



INSTITUTIONEN FÖR TILLÄMPAD IT

HINDER FÖR VIDAREUTNYTTJANDE AV ÖPPNA DATA

Ur ett tjänsteutvecklingsperspektiv

Rickard Andersson

Kristian Mörling

Kandidatuppsats:	15 hp
Ämne:	Informatik
År:	2018
Rapport nr:	2018:095

Sammanfattning

Det sägs att vidareutnyttjande av öppna data från offentlig myndighet kan resultera i positiva effekter såsom ökad insyn och ansvarsskyldighet inom offentlig myndighet samt mer innovativa och effektiva offentliga tjänster. Däremot realiseras inte någon nytta om öppna data inte vidareutnyttjas och offentlig myndighet behöver därför göra mer än att bara tillhandahålla data. För att främja vidareutnyttjande av öppna data behövs insikter i användares behov, samtidigt som de potentiella användarna är en heterogen grupp, bestående av personer med olika bakgrund, motiv och som använder sig av olika typer av teknologi.

Med bakgrund i att insikter i användares behov behövs väljer vi att utforska vilka hinder som upplevs vid vidareutnyttjande av öppna data. Med tanke på användarnas mångfald väljer vi också att begränsa undersökningen till de hinder som upplevs av tjänsteutvecklare.

Empiriskt material bygger på fem intervjuer med personer från två olika länder, där samtliga har erfarenhet av att utveckla tjänster baserade på öppna data. De huvudsakliga resultaten visar bland annat att krav på datakvalitet varierar mellan projekt, att interaktion med offentlig myndighet fungerat både bra och dåligt samt att åsikter om vilken typ av licens som är lämplig skiljer sig.

Nyckelord

Öppna data, Public sector information, PSI, tjänsteutveckling

Title

Barriers to re-use of Open Government Data: A service development perspective

Abstract

It is said that reuse of open government data can result in positive effects like increased transparency and accountability in the government and additionally more innovative and effective public services. On the other hand, no benefit is actually realized if open government data is not re-used and governments therefore need to do more than just publish data. To foster re-use of open data there is a need for insight in the users needs. However, the potential users are a heterogenous group consisting of persons with different backgrounds, motives, and who uses different types of technology.

In the light of a requirement of insight in the user needs, we choose to explore what barriers that are encountered while re-using open data. Due to the diversity among potential users the research is constrained to the barriers that are encountered by service developers.

The empirical material is based on five interviews with people from two different countries, where all of them has experience in service development based on open government data. The main results shows, among other things, that requirements of data quality varies between projects, that interaction with governments has worked both good and bad, and that opinions about which type of license that is suitable varies.

Keywords

Open data, Open government data, Public sector information, PSI, service development

Tack!

Först av allt vill vi tacka Miska Knapek och Elias Arnestrand för värdefull input och hjälp med att komma i kontakt med personer inom öppna data-scenen.

Vidare vill vi tacka alla informanter för att ni tog er tid till att dela med er av era erfarenheter och synpunkter.

Slutligen vill vi även tacka vår handledare, Juho Lindman, för otroligt bra feedback!

Innehållsförteckning

1	Inledning.....	1
1.1	Bakgrund.....	1
1.2	Syfte och frågeställning.....	2
1.3	Avgränsning.....	2
1.4	Undersökningens struktur.....	3
2	Relaterad forskning.....	4
2.1	Öppna data.....	4
2.2	Vidareutnyttjande av öppna data.....	4
2.3	Hinder för öppna data.....	5
2.3.1	Från ett leverantörsperspektiv.....	6
2.3.2	Från ett användarperspektiv.....	7
3	Teoretiskt ramverk.....	10
3.1	Kvalitet och åtkomst för data och metadata.....	12
3.2	Interaktion.....	12
3.3	Kunskap och färdigheter.....	13
3.4	Teknologi och verktyg.....	13
3.5	Licensiering.....	14
4	Metod.....	15
4.1	Studiens karaktär.....	15
4.2	Genomförande av intervju.....	15
4.3	Urval.....	16
4.4	Analys.....	17
5	Resultat.....	19
5.1	Kvalitet och åtkomst för data och metadata.....	19
5.2	Interaktion.....	21
5.3	Kunskap och färdigheter.....	23
5.4	Teknologi och verktyg.....	24
5.5	Licensiering.....	25
6	Diskussion.....	28
6.1	Resultatanalys.....	28

6.1.1	Kvalitet och åtkomst för data och metadata.....	28
6.1.2	Interaktion.....	29
6.1.3	Kunskap och färdigheter.....	30
6.1.4	Teknologi och verktyg.....	30
6.1.5	Licensiering.....	32
6.1.6	Sammanfattade upptäckter.....	32
6.2	Reflektion över undersökningen.....	34
6.3	Förslag till vidare forskning.....	34
7	Slutsats.....	35
8	Referenser.....	36
9	Bilagor.....	39

1 Inledning

I detta kapitel introduceras det undersökta problemområdet. Vidare presenteras den frågeställning som legat till grund för undersökningen, följt av en redogörelse för hur olika begrepp avgränsats. Slutligen beskrivs hur uppsatsen och undersökningen strukturerats.

1.1 Bakgrund

“Open government data” har sedan 2008 blivit ett populärt begrepp för att beskriva data från offentlig myndighet som publiceras öppet i den bemärkelsen att den fritt kan vidareutnyttjas och distribueras vidare (Ubaldi, 2013). Offentliga myndigheter skapar och samlar in stora mängder data och motivationen till att publicera den öppet ligger i att den kan utgöra ett värde inom många olika områden och för många olika grupper av människor (Open Knowledge International, 2016; Janssen, 2011). Fortsättningsvis kommer “Open government data” att benämnas med svenska termen öppna data. Denna term har egentligen en bredare betydelse vilket beskrivs mer detaljerat i kapitel 2.

Den data som offentliga myndigheter innehar är heterogen och kan exempelvis röra områden som väder, geografi, statistik, trafik, och offentlig budget (Janssen, Charalabidis & Zuiderwijk, 2012). Det sägs att vidareutnyttjande av öppna data kan resultera i positiva effekter såsom ökad insyn och ansvarsskyldighet inom offentlig myndighet, mer deltagande och informerade medborgare, mer innovativa och effektiva offentliga tjänster samt ekonomisk tillväxt (Janssen et al., 2012; Ubaldi, 2013). Det har även publicerats rapporter där kvantifierbara effekter estimeras. Ett exempel återfinns i en studie av Carrara, Chan, Fischer och van Steenbergen (2015a) där det estimeras att den direkta marknadsstorleken för öppna data inom EU28+-länderna kommer nå 75,7 miljarder euro år 2020. I beräkningen av denna summa tas enbart hänsyn till nyttor som förverkligas i marknadstransaktioner. Därmed omfattas summan ej av faktorer såsom nya jobbopportuniteter, nya produkter och tjänster, besparad tid vid användning av produkter och tjänster baserade på öppna data samt ökad effektivitet inom offentlig sektor. Carrara et al. (2015a) poängterar dock att detta just är en prognos och att det i dagsläget saknas data för att i efterhand utvärdera kvantifierbara effekter.

Viktigt att påpeka i relation till nyttan med öppna data är att data i dess egenhet inte utgör något värde, utan att värde först skapas vid dess användning (Janssen et al., 2012). Med detta sagt behöver offentliga myndigheter göra mer än att tillhandahålla data för att säkerställa att öppna data utgör ett värde. För att åstadkomma detta påstår Ubaldi (2013) att de viktigaste uppgifterna för offentlig myndighet verkar vara att identifiera värdefulla data, att säkerställa och förbättra datakvalitet,

samt att främja vidareutnyttjande och efterfrågan av data. Vidare menar Ubaldi (2013) att en förutsättning för detta är att det finns god förståelse för hur data skapas, skyddas, delas och vidareutnyttjas, men även för användarnas behov när det gäller innehåll och format. Detta diskuteras också av Janssen et al. (2012) som hävdar att det idag saknas insikter i användarnas behov och att detta kan vara en förklaring till det glapp som finns mellan potentiella fördelar och det faktiskt realiserade värdet med öppna data. Detta syns även inom forskning, då enbart få studier utforskat öppna data ur användarens perspektiv (Yang & Kankanhalli, 2013).

För att få insikt i användarens perspektiv är det viktigt att förstå vilka användarna är och hur data vidareutnyttjas. Tidigare litteratur tar upp flera olika potentiella användare och användningsområden. I en litteraturöversikt av Safarov, Meijer och Grimmelikhuijsen (2017) nämns exempelvis medborgare, företag, forskare, utvecklare, icke-statliga organisationer och journalister som potentiella användare. I en rapport av Statskontoret (2018) väljer man istället att skilja på *intresserad allmänhet*, *användare av dataset* och *avancerade användare*. Mångfalden i den potentiella användargruppen diskuteras av Susha, Grönlund och Janssen (2015) där de observerat att offentlig myndighet tenderar att använda en och samma strategi för att interagera med samtliga användare. Vidare menar de att en sådan strategi inte fungerar, eftersom olika användare har olika bakgrund, motiv samt använder olika typer av teknologi.

1.2 Syfte och frågeställning

Med bakgrund i att djupare insikter i öppna data-användares perspektiv behövs ämnar vi att undersöka vilka hinder som möts vid vidareutnyttjande av öppna data. Med tanke på den mångfalden som karakteriserar användare av öppna data väljer vi att begränsa undersökningen till de hinder som upplevs av personer som varit inblandade i utveckling av tjänster. Med detta sagt ställer vi följande forskningsfråga.

På vilket sätt upplever tjänsteutvecklare hinder vid vidareutnyttjande av öppna data?

Syftet är inte att få en komplett översikt över de hinder som möts, utan snarare få en djupare insikt i hur hinder som tidigare identifierats upplevs påverka utveckling av tjänster. Vi tror att djupare insikt i hur hinder upplevs kan uppdaga aspekter som är särskilt relevanta inom utveckling av tjänster och därmed utgöra underlag för vidare forskning.

1.3 Avgränsning

Inom ramen för denna uppsats används begreppet "öppna data" för att beskriva data som publiceras öppet av en offentlig myndighet. Definitionen av öppna data

(se kapitel 2) är strikt och krav ställs på bland annat format och licens. För att kunna problematisera aspekter runt vidareutnyttjande använder vi ett bredare perspektiv och diskuterar därför även data som inte helt förhåller sig till denna definition.

I uppsatsen är termen “vidareutnyttja” förekommande. Med detta avses att använda öppna data i något syfte. På engelska används ofta termerna “use” och “re-use”, men på svenska är termen vidareutnyttja förankrad i öppna data-sammanhang.

Med “tjänsteutvecklare” avses personer som varit inblandade i projekt där syftet varit att ta fram en tjänst där öppna data är en väsentlig komponent. Inget krav ställs på hur framgångsrik tjänsten är och därför omfattas även mindre projekt och prototyper.

Begreppet “hinder” används för att beskriva faktorer som på något sätt hämmar vidareutnyttjandet av öppna data. Dessa hinder kan utgöras av både sociala och tekniska aspekter så som exempelvis mänskligt beteende respektive mjukvara och nätverk.

I uppsatsen förekommer begreppen “användare” och “leverantör”. Med användare avses personer som direkt nyttjar öppna data, exempelvis i syfte att utveckla en tjänst eller att analysera data. Med begreppet användare avses däremot inte personer som indirekt nyttjar öppna data, exempelvis en person som använder en tjänst som i sin tur är baserad på öppna data. I de fall där personer som indirekt nyttjar öppna data avses används istället begreppet “slutanvändare”. Med leverantör avses en offentlig organisation eller en organisation som på uppdrag av offentlig myndighet publicerar öppna data.

1.4 Undersökningens struktur

Initialt gjordes en litteraturstudie för att få insikt i vilka hinder som i dagsläget hämmar både användare och leverantörer av öppna data. En redogörelse för dessa, tillsammans med en djupare beskrivning av begreppet öppna data finns i kapitel 2. För att ge struktur till empirisk datainsamling och analys valdes en tematisering över hinder. Denna tematisering beskrivs i kapitel 3. Det empiriska underlaget togs fram genom kvalitativa intervjuer och hur dessa genomförts redovisas i kapitel 4, följt av en beskrivning över resultatet i kapitel 5. Slutligen diskuteras resultatet mot den utvalda tematiseringen och relaterad forskning i kapitel 6 följt av en sammanfattning av de viktigaste resultaten.

2 Relaterad forskning

Inom ramen för detta avsnitt kommer först begreppet öppna data redas ut. Sedan följer en beskrivning av befintlig forskningen rörande hinder för öppna data, både ur leverantörers och användares perspektiv. Att ha förståelse för vad öppna data är, samt vilka hinder som står i vägen för vidareutveckling baserat på öppna data är en viktig förutsättning för att förstå kommande resultat och diskussion.

2.1 Öppna data

I kapitel 1 gavs en kortfattad beskrivning av vad öppna data är för något, samt hur begreppet används inom ramen för denna undersökning. Inom detta kapitel går vi djupare in på begreppet.

Den engelska motsvarigheten “open government data” kan delas in i två olika delar, “government data” och “open data” (Ubaldi, 2013). Den förstnämnda delen avser data som produceras eller uppdras av offentlig myndighet eller en organisation som styrs av offentlig myndighet (Ubaldi, 2013; Open Knowledge International, 2017b). Det finns inget internationellt erkänt begrepp för att referera till data från offentlig myndighet, men termer som förekommer är bland annat “public information”, “government information” och “public sector information” (Vickery, 2011). Den sistnämnda delen avser att data förhåller sig till definitionen av öppen (Open Knowledge International, 2017b). Definitionen av öppen omfattar flera aspekter (exempelvis rörande format och licens) och kan sammanfattas som att vem som helst ska kunna få åtkomst, använda, modifiera och distribuera vidare; men kan behöva vidta åtgärder för att bevara ursprung och öppenhet (Open Knowledge International, 2017a).

2.2 Vidareutnyttjande av öppna data

Som det beskrevs i kapitel 1 utgör inte öppna data något värde om den inte används (Janssen et al., 2012). I det här kapitlet ges en beskrivning på hur öppna data kan vidareutnyttjas, samt hur processen där öppna data transformeras till en tjänst kan se ut.

Ett sätt att betrakta hur öppna data kan vidareutnyttjas beskrivs av Davies (2010). Han särskiljer på fem olika processer: (1) *data till faktum*, att identifiera särskilda faktum i en datamängd; (2) *data till information*, att visa en representation av en eller flera datakällor, till exempel en visualisering eller en rapport; (3) *data till gränssnitt*, att möjliggöra interaktiv åtkomst där en eller flera datamängder kan utforskas; (4) *data till data*, att dela data i ett förändrat tillstånd, exempelvis genom att kombinera öppna data med annan data; (5) *data till tjänst*, att tillhandahålla en tjänst som

bygger på öppna data. Vidare beskriver Davies (2010) att flera av dessa processer kan förekomma i en och samma tillämpning av öppna data. Exempelvis skulle en tjänst (*data till tjänst*) kunna innehålla en del där data interaktivt kan utforskas (*data till gränssnitt*).

En annan studie relaterad till hur öppna data kan vidareutnyttjas har genomförts av Kuk och Davies (2011) där det undersöks hur öppna data kan transformeras till en tjänst. I denna studie beskrivs vidareutnyttjandet som en kedja där transformationer bildar nya artefakter: (1) från rådata till städade data som kan kombineras med andra datamängder; (2) från data som kan kombineras till mjukvara; (3) från mjukvara till delad förvaring av källkod; (4) från delad förvaring till webbaserad tjänst. Studien genomfördes på öppna data-relaterade evenemang i Storbritannien och kan därför nödvändigtvis inte generaliseras till andra kontext. Däremot ger det ändå insikt i hur denna transformationen från öppna data till tjänst kan gå till.

I en studie av Zuiderwijk och Janssen (2014) beskrivs öppna data som en process, men fokus ligger snarare på de aktiviteter som genomförs från att data skapas till att den vidareutnyttjas och någon form av feedback når den som tillgängliggjort datan. De aktiviteter som rör själva vidareutnyttjandet är följande: (1) data hittas via öppna data-portaler; (2) data analyseras, städas eller annoteras; (3) data bearbetas, exempelvis genom att visualisera eller kombinera med annan data.

2.3 Hinder för öppna data

Som tidigare nämnt är fokuset i denna undersökning att utforska vilka hinder som upplevs vid tjänsteutveckling baserat på öppna data. I detta kapitel presenteras relaterade studier där hinder för öppna data utforskats. Dessa omfattar hinder från både leverantörens och användarens perspektiv, där användaren avser någon som direkt nyttjar öppna data (se kapitel 1).

En studie på hög abstraktionsnivå i relation till hinder för öppna data har genomförts av Martin (2014). I denna studie betraktas utvecklingen av öppna data som en innovationsprocess, där första fasen beskrivs som ett nymodigt och lovande koncept och sista fasen innebär ett större genombrott med transformativa effekter i samhället. Martin (2014) argumenterar för att ett större genombrott ännu inte har skett och fokuserar på att förstå vilka hinder som står i vägen för en sådan utveckling, samt vilken påverkan dessa hinder har på praktik och riktlinjer.

Studien bygger på en enkätundersökning gjord i Storbritannien och resulterar i fem huvudsakliga fynd: (1) hinder för öppna data uppfattas vara omfattande, och de rör både sociala och tekniska faktorer; (2) hinder vid efterfrågan och vidareutnyttjande av öppna data uppfattas som övervägande; (3) strukturer inom offentlig sektor och institutioner på makronivå (som exempelvis lagstiftning) uppfattas utgöra ett

hinder; (4) brist på kulturella och finansiella resurser uppfattas vara ett stort hinder; (5) Det finns omstridda hinder, där åsikter om dem skiljer sig. Martin (2014) visar att hinder för öppna data verkar existera och beskriver grovt i vilka områden de manifesterar sig. Ett genomgående tema i studien är att han skiljer på hinder för implementation och hinder för vidareutnyttjande. Med implementation syftar han på att öppna data implementeras i offentliga myndigheter praktik, medan han med vidareutnyttjande menar att individer och organisationer skapar ekonomiska och sociala värden baserat på öppna data. Denna distinktion används för att strukturera kommande två underavsnitt, där hinder ur både leverantörens och användarens perspektiv diskuteras.

2.3.1 Från ett leverantörsperspektiv

Bland tidigare studier där hinder från leverantörens perspektiv utforskats finns både de som exklusivt fokuserar på leverantörens perspektiv (till exempel Barry & Bannister, 2014; Conradie & Choenni, 2014), men också de som tar en bredare ansats och även omfattar användarens perspektiv (till exempel Janssen et al., 2012; Zuiderwijk & Janssen, 2014; Berends, J., Carrara, W., & Vollers, H, 2017). I samtliga av dessa studier används olika kategorier eller teman för att diskutera de hinder som upptäcks.

I en studie av Zuiderwijk och Janssen (2014) avser de att undersöka vilka sociotekniska hinder som finns i det som de kallar öppna data-processen. De menar att öppna data kan beskrivas som en process där data (1) *skapas*; (2) *publiceras*; (3) *hittas*; (4) *används*; (5) *diskuteras*. Vidare argumenterar de för att denna process består av sociala aspekter (till exempel mänskligt beteende, organisationsstruktur, lagstiftning) och tekniska aspekter (till exempel mjukvara, nätverk) vars interaktion påverkar processens utfall. De hinder som identifieras kategoriseras under processens olika faser och hinder som upplevs av leverantören förekommer i alla faser förutom vid användning av öppna data. Exempel på konkreta hinder som diskuteras är bland annat att leverantörer inte har insikt i användarens behov och att det är svårt att säkerställa datakvalitet då krav på datakvalitet kan variera mellan olika användningsområden.

Ytterligare en studie som delvis berör hinder från leverantörens perspektiv har genomförts av Janssen et al. (2012). I denna studie utforskas både upplevda nyttor och upplevda hinder, för att slutligen presentera fem myter runt öppna data. De hinder som identifieras relateras till kategorierna: (1) *institutionella nivå*; (2) *komplexitet vid hantering av data*; (3) *användning och deltagande i öppna data-processen*; (4) *lagstiftning*; (5) *informationskvalitet*; (6) *tekniska nivå*. Den *institutionella nivå* består exklusivt av hinder som upplevs av leverantören och de menar att dessa hinder resulterar i ett motstånd till att gå mot en öppnare kultur. Exempel på hinder som diskuteras är att organisationer kan präglas av en riskavert kultur, att

det kan saknas resurser, samt att enhetliga processer för publicering av data saknas. Även kategorierna för *lagstiftning*, *informationskvalitet* och den *tekniska nivån* omfattas av hinder som kan upplevas av leverantören. Bland dessa nämns exempelvis att föråldrade och fragmenterade system kan försvåra publicerandet av data och att avsaknad av standarder, såsom för metadata, kan upplevas problematiskt.

Berends et al. (2017) tar en annorlunda ansats och diskuterar hinder under följande teman: (1) *politiska*; (2) *organisatoriska*; (3) *juridiska*; (4) *tekniska*; (5) *finansiella*; (6) *kännedom*. Samtliga av dessa kategorier omfattar hinder som kan upplevas av både leverantören och av användaren. Vidare hävdar de att *finansiella* hinder är de mest frekventa från leverantörens perspektiv och nämner exempelvis att leverantörer som historiskt sett har sålt data har svårt att se fördelarna med att släppa data gratis (eller till marginella kostnader). Utöver detta nämner de även att det kan krävas extra finansiering för att implementera riktlinjer för öppna data, exempelvis för infrastruktur, processer och ny personal. Under de övriga kategorierna nämns bland annat att det saknas teknisk kompetens, att det är svårt att förankra öppna data i den dagliga verksamheten, samt att öppna data-initiativ oftast drivs på nationell nivå uppifrån och ner; samtidigt som offentlig myndighet på lägre nivå inte har en kulturell tradition av att dela med sig av data.

2.3.2 Från ett användarperspektiv

Som det nämndes i föregående avsnitt finns det flera exempel på studier som tar upp hinder från både leverantörens och användarens perspektiv (till exempel Janssen et al., 2012; Zuiderwijk & Janssen, 2014; Berends et al., 2017), men också de som enbart har ett fokus på användarens perspektiv (till exempel Zuiderwijk, Janssen, Choenni, Meijer & Alibaks, 2012; Statskontoret, 2018; Magalhães & Roseira, 2016).

Som tidigare nämnt använde Janssen et al. (2012) följande sex kategorier för att diskutera upptäckta hinder: (1) *institutionell nivå*; (2) *komplexitet vid hantering av data*; (3) *användning och deltagande i öppna data-processen*; (4) *lagstiftning*; (5) *informationskvalitet*; (6) *teknisk nivå*. En av de kategorier som enbart involverar hinder som användaren upplever är *komplexitet vid hantering av data*. Inom denna kategori nämns exempelvis att användaren kan uppleva att det är svårt att hitta lämpliga data, att det är oklart hur data potentiellt kan användas samt att det saknas verktyg som stöd för vidareutnyttjande av öppna data. Ytterligare en kategori som enbart omfattar hinder från användarens perspektiv är *användning och deltagande i öppna data-processen*. Inom ramen för denna kategori nämns bland annat att brist på kunskap och färdigheter för att kunna använda och förstå data kan utgöra ett hinder. Bland de övriga kategorierna nämns exempelvis att avsaknad av licens och att restriktiva licenser kan upplevas som problematiskt.

Berends et al. (2012) tar upp hinder från användarens perspektiv under samtliga av kategorierna: (1) *politiska*; (2) *organisatoriska*; (3) *juridiska*; (4) *tekniska*; (5) *finansiella*; (6) *kännedom*. De utgår främst från att användaren är ett företag inom privat sektor och diskuterar därför hur politiska, organisatoriska och finansiella faktorer inom det privata företaget kan stå i vägen för vidareutnyttjande av öppna data. Bland dessa hinder diskuteras exempelvis att beslutsfattare inom det privata företaget behöver vara övertygade om värdet med vidareutnyttjande av öppna data, samt att integration av öppna data i de operationella processerna kan kräva förändringar i befintlig organisationsstruktur. Enligt Berends et al. (2012) är *finansiella* och *juridiska* hinder de minst förekommande hindren bland användare, medan de främst förekommande berör brist på *kännedom* om öppna data. De menar på att potentiella användare (företag och privatpersoner) i stor utsträckning varken känner till att offentligt tillgängliga datamängder existerar eller vilket värde sådan data kan bidra med. Ytterligare hinder från användarens perspektiv som nämns är bland annat att användare kan vara motvilliga till att vidareutnyttja data om den licens som den publiceras under är otydlig, samt att kombination av data kräver en kompatibilitet mellan licenser, vilket kan vara problematiskt ur internationellt perspektiv där länder använder nationella licenser.

I likhet med Zuiderwijk och Janssen (2014) fokuserar Zuiderwijk et al. (2012) explicit på sociotekniska hinder. Skillnaden är att sistnämnda riktar in sig på användarens perspektiv, samt att målet är att ta fram en heltäckande överblick över hinder. De använder sig av två olika typer av tematiseringar för att strukturera tabeller över hinder. Initialt relateras hinder som identifierats i tidigare litteratur till följande tre kategorier: (1) *åtkomst till data*; (2) *användning av data*; (3) *insättning av data*. Hinder från litteratur kombineras sedan med sådana som upptäcks genom intervjuer och workshops. Samtliga av dessa relateras till följande tio kategorier: (1) *tillgänglighet och åtkomst*; (2) *upptäckbarhet*; (3) *användbarhet*; (4) *begriplighet*; (5) *kvalitet*; (6) *länkning och kombination av data*; (7) *jämförbarhet och kompatibilitet*; (8) *metadata*; (9) *interaktion med leverantör*; (10) *öppnande och uppladdning*. Exempel på konkreta hinder som nämns är att data enbart finns tillgängligt efter förfrågan, att data redan är bearbetad (i kontrast till att få åtkomst till ursprungsdata), att det inte finns något bra applikationsprogrammeringsgränssnitt (API), att data upplevs opålitlig samt att metadata ej beskriver hur data har samlats in.

Magalhães och Roseira (2016) har också utforskat vilka hinder som upplevs av användare men fokuserar specifikt på kommersiellt vidareutnyttjande. För att analysera hinder använder de ett ramverk bestående av följande kategorier: (1) *resurser*; (2) *åtkomst*; (3) *användbarhet*; (4) *interaktion*; (5) *lagstiftning*. Undersökningen resulterade exempelvis i insikter om att de flesta företagen verkar föredra data på fil framför API:er (för ökad kontroll och åtkomst), att man upplever att datakvalitet är

bristfällig, att man spenderar mycket tid på att förstå licenser samt att man önskar en djupare interaktion med leverantörer.

Hinder från användarens perspektiv är även fokuset för en rapport framtagen av Statskontoret (2018). I studien utvärderas svenska myndigheter utefter teman som *tillgänglighet*, *kvalitet*, *användbarhet* och *interoperabilitet*. Utvärderingsmodellen består delvis av allmänna aspekter, men också aspekter som anses vara särskilt relevanta för tre olika användargrupper: (1) *intresserad allmänhet*; (2) *användare av dataset*; och (3) *avancerade användare*. Förstnämnda gruppen beskrivs som personer med enklare behov, såsom studenter, journalister och politiker. Den andra gruppen beskrivs som personer som kan vara i behov av större datamängder för vidare bearbetning, exempelvis forskare, utredare och statistiker. Den sistnämnda gruppen beskrivs som att den främst består av programmerare och tjänsteutvecklare, vilket ligger nära den avgränsning som gjorts i vår undersökning. De hinder som identifierats för intresserad allmänhet omfattar att myndigheters webbplatser sällan är anpassade för mobila enheter, att det är svårt att förstå data eftersom den inte sätts i ett sammanhang, samt att det ibland saknas möjlighet att ta del av data direkt på webbplatsen, exempelvis genom interaktiv åtkomst med visualisering och filtreringsmöjligheter. För användare av dataset utgjordes de största hindren av bristande dokumentation och avsaknad av kontaktuppgifter till personer med kunskap om specifika datamängder. Den mest utsatta gruppen var avancerade användare, vilket enligt Statskontoret (2018) beror på att API:er inte erbjuds i stor utsträckning men också för att standarder för data och metadata saknas.

Med den relaterade forskningen presenterad så kommer vi i nästa kapitel att diskutera den taxonomi som valts ut för att strukturera denna studie.

3 Teoretiskt ramverk

Som tidigare beskrivet ämnar vi att inom ramen för denna uppsats att besvara hur hinder upplevs vid tjänsteutveckling baserat på öppna data. För att kunna besvara frågan används en taxonomi över olika typer av hinder för att strukturera delar av studien. I kapitel 2 beskrevs flera exempel på sådana taxonomier.

Den taxonomi vi valt att använda är baserad på en litteraturöversikt av Kučera (2017). I litteraturöversikten sammanställs en lista med hinder från 11 olika verk. Samtliga av verken är publicerade från 2011 till och med 2017. Verken identifierades genom att söka efter hinder för både vidareutnyttjande och publicering av öppna data. Hinder som förekom i flera av verken, eller som var väldigt lika varandra gavs en gemensam beteckning och kombinerades sedan i en lista med de som enbart förekom en gång. Kučera (2017) fortsatte sedan med att kombinera hinder på hög detaljnivå till mer generaliserade hinder. Slutligen grupperades olika generaliserade hinder till kategorier på ännu högre nivå.

Taxomin från Kučera (2017) innehåller både hinder för vidareutnyttjande och för publicering av data. Eftersom vår studie är begränsad till att enbart omfatta hinder vid vidareutnyttjande har vi omdefinierat och exkluderat vissa av kategorierna. De kategorier som exkluderas är: (1) *negativ och oönskad påverkan*; (2) *nyttor och positiv påverkan*; (3) *koordination, organisation, processer och drift*; (4) *resurser, kostnad och finansiering*; (5) *ledarskap, riktlinjer och strategi*. Anledningen till att dessa kategorier har exkluderats är att dessa typer av hinder främst verkar uppstå hos leverantören medan denna studie är begränsad till att omfatta hinder som upplevs av användaren. Vi tror däremot att dessa hinder kan vara rotorsaken till att andra hinder upplevs av användaren. För att illustrera detta kan en av aspekterna som nämns för kategorien *negativ och oönskad påverkan* användas, i vilken Kučera (2017) resonerar med grund i Barry och Bannister (2014) att statstjänstemän kan känna en oro för att publicera data eftersom att eventuella fel i data kan ställa leverantören i dålig dager. Denna oro är antagligen inget som en användare av öppna data känner men om denna oro förhindrar att vissa data publiceras kan det även leda till att användare inte hittar den data de behöver.

Vidare omdefinieras två kategorier: (1) *intressenters attityd och interaktion*; (2) *lagstiftning, licensiering och integritet*. Anledning till att de omdefinieras är att de berör aspekter som upplevs av både leverantören och av användaren. Vi har därför anpassat dem så att de enbart berör aspekter som direkt har med användaren att göra.

Anledningen till att vi valt att använda den taxonomi som utvecklats av Kučera (2017) är dels att det är en litteraturöversikt och bygger därmed på insikter från flera tidigare verk. Vidare bygger den på verk fram till 2017, vilket innebär att den

potentiellt kan omfatta nya insikter som inte existerar i taxonomier som tagits fram tidigare. Trots att vi väljer att använda denna taxonomi som teoretiskt ramverk finns det även motargument. Till exempel det faktum att den är utformad för att betrakta hinder både för leverantörer och för användare. Att vi väljer att exkludera delar som verkar mest relevanta för leverantörer kan resultera i att viktiga perspektiv missas. Vidare bygger taxonomin på en enskild persons tolkning av tidigare litteratur. Trots motargumenten väljer vi ändå att använda Kučeras (2017) taxonomi, främst för att den bygger på verk fram till 2017.

Kategorier	Kortfattad beskrivning
Kvalitet och åtkomst för data och metadata	Åtkomst till data och metadata, samt kvalitetsaspekter i data och metadata.
Interaktion	Samspelet mellan leverantör och användare.
Kunskap och färdigheter	Avsaknad av kunskap och färdigheter för att kunna använda öppna data.
Teknologi och verktyg	Tekniska aspekter, såsom leverantörens infrastruktur. Brist på verktyg som stöd för vidareutnyttjande av öppna data.
Licensiering	Avsaknad av licens, eller olämplig licens som hämmar vidareutnyttjandet av öppna data.
Bygger på kategorisering av Kučera (2017), som utvecklats baserat på följande källor: Huijboom och Van den Broek (2011); Janssen et al. (2012); Martin, Foulonneau, Turki, och Ihadjadene (2013); Ubaldi (2013); Barry och Bannister (2014); Conradie och Choenni (2014), Carrara, Fischer, och van Steenbergen (2015b), Carrara, Fischer och van Steenbergen (2015c), Carrara, Fischer, Oudkerk, van Steenbergen och Tinholt (2015d), Carrara, Nieuwenhuis och Vollers (2016), Berends et al. (2017)	

Tabell 3.1. Utvalda kategorier av hinder

I tabell 1.1 visas en överblick av de slutgiltiga kategorierna tillsammans med kortfattade beskrivningar. I kommande delkapitel beskrivs de utvalda kategorierna från Kučera (2017) mer detaljerat, tillsammans med anpassningar som gjorts för att kategorierna ska vara lämpliga för den avgränsning som görs inom denna studie.

3.1 Kvalitet och åtkomst för data och metadata

Beskrivningen av *kvalitet och åtkomst för data och metadata* består främst av tre olika delar. Den första delen grundas på Carrara et al. (2016) och World Wide Web Foundation (2015). Denna del handlar om att utbudet på öppna data fortfarande är begränsat och att majoriteten av den öppna data som finns inte förhåller sig till den strikta definitionen av öppen. Den andra delen handlar om att den öppna data som finns publicerad nödvändigtvis inte är åtkomlig. Beskrivningen av denna del bygger exempelvis på Carrara et al. (2016) och Janssen et al. (2012). Kučera (2017) diskuterar att användare upplever att det är svårt att hitta den data som de behöver men också att kostnader för åtkomst till data kan begränsa vidareutnyttjande. Vidare diskuteras även att inte all öppna data publiceras i maskinläsbart format. Den sista delen berör mer kvalitetsrelaterade aspekter, både för data i sig och dess tillhörande metadata. Denna del grundas exempelvis på Berends et al. (2017) och Janssen et al. (2012). Ingen definition ges på exakt vilka aspekter av datakvalitet som omfattas; men ofullständighet, felaktighet och avsaknad av data eller metadata nämns som exempel. Vidare diskuteras även att data och metadataas mångfaldiga karaktär i kombination med brist på standardisering är ett vanligt förekommande hinder i den analyserade litteraturen.

Sammanfattningsvis definieras denna kategori till att handla om hinder runt åtkomst till data och metadata, antingen genom att den inte existerar, att den är svår att hitta eller att den inte är åtkomlig på grund av olämpligt format. Vidare omfattas även brister i kvalitet, både på data och tillhörande metadata.

3.2 Interaktion

I diskussionen om *intressenters attityd och interaktion* omfattas attityd och inställning hos bland annat leverantören och användaren. Dessa aspekter grundas på exempelvis Carrara et al. (2015b) och Janssen et al. (2012). I samband med detta diskuteras dels att publicerandet av öppna data kan begränsas av inställningen hos statstjänstemän och leverantörens rådande organisationskultur. Vidare diskuteras även att avsaknad av incitament och tidsbrist kan begränsa användaren. Kategorin utgörs även av aspekter som rör själva interaktionen mellan olika intressenter. Dessa aspekter grundas exempelvis på Carrara et al. (2015c), Ubaldi (2013) och Berends et al. (2017). Kučera (2017) menar att kommunikationen mellan olika intressenter kan vara icke-existerande eller bristfällig och att detta bland annat kan resultera i att myndigheter inte prioriterar publicering av den data som användare behöver.

För att sammanfatta definieras denna kategori till att handla om hur interaktion mellan användare och leverantör kan hämma vidareutnyttjande. Attityd, inställning och organisationskultur hos leverantören omfattas inte, då det inte bör vara ett

direkt hinder för användaren. Dessa aspekter bör istället kunna orsaka hinder som manifesteras på andra sätt, exempelvis genom att data ej publiceras. Vidare argumenterar vi för att brist på incitament är något som främst utgör ett hinder för tjänsteutvecklare som ännu inte anammat öppna data. Slutligen anses brist på tid vara en för marginell aspekt för att motivera ett fokus på användarens attityd inom denna kategori. Eftersom kategorin begränsas till att enbart beröra interaktionen mellan användare och leverantör benämns istället kategorin som interaktion.

3.3 Kunskap och färdigheter

Inom denna kategori beskriver Kučera (2017) att brist på kunskap och färdigheter kan hindra både leverans och vidareutnyttjande av öppna data. Dessa argument grundas exempelvis på Janssen et al. (2012), Carrara et al. (2015b) och Carrara et al. (2016). Vidare hävdas att en förutsättning för att kunna vidareutnyttja öppna data är att användaren har viss kompetens, exempelvis av teknisk karaktär men också inom statistik, företagande och den domän som datan berör. Slutligen diskuteras det att brist på tekniska färdigheter hos leverantören kan begränsa själva publicerandet av öppna data.

Sammanfattningsvis definieras denna kategori till att handla om hur brist på kunskaper och färdigheter hos användaren kan hämma vidareutnyttjandet av öppna data. Brister i kunskap och färdigheter hos leverantören omfattas inte, då detta istället bör orsaka andra former av hinder, såsom brister i datakvalitet, vilka direkt skulle kunna påverka användaren.

3.4 Teknologi och verktyg

I beskrivningen av *teknologi och verktyg* omfattas tekniska aspekter och brist på verktyg och support. Dessa aspekter bygger på exempelvis Huijboom och Van den Broek (2011), Janssen et al. (2012) och Ubaldi (2013). Kučera (2017) menar på att vidareutveckling kan hämmas av tekniska aspekter, såsom utdaterade och fragmenterade applikationer hos leverantören eller att dagens nätverksteknologi inte är tillräcklig. Vidare menar han även att brist på verktyg och support kan begränsa vidareutnyttjande.

För att sammanfatta så definieras denna kategori till att omfatta tekniska aspekter, både i allmänhet och runt leverantörens infrastruktur. Kučera (2017) klassificerade tekniska problem med öppna data-portaler under en annan kategori men vi väljer istället att inkludera det inom denna. Vidare omfattas även hinder som kan upplevas på grund av brist på verktyg och stöd för vidareutnyttjande, såsom verktyg för bearbetning och visualisering.

3.5 Licensiering

I samband med kategorin *lagstiftning, licensiering och integritet* beskrivs tre olika aspekter. Dels handlar den om att utbud av öppna data kan vara begränsat på grund av att nationell lagstiftning hindrar att vissa data släpps. Kučera (2017) menar på att det kan röra data som kan knytas till individen och att anonymisering av sådan data är problematisk. Detta resonemang grundar han på exempelvis Ubaldi (2013) och Carrara et al. (2015b). Vidare omfattar kategorin problematik rörande den licens som data publiceras under. Denna del baseras exempelvis på Janssen et al. (2012) och handlar om att licenser som säkerställer rättighet för vidareutnyttjande kan saknas, eller att de kan vara alldeles för restriktiva. Slutligen diskuteras även att nationell lagstiftning som stöttar publicerandet av öppna data kan saknas eller vara för ospecifik, vilket grundas på Berends et al. (2017).

Sammanfattningsvis definierar vi denna kategori till att enbart handla om hinder som upplevs i samband med att licenser saknas eller upplevs vara för restriktiva. I vår definition inkluderas inte hinder som rör hur brister på eller otydligheter kring nationell lagstiftning kan begränsa publicerandet av öppna data. Vi argumenterar för att denna typ av hinder främst påverkar leverantören, även om de kan vara rotationsorsaken till andra hinder som användaren upplever; exempelvis att de har problem med att hitta data som de behöver. Till följd av förminskningen i kategorins omfång benämns denna kategori istället som licensiering.

4 Metod

Inom detta avsnitt diskuteras och motiveras studiens karaktär, genomförande av datainsamling och analys samt vilka forskningsetiska aspekter som hänsyn tagits till.

4.1 Studiens karaktär

Tanken med studien är som tidigare nämnt att komma fram till vad tjänsteutvecklare har för uppfattningar och erfarenheter om hinder som möts vid utveckling av tjänster baserade på öppna data. Syftet med studien är inte att kartlägga alla hinder som förekommer vid tjänsteutveckling på öppna data, utan att utforska hur de upplevs. I och med att vi är intresserade av människors erfarenheter och upplevelser så är det enligt Patel och Davidson (2011) svårt att på förhand veta vilka svar som informanter kan tänkas ge. Av denna anledning har en kvalitativ studie genom intervjuer valts framför en kvantitativ metod.

4.2 Genomförande av intervju

Datainsamlingen är grundad på semistrukturerade intervjuer. Intervjufrågorna utformades med en låg grad av strukturering, alltså att frågorna till stor del var öppna. Patel och Davidsson (2011) menar att detta är ett bra sätt för att ge informanterna större utrymme till att formulera svaren med egna ord, vilket lämpade sig bra för vår undersökning. Intervjufrågorna konstruerades med en viss bestämd ordning, vilket Patel och Davidsson (2011) benämner som att det finns en viss grad av standardisering. Med en viss bestämd ordning menar vi att det fanns en övergripande struktur som utgjordes av inledande frågor, huvudsakliga frågor och avslutande frågor. De huvudsakliga frågorna konstruerades efter de kategorier som definierats i kapitel 3. Dessa frågor hade en lägre grad av standardisering då någon viss ordning på kategorierna inte var betydelsefull. De inledande frågorna handlade om informantens bakgrund, vilket var viktigt för att säkerställa att informanten hade rätt kompetens. Frågorna i den avslutande delen inkluderades för att fånga upp eventuella områden och aspekter som missats. Slutligen sammanställdes samtliga intervjufrågor i två manus, ett på svenska och ett på engelska (se kapitel 8).

Eftersom att intervjuare och informanter befann sig i olika länder så sköttes intervjuerna virtuellt. En intervjuare skötte inspelning och anteckning av intervjun medan den andra intervjuaren ställde intervjufrågorna. Att anteckna intervjuerna ansågs nödvändigt för att dokumentera en första uppfattning på det som informanterna framfört. Intervjuerna som maximalt tog 60 minuter röstinspelades, dels i syfte att kunna förenkla analysarbetet men också i syfte att i efterhand kunna säkerställa att anteckningarna var korrekta. Anteckningarna fungerade även som ett slags stöd för att dokumentera intervjun utifall att respondenten inte godkände inspelning

av intervjun. Detta blev dock inte fallet, utan samtliga informanter gav ett medgivande till ljudinspelning.

För att säkerställa att intervjufrågorna var av god validitet, det vill säga att frågorna täcker det vi avser att undersöka, så genomfördes en pilotintervju. Detta är enligt Bell och Waters (2015) bra för att komma fram till vilka frågor, ifall några, som behöver omformuleras. Pilotintervjun ledde bland annat till insikten om att en fråga var vagt formulerad, vilket ledde till att frågan formulerades om. Eftersom ingen markant förändring i manus gjordes så har material från pilotintervjun inkluderats i resultatet. Efter pilotintervjun genomfördes ytterligare fyra intervjuer med övriga informanter. Under intervjuernas gång ställdes även stängda följdfrågor som inte fanns nedskrivna i manuset för att tydliggöra och exemplifiera saker som sagts, men också för att leda tillbaka informanten till det huvudsakliga temat. Vidare ställdes de huvudsakliga frågorna i kontexten av ett enskilt projekt som informanten arbetat med, men åsikter och erfarenhet från andra projekt diskuterades också.

Vad gäller forskningsetiska aspekter så har varje intervju påbörjats genom att presentera studiens syfte, deltagarnas anonymitet i studien, deltagarnas frihet till att avbryta eller att inte svara på vissa frågor och information kring vilka som har tillgång till det inspelade materialet. Inför intervjuerna blev informanterna tillfrågade om de godkände att intervjuerna spelades in. Medgivande gavs muntligt, eftersom att intervjuer hölls virtuellt.

4.3 Urval

Då det inte har förekommit någon begränsning på vilken typ av tjänster som utvecklats eller vilken typ av data som nyttjats utgörs den totala populationen av samtliga personer som varit inblandade i projekt, där syftet varit att ta fram en tjänst baserad på öppna data. Då vi initialt saknade kontakter inom området har vi inte slumpmässigt kunnat välja informanter, utan har istället gått efter tillgänglighet. Informanter identifierades via ett led av kontakter som till en början tillhandahölls av vår handledare. Genom dessa kontakter identifierades informanter både i Sverige och Tyskland, vilket var lämpligt då vi befann oss i dessa länder. En intressant aspekt i relation till detta är att länderna fått liknande poängsättning i en rapport från World Wide Web Foundation (2017). Rapporten beskriver bland annat hur långt olika länder har kommit med sina öppna data-initiativ. Poängsättningen grundas i flera olika mätvärden. Sverige och Tyskland skiljer sig inom vissa mätvärden, men i den totala poängsättningen är länderna jämna.

Nedan följer en presentation av informanterna som deltagit i undersökningen.

Informant A: Informantens relation till öppna data är bland annat utveckling av en app med flera tusen unika dagliga användare. Informanten har också varit delaktig i

flera projekt tillsammans med företag, där både öppna data och kommersiella data varit inblandat. Informanten är yrkesverksam i Sverige.

Informant B: Informantens första kontakt med öppna data var på ett Hackathon, där informanten påbörjade utvecklingen av en app som idag har en användarbas på flera tusen användare. Informanten har även varit inblandad i fler projekt där öppna data varit en viktig komponent. Informanten är yrkesverksam i Sverige.

Informant C: Informanten är en del av ett mindre community med 10 till 15 regelbundna deltagare. Detta community drivs i en tysk stad med ca 300 000 invånare och man arbetar främst med data på kommunal nivå. De organiserar bland annat workshops där de bland annat bygger applikationer baserade på öppna data.

Informant D: Informanten är idag med i ett community som jobbar med att organisera workshops där de bland annat arbetar med att ta fram projekt för att visa värdet med öppna data. Detta community drivs också i en tysk stad med ca 300 000 invånare (dock inte samma stad som *Informant C*).

Informant E: Informanten har varit delaktig i samma community som *Informant C*. De har dock inte uteslutande arbetat med samma projekt. Informantens roll i detta community har främst varit att programmera och att organisera Hackathons.

Undersökningen skulle kunna ha gynnats av fler informanter, men på grund av brist på tid har antalet informanter begränsats till fem. Under intervju med *Informant D* uppstod tekniska problem, vilket resulterade i att intervjun blev kortare än planerat. Samtliga frågor besvarades, men svaren blev inte lika utvecklade som under övriga intervjuer. Vidare kan det finnas en viss partiskhet i urval av informanter. Dels för att vi på förväg kände till att *Informant D* har en god relation med leverantören, men också för att *Informant A* varit aktiv inom en specifik datadomän länge och har därför god relation med leverantören av data.

4.4 Analys

Efter att intervjuer genomförts påbörjades arbete med att transkribera intervju-materialet. Detta för att det ansågs vara viktigt att arbeta med transkribering när intervjun fortfarande var färsk i minnet. Så snart transkribering av intervju var färdig påbörjades analys. Intervjuerna skedde löpande under tre veckor, vilket även förde med sig att analys varvades med att nya intervjuer hölls.

Analys av intervjumaterial gjordes genom att relatera olika utsagor i transkriberingarna till kategorierna i det teoretiska ramverket. Kopplingen mellan utsaga och kategori gjordes genom att kommentera det transkriberade materialet i ett ordbehandlingsprogram. Detta genomfördes individuellt och sedan diskuterades vilka utsagor som identifierats och huruvida de skulle inkluderas i resultatet eller ej. Vidare

sammanställdes utvalda utsagor från samtliga intervjuer i ett gemensamt dokument. Slutligen grupperades utsagorna för att identifiera underliggande teman inom respektive kategori från det teoretiska ramverket. Detta gjordes genom att skriva ut de utvalda utsagorna på papper, för att sedan gruppera dem för hand. Gruppering gjordes individuellt och följdes av en diskussion där grupperingar jämfördes. Vissa av grupperingarna rörde teman som inte var relaterat till forskningsfrågan och togs därför ej med i resultatet.

Anledningen till att stora delar av analysen gjordes individuellt var att förhindra att den ena personen skulle påverka den andras tolkning, men också för att inte begränsa oss till en enskild persons tolkning. Slutligen förde det även med sig att vi fann ett större förtroende för de fall där samma tolkning gjorts.

5 Resultat

Under vår undersökning har vi tagit del av hur olika aspekter har, eller inte har hämmat utvecklingen av tjänster baserade på öppna data. Vidare har vi även fått höra om informanternas åsikter kring hinder från tidigare litteratur. I detta avsnitt presenteras resultatet, strukturerat utefter de kategorier som beskrevs i kapitel 3.

Eftersom undersökningen genomförts med informanter från både Sverige och Tyskland är intervjumaterialet på både svenska och engelska. För att undvika förvanskning av informanternas utsagor har citat inte översatts till svenska.

5.1 Kvalitet och åtkomst för data och metadata

I relation till kvalitet på data uttryckte tre av informanterna att de var förhållandevis nöjda inom de projekt som huvudsakligen diskuterats. En av dem menade på att de inte ställde alls höga krav på kvalitet inom projektet, utan snarare nöjde sig med den kvalitet som fanns tillgänglig.

Samtidigt hade flera av informanterna negativa erfarenheter också. Exempel som nämndes var att data varit strukturerad på olämpliga sätt, vilket har lett till svårigheter såsom att visualiseringsprogram har kraschat och att man behövt formatera om data manuellt. Ytterligare ett exempel som togs upp var att det förekommit rena fel i den data som vidareutnyttjats. De kunde baserat på egen erfarenhet konstatera att värdena var orealistiska. En informant har varit med om att de beslutat att inte genomföra projekt då de till följd av fel inte kunde lita på den data som tillhandahållits.

“So yeah that's definitely, we've decided to not do projects because we thought no... uhm... the weather data isn't good enough. We'll do something else with it.” (Informant C)

Samma informant berättade även att de satt upp sensorer för att samla in data på egen hand, just för att de inte kunde lita på den som myndigheten tillhandahöll.

Ytterligare en intressant aspekt som dök upp var att en informant tycker att fel i data kan accepteras så länge som konkurrerande tjänster har samma brist. Motivet till denna inställning är att slutanvändarna av tjänsten förmodligen inser att bristen inte beror på tredjepartsutvecklaren, utan på leverantörerna av den data som tjänsten baseras på.

“Så länge att den kvaliteten som vi har vi vårt data, i vår app, inte skiljer sig väsentligt från, från data som finns i konkurrerande appar eller andra appar man kan använda. Då gör inte mig nånting att datat är av

dålig kvalitet, [...] Då faller liksom ingen... Ingen skuld faller på oss.”
(Informant B)

Vidare diskuterade två av informanterna att krav på datakvalitet är relativ. Den ena menade på att data av lägre kvalitet kan vara acceptabelt i mindre projekt, men att det blir extremt viktigt i tjänster med många användare. Den andra informanten tog inte explicit upp projektets omfattning som den avgörande faktorn, utan snarare vad det är för typ av projekt.

“The easiest part is to say. Well, the quality is not enough. And bla bla bla... But honestly there’s not such a thing. The quality of the data is very stuck to your project. So, one project I need real-time data, very granular and with a lot of metadata. And on the other project I just need a CSV-file” (Informant D)

Vad gäller kvaliteten på metadata så var det bara en informant som explicit uttryckte en problematik. Informanten har upplevt att det ofta finns information om när data samlats in men att det brukar saknas information om kontext och hur pålitlig datan är. Som exempel tog informanten upp data över temperaturer, där informanten erfarit att det saknas information om var mätningarna genomförts, samt om de gjorts inomhus eller utomhus.

Flera av informanterna upplevde att det var enkelt att hitta den data de behövde, förutsatt att den fanns publicerad. Två informanter nämnde dock att det varit problematiskt i vissa fall där likartade data finns utspridd på flera olika hemsidor. Som exempel nämndes trafikdata, där data som behövdes publicerades av många olika trafikbolag. Ytterligare en aspekt som nämndes i relation till att data inte publiceras på ett centraliserat sätt var att en informant poängterade att det är fördelaktigt att ha insikt i vilka offentliga myndigheter som finns och vad deras ansvarsområden är.

Däremot så hade flertalet informanter erfarit att data som de behöver inte finns publicerad som öppna data. Något intressant i samband med detta är att de flesta hittat andra lösningar runt det, exempelvis genom att med hjälp av mjukvara läsa av information från en myndighets hemsida. Detta sågs dock inte som någon hållbar lösning. En informant uttryckte sig såhär:

“[...] That we had to scrape it... has been... that’s always kind of a thing where you say “No, we don’t really wanna do that”, because it means a lot of maintenance work” (Informant C)

Vidare upplevde flera informanter att det var svårt att få leverantörer att publicera specifika data som inte redan finns tillgänglig. En informant uttryckte även att det fortfarande finns förhållandevis lite öppna data, samtidigt som en annan informant menade att leverantörer borde fokusera på att publicera mer data, istället för att

fokusera på kvalitet. Två informanter hade en del insikt i leverantörernas organisationer och poängterade att en orsak till att vissa data inte publiceras är att delar av den tekniska infrastrukturen är föråldrad och att de inte kan lyfta ut data därifrån.

Även format på data har av flera informanter upplevts vara ett hinder. I dessa fall har det handlat om att öppna data publiceras i ett format som inte är maskinläsbart. Två av informanterna har valt att inte påbörja eller fortsätta projekt då data som behövdes inte fanns tillgängliga i maskinläsbart format. En av informanterna upplever att efterfrågad data ofta finns maskinläsbart, medan mindre eftertraktade data tillhandahålls i icke-maskinläsbara format. Slutligen var det en informant som tyckte att format på data får allt för mycket uppmärksamhet och att det egentligen är viktigare att prata om interoperabilitet.

“Men det som jag tycker är det viktigaste när man funderar på att öppna upp och kanske göra något nytt, det är ju snarare, vad heter det, interoperabilitet mellan data, så att olika dataset kan kombineras med varandra. Att det snarare är den delen är, som man kanske ska lägga krut på istället, och då kanske ett API kan vara i en excellista, och ett annat kan vara i en JSON, eller, det spelar ingen roll men det kan åtminstone kombineras ihop.” (Informant A)

5.2 Interaktion

Angående interaktion med leverantörer uttryckte tre av informanterna att de är ganska nöjda och upplever att de har en god relation. Exempelvis nämnde en informant att leverantören oftast är bra på att informera vid större förändringar i API:er. Ytterligare en informant upplevde att leverantören alltid visat intresse för tankar och åsikter. Samma informant påpekade dock att relationen med leverantören kanske inte är jämförbar, då informanten nyttjat offentlig data en längre tid, även längre än vissa leverantörers öppna data-initiativ existerat.

Samtidigt som de flesta överlag hade positiva upplevelser så kom det även upp negativa historier. I ett projekt som två av informanterna varit delaktiga i hade de varit i kontakt med leverantör som lovade att de skulle få tillgång till vissa data (som ej fanns publicerad i någon öppna data-portal) vid ett specifikt tillfälle. I samband med detta tillfälle arrangerade de ett hackathon där de tänkt att bygga en produkt baserat på denna data. Bara några dagar innan ändrade sig leverantören och nekade åtkomst till datan. Istället fick de lämna in en skriftlig ansökan och fick först tillgång långt senare då datat inte var intressant längre. En av informanterna uttryckte sig på följande sätt om denna händelse:

“Because that was kind of the like our event that we wanted to do. To get the data live and hack on it. Then something like in the last few

days coming up the [specifikt tillfälle]. Something changed there. Someone higher up in this agency said, “No, we can't do that anymore”. [...] That of course takes a lot of motivation of it as well. Especially when you have the frustrating experience of talking to someone and making kind of like an arrangement, and then suddenly it doesn't, you know, seem important.” (Informant C)

Ytterligare en annan informant berättade om en negativ upplevelse där man efterfrågat publicering av vissa data i leverantörens API:er. Informanten visste med sig att denna data fanns tillgänglig då den visades i en av leverantörens egna tjänster. Begäran nekades och efter att informanten byggt en egen lösning publicerades datan via deras API:er trots allt.

“När jag ringde första gången då hade dom inte bestämt sig för att släppa det än, och då sa dem att de inte har några planer på att släppa det. Vi byggde och sen så släppte dom den ändå. Det var ju pengar i sjön kan man tycka.” (Informant B)

Under intervjuerna förekom det även fler negativa upplevelser, exempelvis att det är svårt att påverka hur leverantör publicerar sin data, att leverantören inte alltid informerar vid förändringar i API:er, samt att förbättringsförslag till leverantörens öppna data-portal nekades. I relation till kommunikation från leverantör och förbättringsförslag poängterade en informant hur viktigt det är att som användare av en öppna data-portal behöva registrera sig. Detta för att möjliggöra att en kommunikationskanal faktiskt existerar.

En av informanterna upplevde att öppna data inte alltid är förankrat i hela leverantörens organisation. Som exempel gavs att en slutanvändare till informantens tjänst kontaktat leverantören för ett fel som inte hade med själva tjänsten att göra. Slut användaren blev då uppmanad att använda leverantörens officiella tjänst istället för den som informanten utvecklat. I relation till detta verkade också flera av informanterna uppleva att det kan finnas två sidor av leverantörens organisation, en som förstår vad öppna data handlar om och en som inte alls är engagerad. En informant uttryckte sig på följande sätt:

“Of course we have a good relationship though [...] the public administration that we are in dialogue with, they are the good guy. So they want it, but they know how hard it is and ehm, we are not addressing them and pushing them, we are addressing the other one who are not in dialogue of openness and all these things, so.” (Informant D)

Två av informanterna verkar även uppleva att vissa leverantörer, eller vissa personer inom leverantörens organisation ser på öppna data som något de bara vill få överstökad. En av dessa uttryckte sig på följande sätt:

“Jag tror inte att alla har förståelse för data. Det kan jag uppleva när man kommer lite utanför den här innersta kretsen där alla vet exakt vad dom pratar om [...] för en del organisationer på det sättet så blir det mer som en, öppen data är mer som en liten checkbox som är ganska nice att ha klickat för. Och sen är man ganska nöjd du vet, så man behöver inte ta det vidare från det.” (Informant A)

Vidare upplever två informanter att det finns en form av osäkerhet i myndigheten som leverantör av data, att det inte finns någon garanti för att de kommer att fortsätta med underhåll av data. En av dem tror att detta kan vara en anledning till att inte företag börjat nyttja öppna data i större utsträckning.

5.3 Kunskap och färdigheter

Enbart en av informanterna uttryckte explicit att det funnits en brist på någon kunskap vid utvecklingen. I det fallet var det enbart programmerare som arbetade på projektet och de upplevde det problematiskt att visualisera data på intressanta sätt. Informanten uttryckte:

“So it certainly would’ve been cool if there were, would’ve been someone who know about demographics, who have studied sociology, or someone, yeah, who is more focused on politics, or journalist who know how to do the storytelling, like how to put it in a story, give it a context, and not just have the data visualized in some way.” (Informant E)

Det diskuterades även vilken kunskap som anses vara viktig. En av informanterna hade uppfattningen att all form av kunskap kunde bidra till deras projekt på något sätt, men poängterade samtidigt att programmering, statistik och design var mest viktigt. Det förekom också att informanter betraktade akademiska kunskaper som programmering, design och statistik som något sekundärt. En av dem menade istället att det viktigaste är att förstå människors behov medan en annan menade att kreativitet, eller att man har en idé är det allra viktigaste. Sistnämnda informant menade att saker som teknisk kompetens och förståelse för det område som datat rör snarare är ett nödvändigt ont.

Slutligen verkade en informant främst uppleva att det viktigaste var teknisk kompetens, såsom programmering och förståelse för strukturerade data. Ytterligare en informant poängterade att det är möjligt att genomföra projekt med enbart programmerare men att det inte är optimalt.

5.4 Teknologi och verktyg

I relation till de projekt som primärt diskuterades under intervjuerna upplevde två av informanterna att teknisk infrastruktur inte påverkat utvecklingen nämnvärt. Gemensamt för dessa två var att de inte använde sig av något API, utan fick åtkomst till data på fil. De menade att de i deras fall enbart laddade ner en fil för att integrera den i sina projekt och interagerade därmed inte direkt med leverantörens infrastruktur i större utsträckning. En annan informant har inom det projekt som huvudsakligen diskuterats varit med om att källkod behövde modifieras till följd av förändringar i API:er, såsom att accepterade format för tid ändrats.

Samma informant som varit med om att förändringar i API:er ställt till problem, upplevde i övrigt att det mestadels fungerat bra men poängterade också att det finns ett högt krav på tillgänglighet. I detta fall hämtades data direkt från leverantörens API och därför behövdes åtkomst till data dygnet runt. Ytterligare en informant diskuterade tillgänglighet på API:er och menar på att hög tillgänglighet är en förutsättning om leverantörer vill att människor ska bygga storskaliga tjänster baserade på data, istället för mindre projekt och prototyper.

“Att de förväntar sig att folk ska bygga tjänster på det, då måste det sitta i alla led egentligen, och det fungerar inte att ha kontorstider på drift, för att, kommer du upp i 40 000 unika användare per dag som genererar alltså, det är flera 100 000 frågor per dag som sker, ja det sker 24 timmar om dygnet ju. Så, men då är ju frågan om man ska förvänta sig att datat finns tillgängligt 24/7 eller inte [...] vilket jag säger är en förutsättning om man vill att folk bygger tjänster på det, men annars så är det ju fine ju” (Informant A)

Samtidigt poängterar informanten att det viktigaste är att det är tydligt vilka förutsättningar som gäller, där informanten också upplever att det finns en brist idag.

I relation till verktyg nämnde inte någon informant uttryckligen att de upplevt en brist. När vi ställde frågan berättade de flesta om olika verktyg som de använt, både i öppna data-portalen och som de använt som en del av sina tjänster. Exempel på verktyg som nämndes var dokumentation, datakataloger, bibliotek för JavaScript och verktyg för att konvertera data mellan olika format. En informant förväntade sig inte några verktyg från leverantören, medan en annan påpekade att verktyg antagligen inte är så viktigt om det finns en tydlig idé med vad man vill göra. Den sistnämnda informanten menar att verktyg för exempelvis visualisering och bearbetning skulle kunna vara värdefullt för den som behöver hitta inspiration till en idé, men är skeptisk till hur väl det skulle fungera för att involvera människor i öppna data. Samma person ifrågasätter även om ansvaret för att tillhandahålla verktyg för inspiration verkligen ligger på öppna data-portalen.

“Om man har de där verktygen, osv, det är möjligt att det triggar någon till att, att få en idé eller tanke på någonting. Eh, men, jag är lite osäker på om det är syftet med sån öppna data-portal egentligen, eller om det syftet ligger hos någon annan” (Informant A)

Vidare hade två av informanterna varit med om att data enbart publicerats i ett specifikt verktyg eller anpassats för en viss applikation, istället för att den publiceras i ett maskinläsbart format. Båda var kritiska till detta, varav den ena uttryckte sig på följande sätt:

“[...] de hade mycket av sin öppna data, ledde bara till en sida där man kunde se en karta och klicka runt i [...] Men då har de ju missat att det är ju inte dem som ska stå för innovationen, de ska ju stå för datat på ett sätt som gör innovatörers liv lätta, sen kan innovatörer som jag, och andra programmerare och entreprenörer låta maskiner bearbeta den här datan på ett enkelt sätt, och sen skapa vettiga gränssnitt mot slutanvändaren.” (Informant B)

En annan aspekt som diskuterades var API:er med möjligheten att välja ut specifika delmängder av data (vilket inom ramen för denna studie betraktas som ett verktyg). En informant var positivt inställd till detta och menade på att det som gjorde det möjligt att utveckla tjänsten var att det fanns ett bra API. Informanten menar att innovation ofta inte kommer ur rådata, utan snarare ur ett praktiskt och användbart API. Det förekom dock även en mer kritisk inställning till denna typ av API:er. En informant upplevde att rådata i vissa fall döljs bakom API:er med fördefinierade frågor, till exempel att hämta en resväg mellan två geografiska placeringar. Informanten menar att detta kan hämma innovation genom att leverantören i och med den fördefinierade frågan redan bestämt hur datat ska användas.

“Öppen data är också data som ligger bakom de här API:erna som är så här frågebaserade API:er [...] Problematiken med dem är att de oftast är tänkta att användas på ett specifikt sätt [...] Vilket det då gör det ganska så svårt att bygga tjänster som inte är tänkt att byggas exakt så som skaparna av det här API:et har tänkt att det ska vara” (Informant A)

Samma informant föredrog också rådata framför API:er för att undvika beroende till leverantörens tekniska infrastruktur.

5.5 Licensiering

Majoriteten av informanterna har inte upplevt att licens för den data de vidareutnyttjat har utgjort något hinder. När vi frågade en av dem hur licensen påverkat utvecklingen svarade informanten så här:

“Nej det har den absolut inte. Vi har i princip frågat, “Ja men får vi använda den här datan på det här sättet som vi har tänkt?” - När vi var på tävlingen och då har dem som har ägt datat sagt att “Ja, ja absolut. Det är precis det den är till för. Kör på bara.” Och så har man godkänt någon typ av villkor kanske” (Informant B)

Två av informanterna hade enbart blivit påverkade ur det perspektivet att de behövt källhänvisa till data, vilket de inte upplever som något problematiskt. En av dem uttryckte sig på följande sätt:

“So you only have to say that you got the data from their website and that’s how we, in our project we just put a little infobox and if you click on it, you can see that we get the data from [leverantör av data]. Yeah, and that’s kind of it.” (Informant E)

Under samtalet uttryckte sig ytterligare en informant oproblematiskt om källhänvisning. Det är dock oklart om detta är något som upplevts inom ramen för projektet som vi huvudsakligen samtalade om, eller om det var mer av en allmän åsikt. Informanten menade i alla fall på att det enda som behöver göras är att klistra in en text på en “Om oss”-sida eller liknande.

Två av informanterna hade däremot uppfattningen att licenser antagligen skulle varit mer problematiskt om de hade haft kommersiella intentioner med sina projekt. Inga av dessa hade dock haft några kommersiella intressen och upplevde därför att det fungerat bra.

Under intervjuerna framkom det dock även mer problematiska aspekter i samband med licenser. En av informanterna hade varit med om att det fanns krav på att data som tillhandahölls skulle tas bort om avtal med leverantören hävdades. Sådana krav gör det enligt informanten problematiskt att lagra data på slutanvändares enheter. Informanten nämnde dock att detta gällde tidiga licenser, så det är möjligt att denna form av krav inte är vanligt förekommande längre. Samma informant var även kritiskt inställd till licenser med krav på källhänvisning och menade på att sådana licenser utgör ett hinder för kombination av data.

“Jag förstår inte varför det är så många som har CC-BY-licenser också, där man måste visa var källan kommer ifrån. För det gör det ju också väldigt svårt att kombinera en stor mängd data, då man måste källhänvisa till varje minsta detalj överallt [...] Så ett typiskt sånt exempel är ju att, att mycket av den här datan som finns går ju inte lägga in i så här kartprogram om man vill återpublicera det som egna kartor. För att det ju, det skulle ju vara mer källhänvisningar än kartor i slutet av dagen då” (Informant A)

Samma informant poängterar även att källhänvisning borde vara problematiskt även för leverantörer. Informanten menar att leverantören antagligen inte vill synas i alla samband där deras data presenteras och att de istället begränsar i vilka syften data får vidareutnyttjas, men att detta är problematiskt då dessa syften kan vara öppna för tolkningar.

“[...] Istället försöker man då licensiera bort det där genom att man får bara använda den här datan i det här nobla och ädla syftet. Men det är ju en tolkningsfråga vad det här” (Informant A)

Ett annat ämne som togs upp var hur enkelt eller svårt det är att förstå sig på licenser. En informant upplevde att det kan vara svårt att sätta sig in i licenser, medan två andra betonade vikten av att det måste vara enkelt att förstå licenser. En av dem menade att det inte kan vara ett krav att hobbyutvecklare ska besitta juridisk kompetens och att de därför inte ska behöva sätta sig in i komplicerade avtal.

Slutligen uttryckte även en av informanterna att det är viktigt att licenser är kompatibla med varandra, även på internationell nivå. Den myndighet som publicerar data som informanten främst använder hade initialt använt en nationell licens, men hade nu gått över till att publicera data så att den tillhör allmän egendom (eng. public domain). Detta upplevde informanten som något positivt och menade på att den nationella licensen var problematisk ur kompatibilitetsperspektiv.

6 Diskussion

I föregående kapitel gavs en omfattande beskrivning av resultatet från den empiriska undersökningen. Detta kapitel består av en sammanfattande diskussion där de viktigaste resultaten diskuteras mot teori och relaterad forskning. Vidare reflekterar vi över hur studien gått och dess generaliserbarhet, för att slutligen ge förslag på vidare forskning.

6.1 Resultatanalys

6.1.1 Kvalitet och åtkomst för data och metadata

I kapitel 3 beskrivs det att denna kategori omfattar problematik runt åtkomst till och kvalitet i data och metadata. I resultatet såg vi att denna typ av hinder förekommer, men att olika aspekter har upplevts som mer eller mindre problematiska och att krav på kvalitet kan variera mellan olika projekt.

Flera informanter har varit nöjda med kvalitet på data inom de projekt som huvudsakligen diskuterats, men har samtidigt haft negativa erfarenheter i andra samband. Zuiderwijk och Janssen (2014) diskuterar att krav på datakvalitet är relativt och att det beror på i vilket syfte öppna data vidareutnyttjas. En variation i krav förekom även bland de intervjuade informanterna. Detta framgick bland annat då två informanter diskuterade att kvalitetskrav kan variera beroende på projektets typ och omfattning, men också genom att det förekom skilda åsikter kring hur acceptabelt fel i data är. En informant menade att fel i data spelar mindre roll så länge som konkurrerande tjänster har samma fel, medan en annan har valt att inte genomföra projekt till följd av fel i data.

Ytterligare ett hinder inom denna kategori som bland annat diskuteras av Janssen et al. (2012) är att det upplevs som svårt att hitta relevant data. Av informanterna som intervjuades upplevde de flesta att det är enkelt att hitta data, men att det kan vara svårare när likartade data inte finns centraliserad på en plats, exempelvis trafikdata inom länder med många olika trafikbolag. För informanterna har det varit mer problematiskt i de fall där data inte funnits publicerad som öppna data. Samtidigt har man i flera av dessa fall hittat vägar runt avsaknaden av data, exempelvis genom att via mjukvara läsa av information från myndigheters webbplatser. Det innebär att den information som de är ute också har funnits tillgänglig på en myndighets webbplats, men att myndighetens intention inte varit att publicera den som öppna data. I fall där information inte finns tillgänglig på en hemsida bör det inte finnas någon liknande väg runt problematiken. Ett hinder som bland annat diskuteras av Carrara et al. (2016) är att data fortfarande förekommer i icke-maskinläsbara format. Detta är något som också varit problematiskt bland informanterna och har i vissa fall hindrat projekt från att fortskrida.

Sammanfattningsvis verkar hinder inom denna kategori upplevas som problematiska vid utveckling av tjänster. Att data enbart finns i icke-maskinläsbara format har hindrat projekt från att fortskrida. I vissa fall där data som behövdes inte funnits publicerad som öppna data alls har informanterna hittat andra lösningar för att få tillgång. Vidare verkar krav på kvalitet variera från projekt till projekt.

6.1.2 Interaktion

Som förklarat under kapitel 3 avser denna kategori interaktion mellan leverantör och användare kan hämma tjänsteutveckling. I resultatet såg vi att interaktionen har fungerat både bra och dåligt, men också att informanterna kan uppleva att öppna data inte är helt förankrat inom leverantörers organisationer.

Ett exempel på en studie som tar upp hinder inom denna kategori är Berends et al. (2017), som menar att det bland annat kan saknas dialog mellan olika intressenter och att användarens behov inte prioriteras. Bland de informanter som intervjuats har det däremot funnits en dialog och överlag har de flesta varit nöjda med interaktionen. Trots att informanterna var nöjda togs dock även en del negativa erfarenheter upp. Exempel på sådana negativa upplevelser är att leverantören inte varit transparent med vilken data de planerar att lansera via sina API:er, att de inte informerat vid förändringar av API:er samt att leverantören nekat åtkomst till data trots tidigare överenskommelse.

Vidare upplevde flera informanter att det ibland verkar finnas två sidor av leverantören; en sida som är intresserad av och arbetar aktivt för öppna data, och en sida som inte inser värdet i det och som snarare ser det som något jobbigt. De menade på att det inte verkar vara förankrat i hela organisationen och att en del enbart ser öppna data som något som måste göras. Denna uppfattning verkar stämma väl överens med den diskussion som förs av Berends et al. (2017) om att öppna data-initiativ ofta drivs på nationell nivå uppifrån och ned, och att offentlig myndighet på lägre nivå inte har en kulturell tradition av att dela med sig data. En liknande osäkerhet kan ses hos de två informanter som inte riktigt litar på offentliga myndigheter i rollen som leverantör av data. En av dem menar på att detta kan vara en anledning till att få företag börjat anamma öppna data.

För att sammanfatta verkar det som att hinder inom denna kategori har upplevts som problematiska vid utveckling av tjänster. Bland de intervjuade informanterna verkar det dock som att det funnits en dialog, men att den inte alltid varit helt tillfredsställande då leverantören exempelvis inte varit transparent, eller att de gjort tomma överenskommelser. Vidare upplever vissa informanter att öppna data inte är helt förankrat i offentliga leverantörers organisationer.

6.1.3 Kunskap och färdigheter

Under kapitel 3 definierades denna kategori till att omfatta avsaknad av kunskap och färdigheter för att kunna vidareutnyttja öppna data. Inom de projekt som huvudsakligen diskuterats nämndes överlag inte någon explicit brist på kunskap, medan åsikterna om vilken kunskap som var viktig varierade.

Att avsaknad av kunskap för att använda och förstå data kan utgöra ett hinder har bland annat identifierats av Janssen et al. (2012). I kontrast till detta var det enbart en informant som uttryckte en explicit brist på kunskap i relation till de projekt som huvudsakligen diskuteras. Detta behöver dock inte innebära att övriga informanter inte upplevt en brist på kunskap. Det skulle kunna vara så att brist på kunskap och färdigheter utgör ett större hinder inom användargrupper där syftet inte är att utveckla en tjänst. Exempelvis de användartyper som Statskontoret (2018) beskriver som intresserad allmänhet och användare av dataset eller hos tjänsteutvecklare som ännu inte anammat öppna data. I studien av Janssen et al. (2012) fokuserar de inte på någon specifik användargrupp och det är därför svårt att konstatera om brist på kunskap och färdigheter även upplevs av personer inom tjänsteutveckling.

När det kommer till vilka färdigheter som behövs för att vidareutnyttja data, nämner Carrara et al. (2015b) bland annat att färdigheter inom teknologi och statistik är en viktig förutsättning. Flera informanter var eniga om att denna typ av kunskap behövs, men två informanter poängterade att detta inte var allt. En av dem nämnde att det behöver finnas någon form av kreativitet eller en idé, medan den andra menade att det viktigaste är att det finns förståelse för vilka behov som finns.

Sammanfattningsvis har enbart en informant nämnt att brist på kunskap och färdigheter utgjort ett hinder. Trots detta kan vi ej dra slutsatsen att brist på kunskap och färdigheter inte utgjort ett hinder för de övriga. Vidare var flera informanter överens om att akademisk kunskap, såsom inom teknologi och statistik behövs, även om den inte alltid betraktas som det mest viktiga.

6.1.4 Teknologi och verktyg

Under kapitel 3 definierades denna kategori till att avse hur tekniska aspekter kan utgöra hinder, både av generell karaktär och i samband med leverantörens tekniska infrastruktur. Vidare omfattas även hur vidareutnyttjande av öppna data kan hämmas som en följd av brist på verktyg. I resultatet nämndes inte någon brist på verktyg, samtidigt som erfarenheter och åsikter om leverantörens tekniska infrastruktur varierade.

De informanter som ej använt API:er upplevde inte att leverantörens tekniska infrastruktur påverkade märkbart. Detta är inte förvånande då de enbart interagerar med den genom att ladda ner en fil. Av de som använt API:er fanns en annan uppfattning och de menade på att tillgänglighet på leverantörens infrastruktur spelar stor

roll. En av dem ställde höga krav på tillgänglighet, medan den andra påpekade att hög tillgänglighet är en förutsättning för att storskaliga tjänster ska kunna utvecklas baserat på datan. Samtidigt poängterade en av dem att det viktigaste är att det finns en tydlighet runt vilken tillgänglighet som erbjuds.

Brist på verktyg som stöd för vidareutnyttjande av öppna data har i tidigare forskning tagits upp som ett hinder, bland annat av Janssen et al. (2012). I kontrast till detta nämnde inte någon informant att de upplevt en brist på verktyg. Det är möjligt att denna typ av verktyg främst behövs av andra typer av användare som inte har motivet att utveckla tjänster baserat på data. Ett uttryckt behov av verktyg för andra användargrupper kan ses i rapporten av Statskontoret (2018). De menar att en viktig förutsättning för att intresserad allmänhet ska kunna vidareutnyttja data är att de ska kunna ta del av data direkt på hemsidan, till exempel genom visualiserings- och filtreringsverktyg.

Som tidigare nämnt väljer vi att betrakta API:er som ett verktyg för att kunna välja ut specifika delmängder av data. Bland informanter rådde det skilda åsikter om huruvida API:er är ett värdefullt verktyg eller ett hinder för innovation. Informanten som är negativt inställd till API:er menar att data ibland kan döljas bakom fördefinierade frågor och att potentiella användningsområden bestäms redan när API:er designas. Detta perspektiv på API:er går att relatera till ett hinder som identifierats av Zuiderwijk et al. (2012). Detta hinder rör inte specifikt API:er, men berör att åtkomst till bearbetade data utan möjlighet att få tag på ursprungsdata kan vara problematiskt. Denna kritik mot vissa former av API:er är också intressant med tanke på att Statskontoret (2018) betraktar brist på API:er som det främsta hindret för programmerare och tjänsteutvecklare. Om dessa kritiska synpunkter är mer utbredda bör det också vara viktigt att ta särskild hänsyn till att inte innovation hämmas.

Ytterligare ett argument till att använda rådata istället för API:er som förekom under intervjuerna var att undvika beroende av myndighetens tekniska infrastruktur. Detta stämmer bra överens med ett av de fynd som Magalhães och Roseira (2016) gjorde, där de fann att företag verkar föredra data på fil framför API:er. Anledningen till detta var det fanns ett behov av ökad åtkomst och kontroll.

Sammanfattningsvis nämnde inte informanterna att brist på verktyg har varit problematiskt. Till vilken grad myndighetens tekniska infrastruktur påverkar verkar variera beroende på om data hämtas via API eller fil. Slutligen verkar det finnas skilda åsikter om huruvida ett API med fördefinierade frågor fungerar som ett bra verktyg eller som ett hinder för innovation.

6.1.5 Licensiering

I kapitel 3 definierades denna kategori till att avse hinder kring avsaknad av licens, samt fall där licenser är för restriktiva. I resultatet framgick att de flesta av informanterna ej upplevt licenser som problematiskt men det förekom även mer kritiska åsikter.

Ett exempel på tidigare litteratur där avsaknad av licens eller för restriktiva licenser nämns som hinder är Janssen et al. (2012). Som tidigare sagt upplevde de flesta av informanterna att licenser ej påverkat utvecklingen av tjänster. I värsta fall har de fått inkludera en källhänvisning, vilket ej upplevdes som problematiskt. En informant var dock mer negativt inställd till krav på källhänvisning och menar på att detta är väldigt problematiskt när en datamängd ska kombineras med andra datamängder.

Samma informant påpekade även att licenser ibland kan ha restriktioner för i vilka syften data får vidareutnyttjas, samtidigt som dessa syften är öppna för egna tolkningar. Om licenser är fria för tolkningar bör de även vara svåra att förstå, vilket också lyfts fram som en viktig aspekt. Exempelvis påpekade en informant att licenser kan vara svåra att förstå, medan två andra menade att det är viktigt att de är lätta att förstå. Denna aspekt tas upp av Berends et al. (2017) som menar på att potentiella användare kan undvika att vidareutnyttja data på grund av att licensen är svår att förstå.

Slutligen tyckte en av informanterna också det var viktigt att licenser är kompatibla med varandra, även på internationell nivå. Detta stämmer överens med hinder som tas upp av Berends et al. (2017), där de bland annat menar att nationella licenser kan utgöra hinder för internationell kompatibilitet.

Sammanfattningsvis har licenser överlag inte utgjort något hinder för de intervjuade informanterna. Däremot framkom åsikter om att det är viktigt att licenser är enkla att förstå och att de är kompatibla med varandra. Slutligen fanns det skilda åsikter om huruvida krav på källhänvisning är en problematisk aspekt eller ej.

6.1.6 Sammanfattade upptäckter

För att få en övergripande bild av de upptäckter som gjorts inom ramen för denna studie har de resultat som vi anser är mest viktiga sammanställts i tabell 6.1. De resultat som ansetts som mest viktiga är de resultat där flera informanter haft liknande åsikter eller erfarenheter, där informanter haft skilda åsikter samt aspekter som är intressanta i relation till tidigare litteratur.

Kvalitet och åtkomst för data och metadata
Krav på datakvalitet verkar variera mellan olika projekt.
För det mesta enkelt att hitta data. Kan vara svårare när data ej återfinns på centraliserad plats.
Brist på data i maskinläsbart format har hindrat projekt från att fortskrida.
Flera har erfårit att data de behöver ej finns publicerad. Projekt har i flera fall fortskridit trots avsaknad.
Interaktion
Dialog mellan leverantör och informant har i samtliga fall existerat. Interaktion har fungerat både bra och dåligt.
Flera upplever att det kan finnas två sidor av leverantörers organisationer, där den ena sidan uppfattas som mindre drivande.
Flera upplever en osäkerhet med offentlig myndighet i rollen som leverantör av data.
Kunskap och färdigheter
Kunskap inom exempelvis teknik och statistik upplevs överlag som nödvändigt.
Teknologi och verktyg
Teknisk infrastruktur hos leverantör påverkar de informanter som använder API i större utsträckning än de som hämtar data på fil.
Informanter som använt API:er anser att hög tillgänglighet är viktigt för storskaliga projekt.
Det förekom skilda åsikter om huruvida API:er som verktyg för att välja ut delmängder av data gynnar eller hämmar innovation.
Licensiering
Flera informanter upplevde att licensiering inte utgjort något hinder.
Det förekom skilda åsikter om huruvida krav på källhänvisning är problematiskt.

Tabell 6.1. Sammanfattade upptäckter

6.2 Reflektion över undersökningen

Denna undersökning har gjorts i syfte att utforska hur hinder upplevs vid tjänsteutveckling baserat på öppna data. Empiriskt material bygger på intervjuer som genomförts med informanter i två olika länder som använt sig av olika typer av data. Trots att länderna möjligtvis kan vara jämförbara är det antagligen svårt att generalisera resultaten till tjänsteutvecklare i allmänhet. Eftersom informanterna använt sig av olika typer av data och kommer åt data på olika sätt finns det skillnader inom gruppen. Ytterligare faktorer som påverkar generaliserbarheten är att empiriskt material bygger på förhållandevis få intervjuer, då fler inte varit möjliga att genomföra på grund av tidsbrist. Vidare bygger det empiriska materialet enbart på erfarenheter och åsikter från personer som redan vidareutnyttjar öppna data i sina tjänster. Det skulle också varit relevant att inkludera tjänsteutvecklare som valt att inte integrera öppna data i sina tjänster. Trots att resultaten antagligen inte går att generalisera tror vi att undersökningen har givit insikter om aspekter värda att studera vidare.

6.3 Förslag till vidare forskning

Som det nämndes i kapitel 6.2 finns det skilda erfarenheter bland de informanter som ingått i denna undersökning. Vidare forskning skulle kunna göras runt vilka hinder som upplevs i mer preciserade kontexter, exempelvis en specifik typ av data eller vid användning av en specifik dataportal. Det skulle också vara intressant att utforska vilka hinder som andra användargrupper upplever, exempelvis de som Statskontoret (2018) beskriver som intresserad allmänhet. Under undersökningen framkom det tendenser på att vissa hinder upplevs som mer problematiska än andra. Framtida studier skulle därför kunna ägnas till att försöka rangordna de hinder som upplevs. Mer specifika aspekter som skulle kunna utforskas är vad som egentligen karakteriserar “för restriktiva” licenser, samt vilka användargrupper som har behov av verktyg som stöd för att vidareutnyttja öppna data.

7 Slutsats

Inom denna studie har vi utforskat hur tjänsteutvecklare upplever hinder vid vidareutnyttjande av öppna data. Från tidigare litteratur valdes en taxonomi över hinder ut, bestående av fem olika kategorier: kvalitet och åtkomst för data och metadata, interaktion, kunskap och färdigheter, teknologi och verktyg samt licensiering. Baserat på fem intervjuer har vi fått insikt i hur hinder inom respektive kategori kan upplevas inom tjänsteutveckling. De viktigaste upptäckterna har vidare sammanställts i en tabell strukturerad utefter de fem kategorierna. Upptäckterna visar på likheter och skillnader i jämförelse med hinder från tidigare litteratur, men även på aspekter där det verkar finnas skilda åsikter. Exempelvis huruvida krav på källhänvisning i licenser upplevs som problematiskt.

Resultatet från denna undersökning går antagligen inte att generalisera till tjänsteutvecklare i allmänhet, då det förekommit olikheter inom den undersökta gruppen samt att antalet informanter var relativt få. Trots detta tror vi att de viktigaste upptäckterna är intressanta och att de kan utgöra en grund för vidare forskning.

8 Referenser

- Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity: The International Journal of Government & Democracy in the Information Age*, 19(1/2), 29-152.
- Bell, J., & Waters, S. (2015). *Introduktion till forskningsmetodik*. Lund: Studentlitteratur AB
- Berends, J., Carrara, W., & Vollers, H. (2017). *Analytical Report 5: Barriers in working with Open Data*.
- Carrara, W., Chan, W. S., Fischer, S., & van Steenbergen, E. (2015a). *Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources*. Luxemburg: Publications Office of the European Union
- Carrara, W., Fischer, S., & van Steenbergen, E. (2015b). *Analytical Report 2: E-skills and Open Data*.
- Carrara, W., Fischer, S., & van Steenbergen, E. (2015c). *Open Data Maturity in Europe 2015*.
- Carrara, W., Fischer, S., Oudkerk, F., van Steenbergen, E., & Tinholt, D. (2015d). *Analytical Report 1: Digital Transformation and Open Data*.
- Carrara, W., Nieuwenhuis, M., & Vollers, H. (2016). *Open Data Maturity in Europe 2016*.
- Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, 31(1), 10-17.
- Davies, T. (2010). Open data, democracy and public sector reform: A look at open government data use from data.gov.uk (Masteruppsats). Oxford: University of Oxford. Tillgänglig: <http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf>
- Huijboom, N., & Van den Broek, T. (2011). Open data: An International comparison of strategies. *European Journal of ePractice*, 12(1), 4-15.
- Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, 28(4), 446-456. DOI: 10.1016/j.giq.2011.01.004
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29(4), 258-268. DOI: 10.1080/10580530.2012.716740

- Kučera, J. (2017). *Analysis of Barriers to Publishing and Re-Use of Open Government Data*. 2017 IDIMT: September 6-8. Poděbrady, Tjeckien, Linz: Johannes-Kepler-Universität Linz, 305-314
- Kuk, G., & Davies, T. (2011). *The Roles of Agency and Artifacts in Assembling Open Data Complementarities*. 2011 ICIS: December 4-7. Shanghai, Kina
- Magalhães, G., & Roseira, C. *Exploring the Barriers in the Commercial Use of Open Government Data*. 2016 9th International Conference on Theory and Practice of Electronic Governance (ICEGOV): Mars 1-3. Montevideo, Uruguay, 211-214
- Martin, C. (2014). Barriers to the Open Government Data Agenda: Taking a Multi-Level Perspective. *Policy & Internet*, 6(3).
- Martin, S., Foulonneau, M., Turki, S., & Ihadjadene, M. (2013). Risk analysis to overcome barriers to open data. *Electronic Journal of e-Government*, 11(1), 348-359.
- Open Knowledge International. (2016). Why Open Data?. Hämtad 2018-03-28, från <http://opendatahandbook.org/guide/en/why-open-data/>
- Open Knowledge International. (2017a). Open Definition 2.1. Hämtad 2018-04-30, från <https://opendefinition.org/od/2.1/en/>
- Open Knowledge International. (2017b). Open Government Data and Content. Hämtad 2018-05-11, från <https://opendefinition.org/government/>
- Patel, R., & Davidson B. (2011). *Forskningsmetodikens grunder*. Lund: Studentlitteratur AB
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, 22(1), 1-24.
- Statskontoret. (2018). *utmaningar för att använda myndigheternas öppna data*. Stockholm: Statskontoret
- Susha, I., Grönlund, Å., & Janssen, M. (2015). Organizational measures to stimulate user engagement with open data. *Transforming Government: People, Process and Policy*, 9(2), 181-206. <https://doi-org.ezproxy.ub.gu.se/10.1108/TG-05-2014-0016>
- Ubaldi, B. (2013). *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives* (OECD Working Papers on Public Governance, 22). Paris: OECD Publishing

Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. *Information Economics, Paris*.

World Wide Web Foundation. (2015). *Open Data Barometer Global Report: Third Edition*. Hämtad 2018-05-15, från <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf>

World Wide Web Foundation. (2017). *Open Data Barometer Global Report: Fourth Edition*. Hämtad 2018-05-17, från <https://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport.pdf>

Yang, Z., & Kankanhalli, A. (2013). *Innovation in Government Services: The Case of Open Data*. I Y.K. Dwivedi, H.Z. Henriksen, D. Wastell, & R. De' (Ed.), *Grand Successes and Failures in IT. 2013 IFIP WG 8.6 International Working Conference on Transfer and Diffusion of IT, TDIT: Juni 27-29. Bangalore, Indien, Berlin, Heidelberg: Springer, 644-651*

Zuiderwijk, A., & Janssen, M. (2014). Barriers and Development Directions for the Publication and Usage of Open Data: A Socio-Technical View. I M. Gascó-Hernández (Ed.), *Open Government: Opportunities and Challenges for Public Governance* (s. 115-135). New York: Springer

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of e-Government, 10(2)*, 156-172.

9 Bilagor

Bilaga 1: Intervjumanus – Svenska

Intervjuns struktur

- Presentation
- Förklara studiens syfte
- Förklara intervjudeltagarens rättigheter
- Förklara hantering av uppgifter
- Fråga om Inspelningsmedgivande
- Ställ intervjufrågor
- Avslut

Presentation

Hej _____.

Jag heter _____ och med oss har vi även _____ som främst kommer att ta anteckningar.

Innan vi sätter igång skulle jag bara vilja gå igenom ett par saker.

Förklara studiens syfte

Syftet med vår studie är att utforska vilka hinder som upplevs vid användning av öppna data från myndigheter för att utveckla produkter och tjänster. Under detta samtal vill vi därför höra mer om dina erfarenheter av detta.

Förklara intervjudeltagarens rättigheter

Jag vill även poängtera att ditt deltagande är frivilligt och du bestämmer själv om du vill svara på alla frågor. Självklart kan du när som helst välja att avbryta intervjun om du önskar.

Förklara hantering av uppgifter

Vi kommer inte skriva ut några namn i uppsatsen, varken på personer, företag eller myndigheter. Jag och _____ är de enda som kommer att ha åtkomst till uppgifter om dig, samt det material vi samlar in under samtalet. I uppsatsen kommer du att benämnas som "Informant A", eller liknande.

Fråga om inspelningsmedgivande

För att vid senare tillfälle kunna gå tillbaka och repetera vad som sägs här idag skulle vi vilja spela in samtalet. Det är enbart jag och _____ som kommer ha tillgång till ljudinspelningen.

- Godkänner du att vi spelar in samtalet?

Innan vi sätter igång ljudinspelningen...

- Har du några frågor eller kommentarer?

Nu sätter vi igång ljudinspelningen.

Ställ intervjufrågor

Inledande frågor

- Skulle du kunna berätta lite om din erfarenhet av att använda öppna data?
- Skulle du kunna berätta lite om den/de tjänster du varit med och utvecklat?
 - o Vilken typ av data?
 - o Hur lång tid har ni arbetat med att utveckla tjänsten/tjänsterna?
 - Om många nämn de mest omfattande tjänsterna sett till tid.
 - o Hur många användare?
 - Om många nämn de mest omfattande tjänsterna
 - o Hur stort team?
 - o Utvecklad privat, på ett företag, hackathon?

Kvalitet och åtkomst för data och metadata

- Hur har kvalitet på data och tillhörande metadata påverkat utvecklingen av tjänsten/tjänsterna?
- Hur upplever du att det fungerat att hitta den data ni behövde för utveckling av tjänsten/tjänsterna?

Licens

- Kan du berätta lite om vilken påverkan den licens som datan ni nyttjar (om någon) publicerats under har påverkat utvecklingen av tjänsten/tjänsterna?

Interaktion

- Vad har motiverat dig/er till att utveckla tjänsten/tjänsterna?
- Skulle du kunna beskriva relationen mellan dig/er och den myndighet vars data ni använder för tjänsten/tjänsterna?
 - o Hur skulle du/ni önska att relationen såg ut?

Kunskap och färdigheter

- Vilka färdigheter tror du att du/ni besitter som varit viktiga vid utvecklingen av tjänsten/tjänsterna? Exempelvis tekniska, statistiska/analytiska, kommunikation, företagande, eller det området datat handlar om

- Tror du att några övriga färdigheter skulle förenklat utvecklingen ännu mer? I så fall, vilka?

Teknologi och verktyg

- Hur tycker du att myndighetens tekniska infrastruktur (t.ex. långa svarstider, att API'er är otillgängliga, m.m.) påverkat utvecklingen av tjänsten/tjänsterna?
- Har du/ni, eller skulle du/ni haft användning av några verktyg, exempelvis för visualisering eller bearbetning av data? I så fall, vad för några?

Avslutande frågor

- Med bakgrund mot vad vi diskuterat. Tycker du att vi har missat något, eller finns det något annat du skulle vilja tillägga?

Avslut

Tack så hemskt mycket för ditt deltagande.

Om du kommer på något ytterligare som vi missat att diskutera här idag så får du hemskt gärna komplettera via e-mail.

Bilaga 2: Intervjumanus – Engelska

Interview structure

- Presentation
- Explain the purpose of the study
- Explain interviewees rights
- Explain management of information
- Question about recording permission
- Ask interview questions
- Ending

Presentation

Hi _____.

My name is _____ and _____ is also here who mainly will be taking notes. Before we start, I'd like to go through a few things.

Explain the purpose of the study

The purpose of our study is to explore what barriers that are experienced while using open government data for development of products and services. During this conversation we'd like to know more about your experiences of this.

Explain interviewees rights

I would also like to make it clear that your participation is voluntarily and you can decide yourself if you want to answer all the questions. If you at any point would like to cancel the interview, you are of course free to do so.

Explain management of information

In the thesis we won't print any names, neither on people, business or governmental institutions. Me and _____ are the only ones who will have access to the information about you and the material that we gather here today. In the thesis you will be called "Informant A" or similar.

Question about recording permission

In order to be able to go back and repeat what has been said here today, we'd like to record the conversation. Only me and _____ will have access to the audio recording.

- Do you approve that we record the call?

And before we start recording...

- Do you have any questions or comments?

Now we'll start recording.

Ask interview questions

Initial questions

- Could you please tell us a little about your experience of using open government data?
- Could you tell us about the service(s) that you've developed?
 - o What type of data?
 - o For how long have you worked with the development of the service(s)?
 - If several, please mention the most comprehensive ones
 - o How many users?
 - If several, please mention the most comprehensive ones
 - o How large is your team?
 - o Developed privately, in a business, hackathon?

Accessibility and quality of data and metadata

- How has data quality and the quality of metadata affected development of the service(s)?
- How do you think that it has worked to find the data you needed for development of the service(s)?

Licensing

- About the license under which the data you use has been published. How has it affected development of the service(s)?

Interaction

- What has been your motivation in developing these service(s)?
- Could you describe the relation between you and the government who publish the data you use for the service(s)?
 - o How would you like that relationship to be? Optimally?

Knowledge and skills

- What skills do you think was important in development of the service/services? For example skills regarding technology, statistics/analytics, communication, business-related or skills regarding the area the open data is related to.
 - o Is there any other skill that would have simplified the development even further? If so, which ones?

Technology and tools

- How do you think that the technical infrastructure (like long response times, unavailability of API's) of the governmental institution has affected development of the service/services?
- Did you come across any useful tools, such as tools for visualising or processing the open data? If any, which ones? Is there any tool you did not come across that you think would simplify the work with the open data?

Concluding questions

- From what we have discussed today, do you think that we have forgotten to ask something, or is there anything else you would like to add?

Ending

Thank you very much for your participation.

If you later recall something in addition to what we have discussed today, then you are more than welcome to provide us with these details by email.