellre skatt beroende på hur mycket jag kör
339 på vägen, än betalar någon schablon om jag inte kör alls, det känns ju inte rättvist. Så då
340 kanske jag skulle vara villig att dela min location när jag kör på vägen för att jag skulle kunna
341 möjliggöra så att jag skulle kunna få rättvis skatt. Så det finns ju väldigt många use case. Det

342 viktigaste här är ju att du utgår ifrån vilken vinning är det för den som ger sin data. Det kan ju
343 också vara så här, vill du ge din position till din kommun för alltså var du cyklar till jobbet
344 varje dag för att dem ska kunna, när första snön kommer prioritera var dem ska plöja så att du
345 fortfarande kan cykla. Så länge du är väldigt tydlig med vad datan ska användas till så tror jag
346 säkert att folk är villiga att avvara sin data, om man får ett värde tillbaka. Då tror jag att
347 mobildata i framtiden kommer ha en stor roll för att vi kommer kunna positionera väldigt
348 mycket bättre än vad vi kan idag.
349 SL
350 Yes, och sista frågan då. Vad anser du att det finns för framtida användningsområden med
351 mobildata i syfte att bromsa spridning av smittsamma sjukdomar?
352 R3
353 Ja men det handlar ju precis om samma sak egentligen, det är ju ett annat use case för att veta
354 hur folk rör sig anonymt kan du titta på större samhällen, du kan titta på större grupper av
355 folk. Men du kan ju också välja då att aa men jag vill att eh, dela med mig av min location,
356 min, ni får titta på mig, och jag kommer också rapportera om jag är smittad för att skydda min
357 omgivning eller vad det kan vara för någonting. Så att både i anonym och consentad form så
358 tror jag definitivt att man kan använda det för att bromsa smittspridning men också titta på, i
359 sina modeller hur kommer den här, vad behöver vi att lägga in för åtgärder för hur rör sig
360 människor, om man vet det i mer genreller termer så har man också en, är man väldigt mycket
361 mer förberedd på när, när nästa pandemi kommer.
362 SL
363 Aa men det var de vi hade att säga, har du någonting mer att tillägga eller fråga eller så?
364 R3
365 Nej men det tror jag inte, det är superspännande att ni har valt det här området, nu är det ju
366 väldigt aktuellt förstås med COVID-19 men du vet ju att det kommer ju att komma flera
367 pandemier så, det blir väldigt spännande eftersom jag vara så nära det här arbetet, det var ju
368 väldigt spännande att se. Aa det jag vill ju inte önska en ny pandemi, herregud men om det
369 skulle bli så är det spännande att se vad vi har lärt oss. Det var ju det världen lärde sig nu att
370 vi är inte förberedda på det här.
371 SL
372 Nej precis.
373 R3
374 Så se hur mycket vi kan, och hoppas att det tar väldigt, väldigt länge men att vi fortfarande
375 kan lära nästa, nästkommande generationer hur man inte ska göra i vissa fall.
376 NG
377 Ja verkligen, och en fråga på det, ert samarbete med Folkhälsomyndigheten kommer det
378 upprätthållas eller kommer man ha någon sådan förebyggande plan för hur man eventuellt,
379 alltså kommer det här samarbetet kunna återupprättas för att det finns...
380 R3
381 Vi är fortfarande partners. De använder våran data för aa för att titta på just hur människor, aa
382 men samhällsrörelser påverkar det som dem tittar på. Jag gissar på att dem inte bara har
383 COVID-19 som sitt fokus nu heller, så jag tror definitivt, nu vet dem, dem visste inte vad
384 mobildata var och hur man kunde arbeta med det förut och det vet dem nu. Sen får du ju prata
385 med em om dem ser potential men det tror jag alla gånger att dem gör.
386 NG
387 Mm, aa men jättebra, stort tack återigen för att du tog din tid, det har varit jätteintressant att få
388 höra ett proffs i ämnet.
389 R3

390 Aa jag vet inte hur mycket proffs jag kände mig där borta i april 2020 det var en intressant
391 arbetserfarenhet kan man säga. Men jättekul och som sagt jättekul att ni skriver om det hr och
392 gör att det sprids ännu mer. Jag tycker det är jätteroligt, super aktuellt ämne så lycka till.
393 SL
394 Aa tack så jättemycket, ha det bra.
395 NG
396 Stort tack.
397 R3
398 Ha det så bra hejdå.
399 SL
400 Hejdå.
401 NG
402 Hejdå.

1 Appendix 8 – Transkribering intervju 4

2 SL = Simone Lewenhaupt

3 NG = Nils Gunnarson

4 R4 = Respondent 4

5

6 SL

7 Då börjar jag med att fråga hur gammal du är?

8 R4

9 44 år.

10 SL

11 Och vilket kön identifierar du dig som?

12 R4

13 Manligt.

14 SL

15 Yes, vilket företag jobbar du på och vill du kort beskriva din yrkesroll?

16 R4

17 Ja, så jag och mina kollegor vi jobbar ju då på Telia eller Telia Company heter vi formellt då.
18 Vi jobbar ju med hela Telias footprint som vi kallar det då Footprint, det vill säga
19 avdelningarna vi är verksamma, Finland, Sverige, Norge, Danmark och Estland. Men det är ju
20 alltid under varumärket Telia. Så jag har när det här skedde då med Folkhälsomyndigheten
21 och under hela COVID-pandemin så ledde jag det hära affärsområdet som vi kallar Data
22 Insights och min roll har då kallats Head Of Data Insights. Affärsansvarig för den här
23 portföljen då, produktportföljen som vi kallar för Crowd Insights. Nu har vi dock
24 omorganiserats sedan 1:a april och jag ansvarar nu för ett bredare område och ansvarar nu för
25 vår totala produktportfölj här på Division X på Telia och det innefattar då Crowd Insights men
26 också hela Telias IoT, alltså Internet of Things. Så det är min nuvarande roll som kallas Head
27 of Product heter det.

28 SL

29 Aha, okej. Toppen! Hur länge har du arbetat med mobildata kopplat till COVID-19

30 R4

31 Ja, det har varit ungefär, vad blir det nu när jag började på Telia, så att det är tre och halvt år
32 som jag har jobbat med detta och byggde upp det här affärsområdet. Det var ett
33 innovationsprojekt först på Telia som hade skett några år tider med mobildata och jag kom in
34 då för att börja etablera det här som ett eget affärsområde och ta det från innovation till en
35 riktig produkt så vi kan skala upp i våra marknader.

36 NG

37 Intressant men då går vi över till användning och nytta med mobildata. Vikten av att
38 transformera data till insikt är något ofta något som nämns av Telia Crowd Insights. Hur
39 skulle du kortfattat beskriva processen från data till insikt då?

40 R4

41 Det är en väldigt, väldigt bra fråga och gäller ju inte bara vår produkt Crowd Insights men jag
42 brukar se det som så en insight eller en insikt är alltid svaret på en fråga, en eller flera frågor.

43 Och det inbegriper ju då att det måste finnas en fråga och oftast är det som är det svåra, ofta

44 när man jobbar med nya datakällor. Att förstå vilken fråga man faktiskt vill besvara, eh, och
45 vad gör jag när jag har svaret. Vad vet jag och vad jag ska göra med det, då är det en insikt
46 som då är en, eh, förlåt för att det blir mycket svengelska men då är det något som är
47 actionable. Och där är ju datat en del, såklart, i att kunna forma datan för att kunna då vara ett
48 svar på en eller flera frågor. Och för oss är det ju ofta att våra kunder vill ha
49 verksamhetsrelaterade frågor och något som kan hjälpa den befintliga verksamheten att bli
50 snabbare eller bättre eller skapa nya affärsmöjligheter. Data i sig svarar ju på ingenting, det är
51 ju en möjliggörare så att gå från data till insikt så krävs det ju såklart en relevant datamängd
52 men det krävs också en djup förståelse för problemställningen och sen inför det arbete som
53 krävs för att transformera datan kanske sammanfoga med annan typ av data, visualisering,
54 paketering av detta, förklaring av detta och i vårt fall även prissättning eftersom det är en
55 produkt och hur man vill konsumera detta, användarperspektiv eller kundperspektiv. Det är en
56 oerhört bred fråga och där många företag går bet, att man startar för mycket alldeles för
57 mycket från datahållet. Ehm, och kommer med massa svar men man har inte säkerställt att det
58 finns några frågor. Man kommer med svar som ingen vil ha.

59 NG

60 Jätteintressant svar, ehm, nästa fråga då, big data har använts länge inom folkhälsan. På vilket
61 sätt har möjligheterna för att kunna samordna data och analysera datan med moderna verktyg
62 förenklat erat arbete.

63 R4

64 Ja, nu är ju inte jag någon domänexpert inom folkhälsa men big data kan jag väl uttala mig
65 lite mer specifikt och jag skulle säga att det har varit helt avgörande, den teknikutvecklingen
66 vi har sett dem senare åren. När det pratades mycket om big data här för kanske tio år sedan,
67 nu börjar det bli lite mera mainstream då pratade man om Hadoop som ett exempel, data legs
68 men det var ganska så, väldigt väldigt tekniktungt, krävde oftast väldigt väldigt tung
69 specialiserad kompetens för att sköta servrar och göra verkligen mycket för att hålla
70 infrastrukturen vid liv och vidmakthålla den, eehm och på senare år, senaste 5 åren så har det
71 kommit mycket mer tillgängligt i dem olika molnmiljöerna från Google, Amazon och
72 Microsoft som har gjort att dem tar hand om allt det där jobbiga när har en jättestor och ser till
73 att man har processorkraft för det och tillräckligt många servrar och lagring och så vidare. Så
74 som vi jobbar med big data, vi använder ju då, för den här tjänsten använder vi Google Cloud,
75 dem har en produkt som heter Big Query som är väldigt bra på att hantera väldigt stora
76 datamängder och vi kan helt enkelt ställa frågor till det här datat som om det egentligen vore
77 en liten datamängd och då använder man som analytiker sen urminnes tider, använder man ju
78 ett frågespråk som heter SQL, Structured Query Language, ehh och när big data kom och
79 Hadoop kom så kunde man inte använda SQL utan man behövde skriva annan typ av kod så
80 man behövde specialistkompetens. Nu med funktioner som Big Query så kan analytikerna
81 som kan SQL hantera det själv så jag har någon i min organisation som kan det där lite
82 lågnivå som att hålla servrar vid liv och den här typen av plumbing som krävdes då när det
83 här kom. Så det är ju mycket mera tillgängligt, eeh och det är väl egentligen att dem här
84 jättarna, Google, Amazon och Microsoft, har gjort, ah att i princip så enkelt att man kan jobba
85 med big data som att det vore small data. Så vi kan ställa frågor till otroligt stora datamängder
86 och vi tar ju in flera miljarder datapunkter varje dag då så vi har ju en väldigt ansenlig mängd
87 data och ska man då ställa frågor till det här datat så då är ju det under huden så kanske
88 Google nyttjar 1000 servrar när vi trycker på run liksom och några sekunder senare så får vi
89 ett resultat och vi behöver inte bry oss någonting om allt det där som händer under. Det vi är
90 intresserade av är ju att kunna ställa frågor till datan.

91 NG

92 Juste, nej men intressant. Åh, om vi går vidare då. Hur har arbetssättet sett ut innan man hade
93 tillgång till mobildata och då tänker jag i att spåra rörelsemönster.

94 R4
95 Hmm, tänker du specifikt för folkhälsa usecase eller mer brett?
96 NG
97 Mer brett om du känner dig mer bekväm inom det.
98 R4
99 Oftast har man ju då haft någon typ av proxy information för rörelse som man tar till exempel
100 kollektivtrafik som är ett område vi jobbar mycket med. Dem har ju inte obegränsat med
101 pengar, dem har ju en budget den är ofta ganska stor och delvis skattefinansierad och med
102 privata medel, med ofta biljettintäkter. Då handlar det att göra så mycket som möjligt för dem
103 här pengarna och inom kollektivtrafik handlar det mycket om att rätt tidtabell, ha rätt rutter,
104 trafikera är det en bussrutt, var ska man ha busshållplatsen. Ska man bygga en tunnelbana, en
105 ännu större investering vill man verkligen veta vilken sträcka det ska vara och så vidare. Och
106 då handlar det om att förstå svaret på frågan då vart vill folk resa och hur. Det är ju frågan.
107 Det har ju funnits länge såklart men då har man ju ofta jobbat med en kvalitativ data, på sättet
108 att man har frågat och skickat ut enkäter eller titta på annan typ av proxydata. Ni har säkert
109 sett det när folk har en sån där clicker och faktiskt räknar folk på riktigt liksom och då får man
110 ju data som är... det finns ju appar som samlar in sån data också, man får ju liksom, oftast
111 olika delar av bilden. Om man tänker att hela resebehovet är som en stor målning så är det
112 som att man tittar genom en lupp eller någonting när man använder till exempel
113 enkätundersökningar eller om någon står på T-centralen här i Stockholm och räknar. Man får
114 ju en väldigt liten... en begränsad insyn. Så det är både, vad ska man säga, det som kallas
115 completeness alltså den hur komplett bilden är, den är inte speciellt representativ. Det vill
116 säga att den kanske inte representativ för hela befolkningen om man svarar med en enkät, det
117 är ofta arga personer som svarar på en enkät och alltså inte jättereformativ. Men så är det ju
118 en annan frekvens också så mobildata är ju liksom always on, dels så har man ju ett enormt, vi
119 har ju miljarder datapunkter per dag, det är den största kontinuerliga undersökningen för
120 Sverige hela tiden. Så dels får du ju en komplett bild, du har hög grad av completeness men
121 då får ju också en frekvens. Du kan ju göra en ändring på måndag i hur dina rutter dras till
122 exempel, du testar att trafikera andra rutter och sen har vi typiskt två dagars fördröjning av det
123 här datat då på grund av etiska och lojala skäl. Men på onsdag kan du utvärdera om det är var
124 bra eller dåligt. Genom att egentligen logga in i våra verktyg och klicka, tänk er då att göra
125 den typ av utvärdering förr i tiden, innan det här datasetet. Det hade ju knappt gått och om
126 man skulle gjort det skulle det vara väldigt dyrt så det är definitivt inget man skulle vilja göra
127 eller klara av att göra varje vecka då. Så det är jättespännande utifrån det perspektivet för det
128 är en helt ny typ av datakälla, både i form av completeness i form av representativitet och
129 frekvens då.
130 NG
131 Aa men jättebra svar och sen en av huvudfrågorna kanske då kopplade till vår rapport. Anser
132 du att mobildata har använts under COVID-19 för att bromsa smittspridningen?
133 R4
134 Ja, det skulle jag säga i allra högsta grad och mycket på grund av samarbetet med
135 Folkhälsomyndigheten i Sverige. Vi samarbetar även med liknande, vi hade liknande
136 samarbete med dem andra nordiska länderna. Finland jobbade vi till exempel med
137 statsberedningen, Prime Minister Office kallas dem för. Som hade då en liknande roll som
138 Folkhälsomyndigheten hade här i Sverige. Så dels genom dem här samarbeten som vår data
139 faktiskt har använts och som också står i det här forskningspappret då att det används för att
140 då, som en del i beslutsunderlaget för vilka rekommendationer, restriktioner man vill göra,
141 vilken typ av intervention. Så dels det sen har vi ju också varit väldigt öppna med att bidra
142 med det här data till media också. Vilket jag tycker har varit väldigt viktigt. Jag vet inte om ni
143 kommer ihåg det men det var ju helt galna headlines på rubrikerna på Aftonbladet i början av

144 pandemin. Det var någon som hade varit på Medborgarplatsen i Stockholm och tyckt att det
145 var mycket folk som då blev förstasidans stoff, det är ju inte ens data. Det är ju någon som har
146 tyckt någonting så då ville vi bidra till en datadriven offentlig debatt och diskussion. Kan ju
147 såklart inte konkret säga att det ledde till minskad smittspridning men det är det iallafall med
148 en datadriven dialog som ofta bidrar till det. Sen har vi även gjort lite andra samarbeten, vi
149 gjorde ju ett samarbete med Sahlgrenska, jag vet inte om ni känner till det?

150 NG och SL

151 Vi har fått höra lite kort om det av en annan respondent.

152 R4

153 Aa då jobbade vi iallafall med Ericsson, Chalmers och Sahlgrenska och där använde Ericsson
154 och byggde AI-modeller på vår data och liten annan data som har att göra med intagna
155 patienter och så vidare. Och då så byggde dem en prediktionsmodell som Sahlgrenska
156 använde sig av under pandemin för att förutsäga vårdbehovet. Så dem använde faktiskt det i
157 deras planering, hur dem skulle planera sina resurser, hur många bäddar behöver dem, hur
158 många intagna och det i sin tur ledde till effektiv vårdplanering vilket såklart är bra på många
159 sätt. Vi gjorde en pressrelease på det som jag kan dela med er.

160 NG

161 Ja, jättegärna. Den kikar vi gärna på. Vi har även läst artiklar om att Google har gjort sådana
162 här i mindre skala, separata forskningsgrupper som har använt Googles data för att förutspå
163 sjukhusinläggningar.

164 R4

165 Ja, det finns ju en hel uppsjö forskningspapper kring hur mobilitetsdata har och kunnat
166 användas. Dock inte jättemycket konkreta sätt där det faktiskt har använts och där tycker jag
167 vi sticker ut i Sverige faktiskt med Folkhälsomyndighetens samarbete. Vi på Telia kan inte
168 göra någonting för att styra, det är myndigheterna som gör det. Där har ju vår data varit en del
169 av underlaget för datan som Folkhälsomyndigheten hade så där tycker jag vi var hyfsat unika
170 ändå. Sen finns det otroligt mycket forskning som har gjorts på när man vill antingen en del
171 när man vill ha bättre epidemiologiska modeller eller att man vill visa på en korrelation att det
172 kan vara nyttigt och så vidare. Det är inte alls lika många som berättar att det faktiskt har
173 använts. Så det är jättekul att varit med och där det faktiskt användes av

174 Folkhälsomyndigheten här i Sverige då. Vi samarbetade även med EU-kommissionen också
175 där vi och andra telecomoperatörer i Europa delade då den här anonyma datan med dem och
176 då finns det en entitet som lyder under kommissionen som heter Joint Research Center som
177 bedrev en hel del forskning på vår och andra telecomoperatörers data. Dem har också
178 publicerat forskningspapper kring då smitta och rörlighet och hur kan på olika sätt kan
179 omforma mobilitetsdata för att då ge svar på hur man bör forma åtgärder för att begränsa
180 smittspridning.

181 NG

182 Aa men jättebra och vidare till nästa fråga som vi kanske redan har varit inne och tochat på
183 men vilka slags beslut är det man har kunnat ta baserat på Telias data och då tänker jag under
184 COVID-19?

185 R4

186 Ja och där vi du ju verkligen ta in svar från Folkhälsomyndigheten. Det uttryckte sig en del i
187 vårt forskningspapper, där står det lite om att dem använde det som en del av deras
188 beslutsunderlag tror jag det stod men här får ni fråga dem.

189 NG

190 Sista frågan för det här segmentet. Många länder implementerade
191 kontaktspårningsapplikationer under COVID-19, fanns det tankar på att Telia skulle utforma
192 en sån, tillsammans med Folkhälsomyndigheten?

193 R4

194 Vi blev approcherade av många andra företag, små och stora som ville ha med oss på olika
195 typer av smittspråkingsappar. Vad jag vet iallafall av den delen på Telia som jag
196 representerar, vi gick aldrig vidare med någon av dem. Vad jag minns så var det inte aktuellt
197 med Folkhälsomyndigheten, jag har faktiskt för mig att Anders Tegnell uttalade sig offentligt
198 om det här också, att det inte var något han trodde på. Har jag för mig men det kan ni nog
199 hitta på Google. Det var alltså ingenting vi utforskade speciellt seriöst. Vi tittade ju på det
200 såklart men lade aldrig någon vidare kraft på det.

201 SL

202 Då är det nästa del, förväntningar på samarbetet mellan Telia och Folkhälsomyndigheten och
203 då tänkte jag kolla vad det fanns för mål initialt med samarbetet?

204 R4

205 Målet initialt i en brinnande pandemi, i slutet av mars, början på april 2020 var att hjälpa till
206 och bidra och vi trodde att vi kunde med vår data som redan vara produktifierad, vi hade den i
207 tjänst, vi hade rätt ut allting kring etiska vägval och legalt, GDPR och så vidare. Vi kände att
208 vi ville hjälpa till och vi tyckte att vi kunde göra det och på något sätt lyckades vi få till det
209 här med Folkhälsomyndigheten som var i den här tiden under omänskligt tryck och fick också
210 otroligt mycket förslag från allmänheten. Alla vill ju hjälpa till, vilket är fint på ett sätt men
211 det blir svårt att sälla ut då men på något sätt så lyckades vi få till det här och vi kunde inleda
212 samarbetet och så var det. Vi ville göra något bra och rätt för samhället att bidra till att mildra
213 effekterna av pandemi.

214 SL

215 Mm och tycker du att målen som sattes upp har uppnåtts?

216 R4

217 Ja, det får jag nog säga. Det var ju ett otroligt intensivt samarbete och otroligt häftigt och hela
218 teamet var ju engagerade och man kände ett purpose på ett sätt som jag tror att vi aldrig har
219 gjort eller kanske aldrig igen kommer känna. Jag tyckte att det införlivade att vi hjälpte till
220 och kunde bidra.

221 SL

222 Ja och var det några mål som tillkom eller ändrades några mål under samarbetets gång?

223 R4

224 Det som hände var ju att Folkhälsomyndigheten, jag tror vi har skrivit det i vårt papper också
225 så jag ska inte försöka sätta ord i deras mun men dem var ju inte så vana vid den här typen av
226 data. Eller inte alls skulle jag säga så det initiala jobbet handlade ju väldigt mycket om hur
227 dem skulle kunna använda det här som en del av deras beslutsprocesser.

228 SL

229 Mm, intressant.

230 R4

231 Så målet från vår sida eller det stora målet var ju att på något sätt bidra eller hjälpa till. Så det
232 förändrades ju inte däremot själva huret, hur det sen blev och hur man skulle göra det. Det var
233 ju något som vi iterativt tillsammans då tog fram. Så vi hade ju väldigt massa olika
234 dashboards, visualiseringar och frågeställningar och sen landade det i att nä, vi behöver dem
235 här sakerna. Som Folkhälsomyndigheten och eftersom vi har gjort dem här iterationerna och
236 då så blev det en av våra produkter som nu är standardiserad som vi kallar för Standard
237 Dashboard eller Generic Dashboard och det är nu ett kommersiellt upplägg så till en början så
238 gjorde vi det här pro bono med Folkhälsomyndigheten men sen ville ju dem ha det på riktigt
239 också såklart och vi behövde ha en långsiktighet i det här så blev dem en kund och kan ställa
240 krav på oss gällande service levels och support, och få en längre garanti. Så målet blev väl om
241 möjligt, så tidförvandlades vårt mål till att det är en viktig kund för oss och nu när pandemin
242 förhoppningsvis inte kommer tillbaka, hoppas vi att det här datat fortsatt är viktigt för dem
243 och att då ser vi i motsättningen att det här är en av flera viktiga kunder för oss.

244 NG
245 Aa, juste. Då går vi över till risker och utmaningar med mobildatan. Vilka är dem största
246 utmaningarna för att kunna använda mobildata under COVID-19 anser du?
247 R4
248 Under COVID-19. Alltså kopplat till att hantera pandemin menar du eller?
249 NG
250 Ja, precis.
251 R4
252 En av de största utmaningarna eller den största utmaningen skulle jag vilja säga är ju att
253 många var ovana med den här typen av data och dem här insikterna man kan få ut av den här
254 datan och det var inte ofta någonting man hade jobbat med. I vissa fall, till exempel Oslo
255 kommun som vi jobbade mycket med deras krishanteringsteam dem hade ju varit en kund
256 sedan tidigare så dem kände ju till vårt data. Men i dem allra flesta fall så var det ju att man
257 var ovan och visste inte riktigt och var inte van med den typ av datan. Utmaningen handlar ju
258 då om att i en tid av kris och oerhörd press att kunna lära sig någonting nytt om man ska
259 kunna jobba med det här som då var ganska okänt för många.
260 NG
261 Aa men det är ju en jätteintressant aspekt. Och om man pratar kring
262 kvalitetssäkringsprocessen. Hur har den gått till för att presentera data med så hög kvalité som
263 möjligt?
264 R4
265 Ja, det är i stort sett det viktigaste vi gör. Har man en dataproduct eller en insiktsproduct så är
266 ju kvaliténn då det absolut viktigaste vi har. Så det är ju något vi har jobbat med rigoröst det
267 är ju en kontinuerlig förmåga i vår organisation och vi är aldrig nöjda. Sen har vi ju startat
268 upp olika samarbeten för att få till att vi verkligen säkrar det här perspektivet på att producera
269 bra och jämförbar statistik över tid så vi jobbar ju till exempel sen ett tag tillbaka med SCB,
270 Statistikmyndigheten SCB, dem kallas Statistiska centralbyrån. Så med dem har vi jobbat fram
271 en del av vår det vi kallar vår extrapoleringsmodell egentligen hur vi skapar statistik från data
272 från Telias mobilnät. Där har dem hjälpt oss att förbättra dem och dem har publicerat en
273 rapport om det. Så vi jobbar ju med det här på ett väldigt öppet sätt och det är inget vi
274 försöker ha som en black box och dölja utan där vill vi göra det i öppenhet tillsammans med
275 våra kunder och partners. Vi har personer som är ansvariga för det så vi har en
276 organisationsförmåga kring det. Vi försöker också utbilda våra kunder i att datakvalité inte är
277 en etta eller en nolla utan det handlar om olika typer av osäkerheter. Det finns inga datakällor
278 som är helt perfekta, ger hundra procent precision, representation, presentation. Det finns inte,
279 det finns alltid lite olikheter beroende på om man vill titta brett eller smalt, under långt tid
280 eller kort tid så påverkar det här osäkerheten då. Så att det är också en del i hur vi pratar med
281 våra kunder och aa kontinuerligt arbete. Det är mycket vi kan bli bättre på. Det blir ju en, det
282 finns inget slut på det här, det är kontinuerligt arbete både med egna förmågor och
283 tillsammans med våra kunder.
284 NG
285 Spännande och över till nästa fråga. Det finns ju en misstro i samhället gällande hur
286 verksamheter och myndigheter hanterar personlig data. Är det något ni aktivt arbetar med för
287 att bibehålla samhället förtroende i ert arbete med mobildatan då?
288 R4
289 Ja, det gör vi och det gör vi genom att vara väldigt transparent och prata väldigt mycket om
290 det här. Vi gör det också genom att använda anonym eller det vi kallar anonymiserad och
291 aggregerad data. Så den data vi jobbar med är ju fri från personuppgifter. Så vi jobbar ju
292 mycket med utbildning kring detta. I form av att vi vill vara väldigt öppna om detta och prata
293 om detta när vi pratar med media och vi lägger upp på vår hemsida, sociala medier. Vi har

294 verkligen försökt att bidra till att då öka förståelse för den här typen av data. Varför det är
295 okej och varför det är bra att jobba med den här typen av data. Så det var ett långt svar på
296 frågan men ja, det har vi absolut och det är en jätteviktig del och jag tror verkligen många
297 företag hade tjänat med att använda anonym data. Istället för att faktiskt gå in i och inte
298 anonymisera och istället jobba med persondata och då behöver man jobba med mer givande
299 och det är svårare i ett GDPR-perspektiv och ökar risken i att göra fel. Men datan vi jobbar
300 med är alltid anonymiserad så minskar också riskerna så det är återigen en viktig fråga. Vi
301 försöker utbilda så mycket som vi kan, brett om detta genom att vara transparenta.
302 NG
303 Amen, jättebra och just det där med att den är anonymiserad, finns det brister inom anonym
304 data, finns det potentiella hot som skulle kunna påverka eller avslöja den datan?
305 R4
306 Per definition ska ju anonym data inte kunna identifiera någon sen aggregerar vi ju också så vi
307 tittar alltid på större grupper. Man kan säga att det är 1000 personer som färdades från södra
308 Stockholm till norra Stockholm. Så det finns ju, vi lägger ju in extra då som vi kallar
309 safeguards för att då säkerställa att det aldrig går att gå tillbaka. Den är irreversibel
310 anonymiserad och här är det ju viktigt att man kan prata med hur man vill anonymisera och
311 hur man säkerställer detta. Det är lätt för någon att säga att vi har anonym data. Okej, på
312 vilket sätt. Har ni testat att se om det går att återidentifiera någon. För om det går att
313 återidentifiera någon, då är det inte anonym data. Så det här är ju till exempel vi har ju tagit
314 extern hjälp som har försökt att knäcka detta också och vilket inte har gått. Att säkra
315 anonymiteten är ju mer än att bara säga att man är anonym och tar bort persondata man
316 behöver ha en rigorös process som behöver vara legalsäkrad också så man lever upp till
317 GDPR. Det finns också etiska avväganden i att även om det är anonym data så tillhandahåller
318 vi den inte i realtid då vi anser att det finns för stora sidoeffekter. Genom att ha ett verktyg där
319 man kan förstå var folksamlingar är i realtid. Det är inte det vi sysslar med utan vår data är
320 alltid fördröjd minst 24 timmar ofta två dygn. Också för att säkerställa att vi känner oss okej
321 ur ett etiskt perspektiv.
322 SL
323 Och då går över till sista delen med möjligheter och framtid. Om man spekulativt inte
324 behövde använda anonymiserad data som i vissa mindre demokratiska länder. Tror du att man
325 då mer effektivt hade kunnat bromsa smittspridning?
326 R4
327 Det är en mindre fråga om data och egentligen vad du gör för åtgärder. Ehm, Kina gjorde ju
328 det. Ni får göra en bedömning huruvida det var accepterat eller inte och till vilken kostnad på
329 den personliga integriteten. Det handlar mer om vad man gör med det. Jag tror inte det men
330 det beror på vilket use-case man har och hur en myndighet vill begränsa smittan. Det fanns
331 faktiskt inget som, eller jag kan säga såhär de samarbeten vi har haft med myndigheter och
332 offentlig verksamheten kring just begränsa spridningen där vår anonyma och aggregerade
333 mobildata har varit en faktor så kan jag inte säga att i dem fallen hade gjort någon skillnad om
334 den inte var anonymiserad och aggregerad.
335 SL
336 Aaa, okej. Tror du att exakthet och precision av mobildata kommer förbättras i framtiden?
337 R4
338 Ja, det kommer det absolut göra och det är ju en kombination av två saker. Det ena är att man
339 mer och mer börjar förstå vad man göra med den här typen av data. Det vill säga att man
340 identifierar dem här frågorna som man behöver för att då kunna lösa saker inom olika
341 verksamheter. Och allt eftersom man blir mognare kring det så kommer man också behöva
342 mer avancerad frågor kring precision, till exempel geografiskt. Idag går vi ner på 500 x 500
343 meter men ju mer detaljerad man blir ju mer avancerad användningsfall är ju det då. Så dels

344 tror jag att det blir en mer mognad marknad av den här typen av data och sen också då
345 parallellt av vad det ges för tekniska möjligheter. Det finns ju något som heter 5G som är
346 nästa generations nät, det till exempel medger en bättre precision men kan vara så att man
347 behöver medgivande för att det blir så exakt att det blir svårt att anonymisera. Då behöver
348 man någon gång liksom, beroende på användningsfall behöver jobba med ickeanonym data
349 och då är det helt andra processer såklart. Man behöver medgivande och kunna ta bort data
350 och så vidare. Så det är både mognaden på marknaden och de tekniska möjligheterna och det
351 sammantaget kommer ju användas i framtiden är jag helt säker på.

352 SL

353 Yes och vår sista fråga då. Vad anser du att det finns för framtida användningsområden med
354 mobildata i syfte att bromsa spridningen av smittsamma sjukdomar?

355 R4

356 Ja, det finns ju jättemycket och där finns det ju mycket litteratur och forskning gjord då som
357 har accelerat pandemin och det har producerats en väldigt massa papers kring detta. Mycket
358 handlar ju om att kunna ha mer precisa epidemiologisk modeller. Där man tidigare, klassiska
359 modeller använder ju ganska mycket eller är ganska grova drag och kan använda historisk
360 data hur rörelsen var mellan regioner för kanske några år sedan. Det vi såg med COVID var
361 att det kunde
362 variera väldigt mycket mellan olika delar av en stad till exempel och också väldigt annorlunda
363 över tid. Så det är mobildata men det är ju lika mycket att man börjar jobba med big data
364 istället för att ha grova statistiska modellerna så det är ju mycket forskning och det är ju alltid
365 att dels kunna modellera spridning till att då informera om påverkan på segregation och det
366 finns mycket i litteraturen och det kan ni nog google fram än det jag har i mitt lilla huvud.
367 Men jag ser jättestora möjligheter men framförallt att man utifrån börjar bli van att jobba med
368 big data och sen råkar det nog vara mobildata eftersom det visade sig vara så relevant med
369 COVID men det kan vara andra datamängder också. Sen kanske man ska skapa en vana att
370 jobba med big data och mera AI istället för rena statistiska modeller. Det vore fantastiskt om
371 man kan använda detta för att skapa early warnings system. Problemet med pandemi var ju att
372 det ofta när man upptäckte smittan var det för sent. Tänk om man då med hjälp av mobildata
373 och data från tidig RNA-sekvensering var det i det här fallet där man då identifierade olika
374 stammar, olika ställen med mobildata, rörelsedata skulle man ju kunna förutse vart är nästa
375 smittohärd och inställd förväntan. Vilket skulle leda till en helt annan produktivitet. Det
376 kräver ju att man också jobbar med detta när det inte är pandemi så det krävs ju att man
377 bygger den här beredskapen.

378 NG

379 Precis precis.

380 SL

381 Tack så mycket. Har du något du vill tillägga eller någon fråga till oss?

382 R4

383 Nä, bra frågor och kul att se vad det blir av detta.

384 NG

385 Stort tack för att du tog din tid och delade med dig.

1 Appendix 9 – Transkribering intervju 5

2 SL = Simone Lewenhaupt

3 NG = Nils Gunnarson

4 R5 = Respondent 5

5

6 SL

7 Första frågan är då, hur gammal är du?

8 R5

9 30.

10 SL

11 Vilket kön identifierar du dig som?

12 R5

13 Man.

14 SL

15 Vilket företag jobbar du på och kan du lite kort beskriva din yrkesroll?

16 R5

17 Jag jobbar ju då inte på ett företag utan vid Folkhälsomyndigheten och det jag sa tidigare är
18 att jag har kommit in som utredare med epidemiologisk expertis på basen av att jag är läkare i
19 botten och doktorerar i epidemiologi och global hälsa med fokus på beteende och
20 smittspridning.

21 SL

22 Yes och hur länge har du jobbat med mobildata kopplat till COVID-19?

23 R5

24 Sedan 2020.

25 SL

26 Amen, toppen då tänkte vi gå in på frågor gällande användning och nytta av mobildata. Och
27 då får vi ju se om det är några frågor som du inte riktigt kan besvara är det bara att du säger
28 det för vi har även intervjuat Telia och försöker ställa samma typ av frågor. Vissa passar dig
29 bättre och andra dem.

30 R5

31 Mm, okej.

32 NG

33 Yes men då går vi över till användning och nytta för mobildata då. Första frågan, vikten av att
34 transformera data till insikt är något som ofta nämns av Telia Crowd Insights. Hur skulle du
35 kortfattat beskriva processen från data till insikt?

36 R5

37 Jo men det är något som gäller egentligen samtliga dataanalyser att man kan ju inte bara se till
38 rådata utan det behöver slås ut på något form av sätt som är begripligt. Dels vad det gäller
39 internt hos oss men också även externt. Vi har ju publicerat dem här översikterna över
40 rörelsedata i veckorapporter som har varit till för allmänheten och det som behöver göras är ju
41 dels att förstå det grundläggande, vad är det för data som har insamlats och utifrån det kan vi
42 se vad det finns för styrkor och svagheter. Vi har ju även jämfört med andra datakällor där vi
43 bland annat har tittat på data över flygtrafik, data från Google och Apple. Aa, det finns en hel

44 drös med saker man kan gå in på som är styrkor och svagheter i varje datatyp. Och vad som är
45 viktigt för oss, är nog... Har nog varit att datan är representativ för befolkningen i stort för att
46 vi ska kunna säga att de finns någonting, någon slags insikt vi kan dra utav det här som är
47 generaliserbar. Kanske inte till hundra procent så men därav har vi också behövt arbeta med
48 olika typer av indelningar av datan. Till exempel när vi har jämfört resor som har varit över
49 kommungränser kontra resor över regiongränser. Det är ju till viss mån varit i syfte att följa
50 upp de restriktionerna vi hade tidigare som syftade till att minska längre resor. Och sen så för
51 att få någon bild av resande överlag och sen har vi ju sett på det som vi definierar som total
52 resemått. Som i stort sett är samtliga resor som är större än den här definitionen från en
53 telemast till en annan. Som kommer med vissa svagheter då telemasterna inte är jämnt
54 stationerade över hela landet men det som gäller är att först få en god bild av vilken data man
55 får tillgång till och därifrån bedöma hur går det att koppla den här datan till saker som vi ser
56 generellt i smittspridningen. Det är svårt att göra direkta sådana kopplingar. Så det vi har velat
57 göra är att se följsamheten bland befolkningen kring de åtgärder som vi har eller de
58 rekommendationer som har kommit från myndigheter. Jag vet inte om det ger en viss bild
59 men det finns steg, alltigenom där man behöver ta ställning till hur representativ datan är, hur
60 generaliserbar fynden är och även vilka fynd kan kommuniceras väl. Vad är intressant och det
61 är inte alltid det som är mest sensationellt som är det som faktiskt sen har en avgörande
62 betydelse för smittspridningen. Till exempel kan man ju se om en tusenprocentig ökning i
63 vissa mått så behöver det inte betyda särskilt mycket om basnivån är att fem personer har rört
64 sig på en viss plats medan det som den har ökat upp till är 50 personer. Till skillnad från en 20
65 procentig ökning i ett annat mått kan vara väldigt mycket mer relevant eller 10 procent. Men
66 det som gör resedatan unik är ju att vi inte riktigt har haft någon baslinje sedan tidigare och
67 det har funnits en hel del andra sätt att mäta följsamhet Men det har inte funnits några sätt där
68 man har gjort på det här sättet sedan tidigare. Med mer eller mindre objektiva mått, utan det är
69 snarast självrapporterad data där individer beskriver hur dem efterlever rekommendationer. Så
70 mycket av arbetet har gått ut på att identifiera vad som är en rimlig baslinje, vilka
71 förändringar kan vi förvänta oss till följd av årstid och vilka kan vi förvänta oss beror på... Aa
72 på en följsamhetsrestriktion eller en icke följsamhet för det är ju också så att ett totalt... För
73 det är ju också så att folk arbetar ju fortsatt, alldeles oavsett och det gäller i alla länder. Så, jag
74 får nog stanna där, jag pratar för länge.

75 NG

76 Det är ingen fara alls men vi går även till nästa fråga. Big data har använts länge inom
77 folkhälsan. På vilket sätt har möjligheten att samordna data och kunna analysera datan med
78 moderna verktyg förenklat erat arbete?

79 R5

80 Jag skulle nödvändigtvis inte beskriva det vi har gjort i det här arbetet som big data sen
81 kanske vi givetvis tar del av en produkt som har tagits fram av till exempel Telia. Men vi har
82 inte suttit med den absoluta rådatan där vi kan följa individer. Allting är ju anonymiserat och
83 aggregerat på olika nivåer som vi har avtalat. Så i viss mån kanske det är att vi arbetar i
84 gränslandet i vad som har varit big data men vi har inte några stora databaser som vi har
85 anonymiserat.

86 NG

87 Nej men intressant! Hur har arbetssättet sett ut innan man hade tillgång till mobildata? Och
88 jag tänker med liknande rörelsemönstersrapporter. Är det något ni har gjort?

89 R5

90 Jag har lite svårt att svara på all övervakning men vad man kan säga är ju att det är ett behov
91 som vi har identifierat dela på grund av pandemin men också under pandemin för att det är
92 något som har kunnat bidra till en del intressanta iakttagelser som kan vara relevanta för sånt
93 som som smittspridning av influensa och andra säsongberoende sjukdomar. Med det sagt så

94 har arbete gjorts med att följa till exempel, jag vet inte om ni känner till Google Flu Trends
95 och så vidare, sånt arbete. Även 1177, det finns publicerat från vår enhet om det men det är ju
96 inte rörelsedata. Det är mer utifrån vårt perspektiv som vi har klumpat ihop med rörelsedata
97 som alternativa datakällor till det mer traditionella smittspridningsövervakningen som
98 kommer från framförallt sjukvården.
99 NG

100 Spännade, verkligen och anser du att mobiloperatörsdata har använts under COVID-19 för att
101 bromsa smittspridningen? Vilket kanske kan beskrivas som en av dem viktigaste
102 frågeställningar för vår rapport.
103 R5

104 Ehhmm, ja det kan vi absolut säga att det har varit en pusselbit i att utforma vårt arbete. Vi
105 har tagit in en otrolig mängd olika data och källor. Rörelsedatat har utgjort en pusselbit och
106 det är ju därav vi har publicerat kring den och det har kommenterats en del kring det. Sen är
107 det svårt att säga om det har varit något som... Det är väldigt indirekt hur mobildata har
108 bidragit till minskad spridning men vi har ju fått en förståelse för hur människor kan anpassa
109 sig och så.
110 NG

111 Så man skulle kunna säga att den har bidragit till bromsad smittspridning i kombination med
112 andra datakällor?
113 R5

114 Ja men det får man nog lov att säga
115 NG

116 Okej, tack! Det leder oss in på nästa fråga, vilka slags beslut har man kunnat ta baserat på den
117 här mobildatan? I form av rekommendationer, policies eller restriktioner.
118 R5

119 Jag kan nog inte säga mer än att den har informerat inför beslut och arbetet har ju använts
120 inom myndigheten.
121 NG

122 Sista frågan för det här segmentet. Många länder
123 implementerade kontaktspårningsapplikationer under COVID-19. Fanns det tankar på att
124 Telia och Folkhälsomyndigheten skulle utforma en?
125 R5

126 Det kan jag egentligen inte svara på i och med att jag dels började arbeta mot slutet av 2020
127 men jag får nog vara... Jag kan tyvärr inte svara på det.
128 SL

129 Ingen fara, vi förstår. Då går vi över till förväntningar på samarbetet med Telia och då tänkte
130 jag kolla vilka mål det fanns initialt mellan Telia och Folkhälsomyndigheten?
131 R5

132 Jo men syftet eller arbetet kom ju till för att Telia som ni säkert redan har läst i Pär och Telias
133 gemensamma artikel var ju att de hade en datakälla som vi såg kunde vara nyttig och öka vår
134 förståelse av hur befolkningen rörde sig och hur de tog till sig diverse rekommendationer för
135 att kunna följa det. Och eftersom de hade den datakällan var vi intresserade av att se hur vi
136 kunde nyttja det. Och det är ju något vi har nyttjat i kombination med annan rörelsedata men
137 även data som på olika sätt följt efterlevnad av rekommendationer.
138 SL

139 Och har målen som sattes upp eller dem initiala målen eller tankarna med samarbete uppnåtts
140 enligt dig?
141 R5

142 Ja, jo men det får man nog på ett sätt säga. Vi jobbar ju fortfarande med det här, jag har
143 faktiskt ett möte med Telia i eftermiddag så det får man nog säga även fast vi inte riktigt är

144 klara eller vad man ska säga. Vi får se hur vi kan fortsätta använda datan och vad det finns för
145 behov.
146 SL
147 Är det några mål som har tillkommit eller ändrats under samarbetets gång?
148 R5
149 Nej eller målsättningen har ju varit att använda datan för att förstå följsamhet,
150 rekommendationer och befolkningens rörelser allt igenom. Och det kvarstår.
151 NG
152 Spännade, då går vi över till risker och utmaningar. Vilka är det största utmaningarna för att
153 kunna jobba med mobildata under COVID-19 anser du?
154 R5
155 Jo men det har ju varit det faktum att det är en ny metod som inte har funnits under så lång tid
156 och det finns en ganska kraftig underliggande variation hur människor resor och betar sig
157 under året och det gäller att den information och den tolkning man gör inte försöker gå för
158 långt. Både att inte vara försiktig eller för offensiv heller. Om man ser till
159 exempel, promenader utomhus så fångas ju det också i datan och det är ju något som är alltid
160 genom aldrig har varit ifrågasatt. Så det har inte varit alldeles lätt och då har vi fått jämföra
161 med olika källor och vi har nog bedömt olika källor som olika tillförlitliga. Därav har det inte
162 alltid varit möjligt att dra slutsatser... Samtalet bryts).
163 NG
164 Hör du oss nu?
165 R5
166 Jag hör er, men som sagt måste man vara försiktig med sina slutsatser och vilket håll man
167 anser att det lutar emot och både väga Epidemiologiska principer kring smittspridning
168 gentemot det man ser i i datan.
169 NG
170 Intressant perspektiv och hur har kvalitetssäkringsprocessen gått till för att presentera data
171 med så hög kvalitet som möjligt i de här rapporterna?
172 R5
173 Men det som gäller då är att jämföra med olika datakällor och vara säker på att dem fynd man
174 ser i ett mått överensstämmer till allra minsta att dem är möjliga till det man ser i annat. Det
175 jag kan lägga till här också, en svårighet i all den här datan är att den kommer retroaktivt.
176 Förvisso bara några dagar efter men därför blir det svårt att omedelbart se vad som händer
177 eller vad som följer av en viss rekommendation men det blir viktigt att inte tolka in heller det
178 man får från andra källor i den datan man har framför sig. Att inte komma med inte allt för
179 lätta slutsatser. Nu tappade jag nästan bort frågan men det var en viktig aspekt i det här med
180 att datan kommer retroaktivt.
181 NG
182 Nej men det är ju passande kvalitetssäkringsprocessen att datan kommer retroaktivt... Det
183 finns en misstro i samhället gällande hur verksamheter och myndigheter hanterar personlig
184 data. Är det något ni aktivt arbetar med för att bibehålla samhällets förtroende i ert arbete med
185 mobildata?
186 R5
187 Jag får nog vara ganska kort gällande den frågan men om ni har sett dem här
188 veckorapporterna så står det hur vi har arbetat med den. Vi har ju enbart tillgång till en
189 aggregerade och anonymiserade datan men mer än så kan jag nog inte kommentera där.
190 SL
191 Vi förstår, tack! Nu några sista frågor angående möjligheter och framtid. Om man spekulativt
192 inte behövde anonymisera datan som i mindre demokratiska länder, hade man då mer
193 effektivt kunna bromsa smittspridningen med datan?

194 R5
195 Där här blir ju nästan en fråga vad jag tycker om den kinesiska modellen och det vill jag nog
196 inte kommentera.
197 SL
198 Ingen fara. Tror du exakthet och precision av mobildata kommer förbättras i framtiden?
199 R5
200 Det vi är intresserade bör inte riktigt påverkas av mer exakt eller om det är mer på vilka GPS-
201 koordinater en individ befinner sig på så är det inte så intressant för oss. Men om man ska
202 spekulera, det är inte riktigt mitt fält men om täckningen av mobilmaster blir högre i
203 Norrlands inland så är det klart att det ökar representativiteten av vår data där. Och jag kan
204 svara på föregående frågan lite också och jag tror inte att eller jag tror utifrån det vi har jobbat
205 med så hade mer detaljerad data inte varit bättre men det är fortsatt så att vi har enbart arbetat
206 arbetat med anonymiserad och aggregerad data.
207 SL
208 Sista frågan då. Vad anser du att det finns för framtida användningsområden med mobildata i
209 syfte att bromsa spridning av smittsamma sjukdomar?
210 R5
211 Ja, nu kanske har jag kanske lättare att uttala mig inom perspektivet som doktorand. Åh, jag
212 återkommer till det. Men jag tror det kan finnas en nytta av att följa den här typen av data och
213 även att försöka skapa sig en typ av baslinje som man kan lättare förstå variationerna utifrån.
214 Och som sagt, svårigheten att utifrån den här datan identifiera vilka typer av rörelser som kan
215 vara associerade med ökad risk för smittspridning som man behöver på något sätt identifiera.
216 Och det tror jag är väldigt svårt att göra men jag tror det finns definitivt något värde av den
217 här typen av arbete. Och vi har ju arbetat med publikationer kring detta och kommer försöka
218 beskriva det i samband med konferenser har jag presenterat om rörelsedata och kommer även
219 att göra det i sommar. Jag kan skicka över den postern men där är vår slutsats att det finns
220 någon typ av användning men det blir väldigt spekulativt om när en ytterligare pandemi skulle
221 kunna dyka upp. Svårt utifrån det om det behövs kontinuerlig beredskap och i sånt fall hur.
222 Det är stora frågor och det gäller ju givetvis mycket mer än rörelsedata.
223 NG
224 När du pratar om den här baslinjen och baslinjeperioden. Handlar det om att man vill hitta ett
225 normalt beteende hos den svenska befolkningen innan pandemin var ett faktum? Eller vad
226 syftar baslinjen till?
227 R5
228 Ja, precis och i det även ta med att det normala beteendendet är dels varierande över tid men
229 den typ av... Det finns inget direkt samband där mer rörelser, mer resor innebär mer
230 exponering för smitta. Så det gäller också att förstå vilka typer av ökning skulle kunna öka
231 till en ökad smittspridning i samhället och vilken typ av mönster i rörelse gör inte det. Utifrån
232 det som det här med baslinje blir viktigt.
233 NG
234 Stort tack!
235 SL
236 Det var alla frågor vi hade, är det något som du vill lägga till eller är det något du tycker vi
237 har missat?
238 R5
239 Nej men det tror jag inte. Det finns ju väldigt mycket mer att säga men jag uppskattar mycket
240 om jag får se hur ni tänkte citera någonting. Innan ni är klara med det och jag får fundera om,
241 det blir nog bra. Roligt ämne för er uppsats!
242 NG

243 Nej men det har varit jättekul ämne och vi hoppas att med vårt intervjumaterial kunna besvara
244 vår forskningsfråga. Stort tack för att du tog din tid och superspännande att få höra din take på
245 våra frågor. Vi ser till att skicka över det transkriberade materialet så fort det är klart.
246 R5
247 Det är bra! Så får vi hoppas att jag inte var alltför otydlig och är det så kan ni bara skicka en
248 fråga.
249 SL och NG
250 Yes, absolut. Ha det så bra!
251 R5
252 Detsamma, hej!

Referenser

- Apple. (n.d.). Mobilitetstrendsrapporter, Tillgänglig online:
<https://covid19.apple.com/mobility> [Hämtad 31 mars 2022]
- Arena, F., Pau, G. (2020). An overview of big data analysis, *Bulletin of Electrical Engineering and Informatics*, Tillgänglig online:
<https://doi.org/10.11591/eei.v9i4.2359> [Hämtad 11 maj 2022]
- Bryman, A. & Bell, E. (2017). *Företagsekonomiska Forskningsmetoder*, Stockholm: Liber AB
- Chang, W., & Grady, N. (2019). NIST Big Data Interoperability Framework: Volume 1, Definitions, *Special Publication (NIST SP)*, vol.1, no.3, pp. 53, Tillgänglig online:
<https://doi.org/10.6028/NIST.SP.1500-1r2> [Hämtad 4 april]
- Dolley, S. (2018). Big Data's Role in Precision Public Health, *Front. Public Health* 6:68, Tillgänglig online: [https://www.frontiersin.org/articles/10.3389/fpubh.2018.00068/full#:~:text=intervention%20\(52\).,Big%20Data%20in%20Precision%20Public%20Health,and%20understanding%20disease%20\(57\)](https://www.frontiersin.org/articles/10.3389/fpubh.2018.00068/full#:~:text=intervention%20(52).,Big%20Data%20in%20Precision%20Public%20Health,and%20understanding%20disease%20(57)) [Hämtad 30 mars 2022]
- Edsberg Møllgaard, P., Lehmann, S., & Alessandretti, L. (2021). Understanding components of mobility during the COVID-19 pandemic, *The Royal Society Publishing*, Tillgänglig online: <https://royalsocietypublishing.org/doi/full/10.1098/rsta.2021.0118> [Hämtad 3 maj 2022]
- Europeiska Kommissionen. (n.d.). Digitala lösningar under pandemin, Tillgänglig online:
https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/digital-solutions-during-pandemic_sv [Hämtad 24 mars 2022]
- Favaretto M, De Clercq E, Schneble CO, Elger BS. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade, *PLoS ONE* 15(2): e0228987, Tillgänglig online:
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0228987> [Hämtad 11 maj 2022]
- Folkhälsomyndigheten. (2020a). Anmälningspliktiga sjukdomar, Tillgänglig online:
<https://www.folkhalsomyndigheten.se/smittskydd-beredskap/overvakning-och-rapportering/anmalningspliktiga-sjukdomar/> [Hämtad 24 mars 2022]
- Folkhälsomyndigheten. (2020b). Spridningen av covid-19 är en pandemi, Tillgänglig online:

- <https://www.folkhalsomyndigheten.se/nyheter-och-press/nyhetsarkiv/2020/mars/spridningen-av-covid-19-ar-en-pandemi/> [Hämtad 23 mars 2022]
- Folkhälsomyndigheten. (2022). Veckorapport om rörelse- och rese­mått, Tillgänglig online: <https://www.folkhalsomyndigheten.se/smittykydd-beredskap/utbrott/aktuella-utbrott/covid-19/statistik-och-analyser/analys-och-prognoser/rapport-om-rorelse--och-resematt/> [Hämtad 30 mars 2022]
- Gerlee, P., Karlsson, J., Fritzell, I., Brezicka, T., Spreco, A., Timpka, T., Jöud, A. & Lundh, T. (2021). Predicting regional COVID-19 hospital admissions in Sweden using mobility data, *Sci Rep 11*, Tillgänglig online: <https://www.nature.com/articles/s41598-021-03499-y> [Hämtad 21 april]
- Google. (2022). Se hur människors rörlighet förändras på grund av covid-19. Tillgänglig online: <https://www.google.com/covid19/mobility/> [Hämtad 31 mars 2022]
- Google Trends. (n.d.). Tillgänglig online: <https://trends.google.com/trends/explore?date=all&q=big%20data> [Hämtad 1 april 2022]
- Harris, S. (2020). Why outbreaks like coronavirus spread exponentially, and how to “flatten the curve”, *The Washington Post*, 14 mars, Tillgänglig online: <https://www.washingtonpost.com/graphics/2020/world/corona-simulator/> [Hämtad 23 mars 2022]
- HBL. (2020). Guterres: Pandemin den värsta globala krisen sedan andra världskriget – rika länder måste hjälpa fattigare, Tillgänglig online: <https://www.hbl.fi/artikel/guterres-pandemin-den-varsta-globala-krisen-sedan-andra-varldskriget-rika-lander-maste-hjalpa-fat/> [Hämtad 23 mars, 2022]
- Jacobsen, D. I. (2002). Vad, hur och varför: om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen, Lund: Studentlitteratur
- Jansen, R., Kovacs, K., Esko, S., Saluveer, E., Söstra, K., Bengtsson, L. & Magpantay, E. (2021). Guiding principles to maintain public trust in the use of mobile operator data for policy purposes, *Data & Policy*, Tillgänglig online: <https://www.cambridge.org/core/journals/data-and-policy/article/guiding-principles-to-maintain-public-trust-in-the-use-of-mobile-operator-data-for-policy-purposes/A4A2A8723151F901F5315279E76F651F> [Hämtad 5 april]
- Khoury, M. J., & Ioannidis, J.P.A. (2014). Big data meets public health, *Science*, Tillgänglig online: https://www.science.org/doi/full/10.1126/science.aaa2709?casa_token=LbjUZla5qqYAAAAA%3AKou0KDz4zgUW-eRd1nDLQtk3bk6oyTiO6khQRTR9MmKlQxE-ZRJydpL5XrEYCfuukWbl_YgIt-WkRyLg [Hämtad 5 april 2022]
- Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*, SAGE Publications

- Lindroos-Hovinheimo, S. (2020). Serious Cyberattack Raises Questions About GDPR Application in Finland, Tillgänglig online: https://helda.helsinki.fi/bitstream/handle/10138/326960/verfassungsblog.de_Serious_Cyberattack_Raises_Questions_About_GDPR_Application_in_Finland.pdf?sequence=1 <https://verfassungsblog.de/serious-cyberattack-raises-questions-about-gdpr-application-in-finland/> [Hämtad 21 april]
- Mansour, O., Kajatzi, M., & Ghazawneh, A. (2022). Using Mobile Data for Understanding Population Movement and Disease Transmission during Covid-19 Outbreak in the Nordics, Honolulu, HI: HICSS , 2022. s. 7151-7160, Tillgänglig online: <https://www.diva-portal.org/smash/get/diva2:1637352/FULLTEXT01.pdf> [Hämtad 4 april 2022]
- Mooney, S. J., & Pejaver, V. (2017). Big Data in Public Health: Terminology, Machine Learning, and Privacy, *Annu. Rev. Public Health* 2018. 39:95–112, Tillgänglig online: <https://www.annualreviews.org/doi/pdf/10.1146/annurev-publhealth-040617-014208> [Hämtad 5 april 2022]
- Oates, B. J. (2006). *Researching information systems and computing*, London: SAGE
- Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., Nadai, M., Letouzé, E., Ali Salah, A., Benjamins, R., Cattuto, C., Colizza, V., Cordes, N., Fraiberger, P., Koebe, T., Lehmann, S., Murillo, J., Pentland, A., Pham, P., Pivetta, F., Saramäki, J., Scarpino, S., Tizzoni, M., Verhulst, S. & Vinck, P. (2020). Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle, *Science Advances*, Volume 6, Tillgänglig online: <https://www.science.org/doi/full/10.1126/sciadv.abc0764> [Hämtad 22 april 2022]
- Polson, N. & Scott, J. (2018). *AIQ: Hur artificiell intelligens fungerar*, Daidalos
- Regeringskansliet, (2022a). Merparten av restriktionerna mot covid-19 tas bort den 9 februari 2022, Tillgänglig online: <https://www.regeringen.se/pressmeddelanden/2022/02/merparten-av-restriktionerna-mot-covid-19-tas-bort-den-9-februari-2022/> [Hämtad 31 mars 2022]
- Regeringskansliet, (2022b). Regeringen föreslår att covid-19 inte längre ska klassas som samhällsfarlig sjukdom, Tillgänglig online: <https://www.regeringen.se/pressmeddelanden/2022/03/regeringen-foreslar-att-covid-19-inte-langre-ska-klassas-som-samhallsfarlig-sjukdom/#:~:text=Med%20smittsamma%20sjukdomar%20menas%20alla,ing%C3%A5r%20ett%2060%2Dtal%20sjukdomar> [Hämtad 16 april 2022]
- Rienecker, L., Stray Jörgensen, P. & Hedelund, L. (2014). *Att skriva en bra uppsats*, Liber, Malmö
- Sarangi, S., & Sharma, P. (2019). *Big data. A beginners introduction*, London: Routledge
- SFS 2004:168. Smittskyddslag. https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/smittskyddslag-2004168_sfs-2004-168

- SFS 2004:255. Smittskyddsförordningen. https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/smittskyddsforordning-2004255_sfs-2004-255
- SFS 2018:218. Lag med kompletterande bestämmelser till EU:s dataskyddsförordning. https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-2018218-med-kompletterande-bestammelser_sfs-2018-218
- Statens medicinsk-etiska råd. (2019). Artificiell intelligens – löftesrik teknik med etiska utmaningar (978-91-519-2158-7), *Statens medicinsk-etiska råd*, Tillgänglig online: https://www.smer.se/wp-content/uploads/2019/06/Smer-konferensrapport_2_webb.NY-REV.pdf [Hämtad 24 mars 2022]
- Telia. (n.d.a). Folkhälsomyndigheten och Telia i Datasamarbete, Tillgänglig online: <https://www.telia.se/privat/aktuellt/hemma-i-folknatet/folkhalsomyndigheten-och-telia> [Hämtad 30 mars 2022]
- Telia. (n.d.b). Telia COVID-19 mobilitetsanalys: svenskarna resande: vecka 52, Tillgänglig online: <https://www.telia.se/privat/aktuellt/hemma-i-folknatet/covid-19-mobilitetsanalys> [Hämtad 24 mars 2022]
- Thunström, L., Newbold, S., Finnoff, D., Ashworth, M., & Shogren, J. (2020). The Benefits and Costs of Using Social Distancing to Flatten the Curve for COVID-19, *Journal of Benefit-Cost Analysis*, Tillgänglig online: <https://www.cambridge.org/core/journals/journal-of-benefit-cost-analysis/article/benefits-and-costs-of-using-social-distancing-to-flatten-the-curve-for-covid19/204BD93C135EC727FAEFC62E3BE72C3B> [Hämtad 23 mars 2022]
- Urbaczewski, A. & Lee, YJ. (2020). Information Technology and the pandemic: a preliminary multinational analysis of the impact of mobile tracking technology on the COVID-19 contagion control, *European Journal of Information Systems*, 29:4, 405-414, Tillgänglig Online: <https://www.tandfonline.com/doi/full/10.1080/0960085X.2020.1802358> [Hämtad 17 april 2022]
- Wesolowski, A., Buckee, O. C., Engø-Monsen, K., & Metcalf, C. J. E. (2016). Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data, *The Journal of Infectious Diseases*, Volume 214, Issue suppl_4, Tillgänglig online: https://academic.oup.com/jid/article/214/suppl_4/S414/2527905?login=true [Hämtad 4 april 2022]
- WHO. (n.d). Coronavirus disease (COVID-19) pandemic, Tillgänglig online: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov> [Hämtad 23 mars 2022]
- Willberg, E., Järv, O., Väisänen, T. & Toivonen, T. (2021). Escaping from Cities during the COVID-19 Crisis: Using Mobile Phone Data to Trace Mobility in Finland, *ISPRS Int. J. Geo-Inf*, Tillgänglig online: <https://www.mdpi.com/2220-9964/10/2/103> [Hämtad 31 mars 2022]

Ågren, K., Bjelkmar, P., & Allison, E. (2021). The use of anonymized and aggregated telecom mobility data by a public health agency during the COVID-19 pandemic: Learnings from both the operator and agency perspective, Cambridge University Press, Tillgänglig online: <https://www.cambridge.org/core/journals/data-and-policy/article/use-of-anonymized-and-aggregated-telecom-mobility-data-by-a-public-health-agency-during-the-covid19-pandemic-learnings-from-both-the-operator-and-agency-perspective/390E5BB8FA9D7A5D870439BC5D7F1996#> [Hämtad 4 april 2022]