



LUND UNIVERSITY

Rekommendationer för samverkan och delning av data, teknik och kunskap inom myndighetsdrivna öppna data-ekosystem

Linåker, Johan; Runeson, Per

2022

Document Version:
Förlagets slutgiltiga version

[Link to publication](#)

Citation for published version (APA):

Linåker, J., & Runeson, P. (2022). *Rekommendationer för samverkan och delning av data, teknik och kunskap inom myndighetsdrivna öppna data-ekosystem*. (Teknisk rapport; Nr 109). Department of Computer Science, Lund University.

Total number of authors:
2

Creative Commons License:
CC BY

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00



LTH
LUNDS TEKNISKA
HÖGSKOLA

Rekommendationer för samverkan
och delning av data, teknik och
kunskap inom myndighetsdrivna
öppna data-ekosystem

ISSN 1404-1200

Teknisk rapport: 109, 2022

Datum 2022-01-25

Avdelningen för Datavetenskap
Johan Linåker, Postdoktor
Per Runeson, Professor

Innehållsförteckning

Sammanfattning	3
Rekommendationer	4
Inledning	8
Bakgrund	11
<i>Samverkan kring öppna data inom ekosystem</i>	<i>11</i>
<i>Styrningsstrukturer inom öppna data-ekosystem</i>	<i>12</i>
<i>Fördelar med samverkan kring öppna data</i>	<i>14</i>
<i>Utmaningar med samverkan kring öppna data</i>	<i>15</i>
Affärs- och konkurrensaspekter	15
Tekniska aspekter	16
Kulturella, organisatoriska och juridiska aspekter	16
Metod och forskningsdesign	17
Undersökta fall av myndighetsdrivna öppna data-ekosystem	19
<i>JobTech Dev</i>	<i>19</i>
Om ekosystemet	19
Om plattformen	19
<i>Trafiklab</i>	<i>20</i>
Om ekosystemet	20
Om plattformen	21
<i>HSL Developer Community</i>	<i>22</i>
Om ekosystemet	22
Om plattformen	22

<i>City of Chicago Open Data program</i>	23
Om ekosystemet	23
Om plattformen	24
<i>Road Data Lab (RoDL)</i>	25
Om ekosystemet	25
Om plattformen	25
<i>ESS-CSDL</i>	26
Om ekosystemet	26
Om plattformen	27
Rekommendationer för att skapa och facilitera öppna data-ekosystem	28
<i>Styrningsstrukturer för ekosystem och ägande av plattform</i>	28
<i>Medling och neutralitet som grund för ett sammanhållet ekosystem</i>	30
<i>Ekosystemets syfte och incitament för samverkan och datadelning</i>	31
<i>Ekosystemets geografiska och demografiska omfattning</i>	33
<i>Funktioner för delning och samverkan inom ekosystemet</i>	34
<i>Verktyg och processer för öppen samverkan</i>	35
<i>Öppen programvara som komplement och katalysator</i>	37
<i>Insamling och delning av data inom ekosystemet</i>	38
<i>Balansering av värde och risk genom form på delade data</i>	40
<i>Standarder och format för datadelning inom och utom ekosystemet</i>	41
<i>Mätning och uppföljning av ekosystemets hälsa och utveckling</i>	42
Referenser	45

Sammanfattning

Data utgör idag en kritisk resurs som genom bearbetning och öppet tillgängliggörande kan bidra till och ligga till grund för effektiviseringar, nya innovationer och tillika ett öppnare och mer demokratiskt samhälle. Mycket av detta värde skapas dock först när datan bearbetas, berikas och kombineras med annan data från andra källor utanför sin organisations.

Genom öppen samverkan och delning av data inom s.k. Öppna data-ekosystem kan dataproducenter och konsumenterna bidra till detta värdeskapande, dels genom kollektiva bidrag och berikning av den gemensamma datan, dels genom delning av fler datamängder. Ytterligare värde kan därtill skapas genom att samverkan utökas till att även innefatta formen för hur datan samlas in och delas, samt tekniken som möjliggör detta arbete.

I forskningsprojektet JobTech Research har forskare från Lunds universitet studerat hur öppna data-ekosystem kan förverkligas och faciliteras från ett myndighetsperspektiv. Syftet är att skapa ytterligare värde och ökad användning av öppen myndighetsdata. Delning och samverkan inom ekosystemet möjliggörs ofta genom någon form av teknisk plattform där data tillgängliggörs, ex. via API:er. Plattformen utvecklas och underhålls vanligen av den s.k. plattformsledaren, en aktör som också vanligtvis är den som initierat det öppna data-ekosystemet och leder dess gemensamma arbete.

Genom forskningsprojektet har flertalet myndighetsdrivna öppna data-ekosystem studerats där rollen som plattformsledare helt eller delvis fylls av en eller flera offentliga aktörer. Forskningen har tagit utgångspunkt från JobTech Dev, ett ekosystem som har skapats och faciliteras av Arbetsförmedlingen i syfte att främja utveckling av nya och förbättrade lösningar för digital matchning och vägledning för den svenska arbetsmarknaden.

Denna rapport summerar resultat och lärdomar från forskningsprojektet i form av ett antal konkreta rekommendationer för hur myndigheter kan gå till väga för att utforma sina egna öppna data-ekosystem och hur de kan initiera och facilitera en värdeskapande samverkan inom dessa.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

Styrningsstrukturer för ekosystem och ägande av plattform

- **R1:** Överväga hur ägarskapet till av den organisation som utgör plattformsledare bäst kan anpassas för att främja samverkan och engagemang med och mellan nyckelaktörer inom ett ekosystem.
 - Om ägarskapet innehas av en enskild offentlig organisation behövs en aktiv dialog skapas med nyckelaktörer samt att dessa får möjlighet att utöva inflytande på plattformens innehåll och utveckling genom en öppen samverkan.
 - Om ägarskapet delas mellan ett antal offentliga och/eller privata aktörer behöver samverkan möjliggöras genom ägarskapet, men även mot de aktörer som ej är delägare men som fortfarande är med i ekosystemet.
- **R2:** Identifiera nyckelaktörer inom ekosystemet och möjliggöra för dessa att ha ett informellt eller formellt inflytande över ekosystemets inriktning, plattformens utveckling och den data som delas. Vilka aktörer som utgör nyckelaktörer kan variera men kan inkludera plattformsledarens ägare, dess partners, organisationer med t.ex. stora mängder användare eller andra värdefulla tillgångar (som data och programvara), men även medborgare och slutanvändare.

Medling och neutralitet som grund för ett sammanhållet ekosystem

- **R3:** Se till att det finns en neutral plattformsledare inom ekosystemet som kan bygga förtroende och tillit hos ekosystemets övriga aktörer, samt möjliggöra samverkan

mellan dessa trots ev. utmaningar, som konkurrensförhållanden mellan vissa (nyckel)aktörer.

- **R4:** Ta hänsyn till, diskutera och leda arbetet med att hantera ev. bekymmer och risker som aktörer ser med delning och användning av data, teknik och annan kunskap inom ekosystemet.

Ekosystemets syfte och incitament för samverkan och datadelning

- **R5:** Upprätthålla en tydlig dialog med aktörer inom ekosystemet angående det övergripande syftet och främja ett värdeskapande hos de respektive aktörerna, både sett till att dela data och delta i själva samverkan och kunskapsutbytet däromkring.
- **R6:** Skapa gemensamma incitament trots olika affärsmodeller som tillämpas av ekosystemets aktörer, oavsett om ekosystemet primärt kan ses som affärs- eller samhällsdrivet, och hur datan eller samverkan utgör en komponent där i.

Ekosystemets geografiska och demografiska omfattning

- **R7:** Upprätthålla en tydlig dialog med ekosystemet angående dess omfattning och framtida utveckling. I detta arbete bör plattformsledaren beakta och ev. inkludera aktörer vars användningsfall kan sträcka sig utanför ekosystemets avsedda geografiska och demografiska täckning.

Funktioner för delning och samverkan inom ekosystemet

- **R8:** Överväga behoven hos ekosystemets aktörer och utifrån detta ta sig an funktioner för att bemöta dessa för att främja delning och samverkan inom ekosystem. Funktioner kan inkludera insamling, bearbetning och publicering av data, samt tillhandahållande av verktyg och infrastruktur för att kunna använda och bearbeta datan, och tjänster och applikationer som kan möjliggöra olika användningsfall baserade på datan.

Verktyg och processer för öppen samverkan

- **R9:** Möjliggöra en öppen samverkan inom ekosystemet genom att tillhandahålla öppna kommunikationskanaler, facilitera en öppen kravhanteringsprocess och aktivt engagera aktörerna inom ekosystemet i samarbetet. Öppna och vedertagna verktyg bör väljas i största möjliga mån.

Öppen programvara som komplement och katalysator

- **R10:** Utveckla och samarbeta med ekosystemets aktörer om öppen programvara som kan förbättra användning och innovation baserat på data. Användningsfall kan innefatta insamling och berikning av data, ge inspiration och dokumentation, eller allmänt stöd till utvecklare genom verktyg, bibliotek och infrastruktur. Plattformsledaren kan också överväga att dela sina applikationer som använder data som input, vilket kan inkludera algoritmer, såväl som verktyg och konsumentinriktade produkter.

Insamling och delning av data inom ekosystemet

- **R11:** Underlätta och möjliggöra delning av data, utöver den egna. Detta kan göras antingen genom att hjälpa aktörerna att aktivt eller passivt bidra med sin data. I det aktiva fallet använder aktörerna en teknisk infrastruktur tillhandahållen av plattformsledaren. I det passiva fallet samlar plattformsledaren in data från aktören för att sedan publicera den på plattformen. En ytterligare aspekt är att också överväga hur de kan använda externa ekosystem för att samarbeta kring data.

Balansering av värde och risk genom form på delad data

- **R12:** Överväga och anpassa datans form för att möjliggöra delning med hänsyn till aktuella integritets-, sekretess- och affärsrisker samtidigt som potential för värdeskapande bibehålls. Granularitet, grad av bearbetning och aktualitet kan från detta perspektiv användas som redskap av en plattformsledare och ekosystemets aktörer för att hitta en lämplig form på datan som alla eller en majoritet av ekosystemets aktörer är bekväma med.

Standarder och format för datadelning inom och utom ekosystemet

- **R13:** Anta och främja öppna standarder där de finns för att möjliggöra interoperabilitet mellan tjänster och portabilitet av data inom ekosystemet och utåt. Vid behov bör plattformsledaren också ta en aktiv roll för att underlätta utvecklingen av nya standarder för att uppnå samma mål.

Mätning och uppföljning av ekosystemets hälsa och utveckling

- **R14:** Utarbeta en plan för kontinuerlig mätning och uppföljning av ekosystemets hälsa som används och visas upp internt och externt för att hjälpa styrning av ekosystemet men också attrahera nya aktörer.

Inledning

Data har av många beskrivits som den nya oljan¹. Den utgör ett råmaterial som genom bearbetning kan bidra till och ligga till grund för effektiviseringar, nya innovationer och tillika ett öppnare och mer demokratiskt samhälle (Attard et al., 2015). Som ett led i detta tillgängliggör offentlig sektor mer och mer data som öppna data, dvs. data som fritt får användas utan inskränkningar². Detta arbete drivs på både från övriga sektorer och medborgarna, men även genom direktiv från EU där det senaste kommer att implementeras genom den nya “öppna data-lagen”³ som förväntas träda i kraft inte långt efter publiceringen av denna rapport.

En aspekt som ofta faller bort i diskussioner och praktiskt arbete är hur värdet av datan uppkommer. Liknelsen mellan data och olja är rätt passande då det är först när datan bearbetas, berikas och kombineras med annan data som själva värdet skapas. Processen från att datan samlas in och publiceras till att den kommer i användning kan på detta sätt liknas vid en värdekedja som kan bestå av flertalet olika aktörer som dataproducenter, datadistributörer, tjänsteleverantörer, apputvecklare, och slutanvändare (Lindman et al., 2014). Roller som kan innehas av aktörer både från offentlig och privat sektor, tillika civilsamhälle, akademi och medborgare själva.

Med denna typ av värdekedja kan ytterligare värde skapas kollektivt genom att skapa återkopplingsmöjligheter och gränssytor mellan dessa aktörer för att de ska kunna bidra till underhållet och berikandet av den gemensamma datan. Organisk utveckling av en sådan samverkan kan bidra till att ytterligare datamängder delas. Samverkan kan därtill komma att utökas till att även innefatta formaten och strukturen för de sätt som datan samlas in och delas, samt tekniken som möjliggör detta arbete och kanske till och med

¹ <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

² <https://opendefinition.org/od/2.1/en/>

³ <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2020/09/sou-202055/>

ligger till grund för de tjänster och applikationer som använder sig av datan. Detta kan i sin tur möjliggöra både innovation och interoperabilitet bland aktuella tjänster och applikationer, samt portabilitet av datan vilket möjliggör en långsiktig och hållbar förvaltning av denna resurs.

I forskningsprojektet JobTech Research⁴ har forskare från Lunds universitet studerat hur dessa typer av värdekedjor och former för samverkan kan förverkligas och faciliteras från ett myndighetsperspektiv i syfte att skapa ytterligare värde och användning av öppen myndighetsdata. I stället för värdekedjor har forskarna valt att se på samspelet som ekosystem där olika aktörer samlas under en gemensam vision för vad de vill åstadkomma med datan. Aktörerna inom dessa s.k. öppna data-ekosystem delar och samverkar kring datan, men även formatet och tekniken för hur den samlas in, berikas, delas och tillämpas. Denna delning och samverkan möjliggörs ofta genom någon form av teknisk plattform som samlar och tillgängliggör data, ex. via API:er. Plattformen utvecklas och underhålls vanligen av den s.k. plattformsledaren, en aktör som också vanligtvis är den som initierat det öppna data-ekosystemet och leder dess gemensamma arbete.

Genom forskningsprojektet har olika myndighetsdrivna öppna data-ekosystem studerats med utgångspunkt från ett av dessa, JobTech Dev, som studerats på särskild detaljnivå. JobTech Dev är ett öppna data-ekosystem som har skapats och faciliteras av Arbetsförmedlingen i syfte att främja utveckling av nya och förbättrade lösningar för digital matchning och vägledning för den svenska arbetsmarknaden. Genom ekosystemet har de därför samlat många av arbetsmarknadens olika aktörer, både privata och offentliga, för att dela och samverka kring data, samt format och teknik, som kan bidra till detta syfte.

Utöver Arbetsförmedlingens JobTech Dev, har forskarna bland andra även studerat:

⁴ <https://portal.research.lu.se/sv/projects/jobtech-dev-research-project-exploring-the-role-of-open-governmen>

- Trafiklab, ett nationellt data-ekosystem för delning och samverkan kring kollektivtrafikdata inom Sverige.
- HSL Developer Community, ett regionalt data-ekosystem för delning och samverkan kring kollektivtrafikdata inom Helsingforsregionen i Finland.
- City of Chicago Open Data program, ett lokalt data-ekosystem för delning och samverkan kring data som produceras inom samtliga förvaltningar inom City of Chicago.
- RoDL, Road Data Lab, ett data-ekosystem under uppbyggnad för att dela vägtrafikdata mellan privata aktörer, med tillämpningar inom maskininlärning för autonoma fordon.
- ESS-CSDL, European Spallation Space Control System Data Lab, som är en pilotstudie för delning av Industri 4.0-data som alarm i stora anläggningar, med syfte att skapa innovation kring de stora forskningsanläggningarna.

Denna rapport summerar en del av de resultat och lärdomar som producerats och anskaffats under forskningsprojektets gång. Summeringen tar sin form i ett antal konkreta rekommendationer för hur myndigheter kan gå tillväga för att utforma sina egna öppna data-ekosystem och hur de kan initiera och facilitera en värdeskapande samverkan inom dessa. Nedan ges en ingående bakgrundsbeskrivning runt fenomenet öppna data-ekosystem och samverkan kring öppna data. Härfter ges en kort och översiktlig beskrivning av den underliggande forskningsmetodiken följt av översiktliga beskrivningar av de undersökta fallen av öppna data-ekosystem. Slutligen ges en mer grundlig kontext till respektive rekommendation som kan hjälpa myndigheter att tolka och tillämpa dem.

Bakgrund

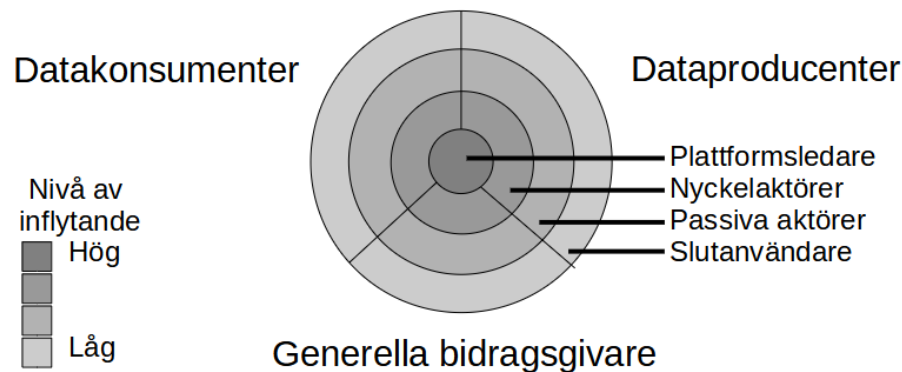
Samverkan kring öppna data inom ekosystem

Samverkan kring öppna data förekommer ofta i det som kallas öppna data-ekosystem, där OpenStreetMap och Wikidata är två populära exempel. Dessa ekosystem representerar en form av nätverkande gemenskaper av aktörer (organisationer och individer), som bygger relationer till varandra på ett gemensamt intresse (Runeson et al., 2021). Det gemensamma intresset för OpenStreetMap är att skapa och tillhandahålla gratis geografisk data. För Wikipedia är det att tillhandahålla en gratis kunskapsbas som kan läsas och redigeras av både människor och maskiner.

Ett öppna data-ekosystem stöds vanligtvis av en underliggande teknisk plattform som gör det möjligt för aktörer inom ekosystemet att t.ex. hitta, arkivera, publicera, konsumera eller återanvända data samt att främja innovation, skapa värde eller stödja nya företag. Aktörer samverkar kring data och relaterad programvara och standarder genom utbyte av information, resurser och artefakter (Linåker & Runeson, 2021). När det gäller OpenStreetMap och Wikidata är deras respektive plattformar öppna tillgängliga, inklusive programvara för att publicera och hantera data.

Flera aktörer kan vara involverade i samarbetet i ett öppna data-ekosystem vilket skapar en värdekedja som sträcker sig från dataproducenter till datakonsumenter (Zuijderwijk et al., 2014). Dessa kan i sin tur delas in ytterligare underkategorier som dataproducenter, datadistributörer, tjänsteleverantörer, apputvecklare, och slutanvändare (Immonen et al., 2014), även om ytterligare kategorier kan förekomma och överlappa (Oliveira et al., 2019). Dessa behövs för att få en fullt fungerande pipeline från data till användartjänst och kan fyllas av aktörer inom eller över organisationer. Datadistributören är vanligtvis en offentlig enhet eftersom delning av öppna data från privata organisationer ännu inte är ett vanligt fenomen (Runeson et al., 2021).

Styrningsstrukturer inom öppna data-ekosystem



Figur 1: Modell som kan visualisera och klassificera aktörer inom ett ekosystem utifrån dess nivå av inflytande och roll. Desto närmre mitten en aktör befinner sig, desto högre inflytande har den över ekosystemets inriktning, plattformens utveckling och den data som delas.

Från ett styrningsperspektiv (till skillnad från det funktionella som diskuterats ovan) brukar det särskiljas mellan tre typer av roller som visualieras i Figur 1 (Iansiti & Levien, 2014). Den första av dessa är *plattformsledaren* som äger och tillhandahåller plattformen och därmed också vanligtvis den som styr och faciliterar samverkan inom ekosystemet (Linåker & Runeson, 2020). Plattformsledare beslutar därmed över styrmodellen för ekosystemet, vilket betyder hur den upprätthåller kontrollen och beslutar om riktningen, men också om styrningsstrukturen, vilket avser fördelningen av rättigheter och skyldigheter mellan ekosystemets medlemmar, och de regler och protokoll som måste följas för att fatta beslut angående ekosystemet (Alves et al., 2017).

Nyckelaktörer och passiva aktörer är två andra roller inom ett ekosystem (Linåker & Runeson, 2020). En *nyckelaktör* är en som vårdar en symbiotisk relation med ekosystemet och dess andra aktörer (Iansiti & Levien, 2014). Vanligtvis har de en nära koppling till plattformsledaren. *Passiva aktörer* är fokuserade på en specifik nisch av marknaden, eller användningsfall, och är i första hand användare av den data som tillhandahålls av ekosystemet snarare än aktiva bidragsgivare i samverkan runt datan och relaterade resurser (Linåker & Runeson, 2020).

Ett sätt att illustrera styrningsstrukturen för ett öppna data-ekosystem är genom en s.k. lökmodellen (se Figur 1) där ekosystemets aktörer kan delas in i lager utifrån hur stort inflytande de har på ekosystemets styrning (Linåker & Runeson, 2020). Ju närmare aktören är kärnan, desto högre inflytande. Själva kärnan består av plattformsledaren som direkt eller indirekt bestämmer vilket inflytande ett ekosystems enskilda aktörer får ha, och därmed vilket lager aktören hamnar på. I de yttre lagren finns de aktörer som är mer passiva användare av datan, och längst ut slutanvändarna till de tillämpningar som skapas av aktörerna i de mellanliggande lagren.

Som illustreras i Figur 1 kan aktörer därutöver även delas in i tre grupper när det gäller deras användning av och bidrag till ekosystemets gemensamma resurser:

- En datakonsument använder den data som är tillgänglig via plattformen. Innefattar den vanlige slutanvändaren.
- En dataproducent bidrar med data, antingen aktivt via en donation till plattformen, eller passivt genom att låta plattformsledaren samla in data från dem, som sedan görs tillgänglig via plattformen.
- En generell bidragsgivare är inte nödvändigtvis en konsument eller producent av data, utan bidrar på något annat sätt till ekosystemets gemensamma resurser och hälsa, t.ex. genom kunskapsdelning eller genom att bidra med nya eller befintliga öppna programvaruprojekt som relaterar till den delade datan.

Rollen som plattformsledare kan vidare variera mellan organisations- och konsortiumcentrerad styrning, respektive nätverksbaserad styrning (Runeson et al., 2021). De tidigare nämnda exemplen med OpenStreetMap och Wikidata är exempel på ekosystem med nätverksbaserad styrning där styrningen är utspridd bland de individer som är medlemmar i ekosystemet. I organisations- och konsortiumcentrerade öppna data-ekosystem är styrningen koncentrerad till en enskild eller grupp av aktörer, vanligtvis offentliga eller privata organisationer med gemensamma affärsintressen i den data som delas inom ekosystemet. Exempel från

forskning finns inom domäner som arbetsmarknad, kollektivtrafik, smarta städer och industri 4.0, där både offentliga och privata organisationer innehar och delar styrning i olika konstellationer (Linåker & Runeson, 2021; Runeson et al., 2021).

Fördelar med samverkan kring öppna data

Samverkan kring öppna data har potential att generera värde likt för öppen programvara och andra typer av öppen innovation (Linåker & Regnell, 2020). Detta kommer som en effekt av att utnyttja den kunskap potentiella arbetskraften som finns inom och utanför ett ekosystem.

Ur ett kostnadsbesparande perspektiv kan den potentiella externa arbetskraften hjälpa till med uppgifter som att tillhandahålla, samla in, bearbeta och publicera data (Runeson et al., 2021). Som nämnts kan ett öppna data-ekosystem liknas vid en värdekedja där datan som råvara berikas och bearbetas på ett kollaborativt sätt (Lindman et al., 2015).

Ur ett kvalitetsperspektiv kan samarbete kring den öppna datan t.ex. bidra till att åtgärda och rätta till fel, men också tillföra information till datan genom annoteringar och andra typer av metadata (Runeson et al., 2021). Som en konsekvens kommer kvaliteten på träningsset för maskininlärning, liksom programvara och processer som använder data att förbättras.

Ur ett innovationsperspektiv kan den potentiella ökade tillgången till högkvalitativ data ytterligare bidra till att tillhandahålla nya och utökade träningsunderlag för maskininlärningsbaserade applikationer såväl som underlag till andra former av datadriven programvara (t.ex. Google Maps). Innovation kan ytterligare accelereras när nya användningsfall och marknader kan utökas eller skapas för aktuella ekosystem, eller till och med leda till att nya skapas. Samverkan kring öppna data kan dessutom sänka den relaterade inträdeskostnaden för aktörer som strävar efter att använda data för att erbjuda tjänster, hjälpa till att katalysera nya entreprenöriella ansträngningar, öka transparens och hjälpa organisationer fatta bättre beslut och effektivare tjänster (Attard et al., 2015).

Utmaningar med samverkan kring öppna data

Att dela och samarbeta om öppna data för med sig ett antal utmaningar, både tekniska och processororienterade, såväl som kulturella och affärsorienterade (Runeson et al., 2021). Nedan lyfter vi fram några av dessa utmaningar som utövare kan behöva överväga inom ett öppna data-ekosystem eller när de funderar på att gå in i eller skapa ett.

Affärs- och konkurrensaspekter

Ur ett affärsperspektiv handlar en viktig utmaning om motiveringen till varför en datamängd i första hand bör delas (Linåker & Runeson, 2021). Det måste finnas incitament som ligger i linje med en aktörs affärsmodell. Aktörer behöver därför förstå ovannämnda fördelar och kunna kontextualisera dessa i sin egen miljö och relatera till relevanta affärsmål.

Fördelarna måste även vägas mot de potentiella kostnaderna och riskerna som kan upplevas med att släppa data (Enders et al., 2020). Potentiella kostnader kan vara relaterade till livscykelhanteringen av datan, det vill säga insamling, bearbetning, kvalitetssäkring, delning och distribution av data (Runeson et al., 2021). Precis som med öppen programvara, hänför sig dessa kostnader såväl som de potentiella fördelarna till mängden samarbete som faktiskt äger rum. Därför måste aktörer inom ett öppna data-ekosystem möjliggöra en hållbar samverkan och delning av data.

En ytterligare utmaning som kan uppstå i ekosystem där konkurrenter ska samverka rör risken att ge bort eller förlora sin konkurrensfördel (Enders et al., 2020), vilket kan vara ett betydande hinder för kommersiell datadelning. För att hantera och möjliggöra ett sådant samarbete kan det finnas behov av en neutral plattformsledare inom ekosystemet som kan förmedla diskussioner, skapa en gemensam vision för ekosystemet och hjälpa aktörer att dela data som alla är bekväma med (Linåker & Runeson, 2021). Det senare, vanligtvis kallat selektivt avslöjande, kan t.ex. innebära att endast vissa abstraktioner av data delas (Enders et al., 2020).

Tekniska aspekter

De potentiella samverkansaspekterna väcker också ett antal mer tekniska utmaningar (Runeson et al., 2021). En sådan utmaning gäller insamlingen av data och hur man säkerställer dess kvalitet. Vissa forskare indikerar att datauppsättningar av otillräcklig kvalitet kan misstolkas eller missbrukas (Zuiderwijk & Janssen, 2014), snarare än att förbättra kvaliteten på programvaran. Gemensamma domänmodeller och standarder för hur data delas och används, samt transparenta processer och verktyg för insamling och berikning av data, bidrar till att möta denna utmaning (Runeson et al., 2021).

Att introducera feedback-loopar både inom ett öppna data-ekosystem och mot slutanvändarna är ett annat sätt (Rudmark & Andersson, 2022). Versionering av data är en annan praktisk aspekt som behöver beaktas för att möjliggöra decentraliserat samarbete, liknande det som kan observeras i samverkan runt utvecklingen av öppen programvara (Worthington et al., 2022).

Kulturella, organisatoriska och juridiska aspekter

Kulturella och organisatoriska aspekter utgör en annan typ utmaningar (Runeson et al., 2021), t.ex. att synkronisera strategiska och operativa nivåer inom en organisation kring beslut om vilken data som ska delas. Individer kan ha olika åsikter eller förståelse för de risker och fördelar som delning skulle innebära. Dessa utmaningar kan också manifesteras i insamlingen och samverkan kring data, eftersom återigen individer med olika bakgrund och kulturer kan ha olika perspektiv.

De legala aspekterna utgör ofta utmaningar för datadelning (Runeson et al., 2021). Organisationer kan vara ovilliga att dela data på grund av osäkerhet om ansvar och vad licenser kan innebära i praktiken. Risk för juridiska komplikationer på grund av GDPR är ett specifikt och vanligt problem enligt europeisk lagstiftning. Företag, till exempel, samlar in och underhåller betydande egenutvecklade personal- och kunddata som de kan vara ovilliga att dela genom samarbeten (Lindman et al., 2015). Att möjliggöra för individer att få kontroll över hur deras data delas kan vara ett sätt att ta itu med denna ganska svåra utmaning (Alorwu et al., 2022).

Metod och forskningsdesign

Denna rapport är baserad på forskning utförd och publicerad inom ramen för ett tvåårigt forskningsprojekt där forskare från Lunds universitet har studerat hur myndigheter kan dela och facilitera en samverkan kring data och relaterade resurser inom s.k. öppna data-ekosystem⁵. Projektet är huvudsakligen finansierat av Arbetsförmedlingen i syftet att ta utgångspunkt i deras egna, relativt nystartade öppna data-ekosystem JobTech Dev som beskrivs i mer detalj nedan. Ambitionen har varit att studera men också föra över kunskap på sådant sätt att samverkan inom JobTech Dev kan främjas och förbättras, samt att möjliggöra för andra myndigheter att dra lärdomar och inspireras av JobTech Devs resa.

Inom projektet har en forskare fungerat som följeforskare hos JobTech Dev och därigenom, förutom att studera dess utveckling, även kunna bidra med erfarenheter och kunskap från ny och befintlig forskning. För att skapa en referensram till JobTech Dev har ytterligare fem öppna data-ekosystem studerats där den förstnämnda tillsammans med två ytterligare forskare från Lunds universitet respektive RISE har samverkat. För en mer ingående presentation av dessa respektive ekosystem hänvisas läsaren till efterföljande kapitel. Studierna av de aktuella ekosystem finns publicerade i två vetenskapliga artiklar.

I den första artikeln studeras och jämförs JobTech Dev, Trafiklab, HSL DevCom och City of Chicago (Linåker & Runeson, 2021). Studien baseras primärt på intervjuer med personer från plattformsledarna inom respektive ekosystem samt erfarenheter från den följeforskning som bedrivits hos JobTech Dev. Resultaten utgör en serie rekommendationer för myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem, kring hur de kan främja en delning och samverkan runt data och relaterade resurser inom dessa ekosystem. Dessa rekommendationer utgör den

⁵ <https://portal.research.lu.se/sv/projects/jobtech-dev-research-project-exploring-the-role-of-open-governmen>

huvudsakliga grunden för rekommendationerna som presenteras i denna rapport.

I den andra artikeln studeras och jämförs JobTech Dev, ESS-CSDL och RoDL (Runeson et al., 2021). Även här baseras studien primärt på intervjuer med personer från de respektive plattformsledarna inom ekosystemen kompletterat med erfarenheter som anskaffats från den följeforskning som bedrivits hos JobTech Dev, men även inom de två övriga ekosystemen. Dessa fallstudier följde på av en serie fokusgruppsmöten där personer från privat och offentlig sektor fick diskutera utmaningar och vinster de såg med delning och samverkan kring öppna data. Resultatet från studien är en modell över de olika aspekter ett öppna data-ekosystem måste beakta för att främja en hållbar utveckling och samverkan kring gemensamma resurser.

I denna rapport presenteras en sammanställning baserad på dessa två underliggande studier (Linåker & Runeson, 2021; Runeson et al., 2021). För en mer ingående presentation av metod och resultat hänvisas läsaren till dessa.

Undersökta fall av myndighetsdrivna öppna data-ekosystem

JobTech Dev

Om ekosystemet

JobTech Dev⁶ är ett ekosystem som skapades 2018 för att samla aktörer som verkar inom eller i relation till den svenska arbetsmarknaden och skapa ett samarbete via en gemensam plattform av öppna data, anslutna API:er och kompletterande öppna programvaruprojekt. Ekosystemet och dess plattform är underhålls och faciliteras av Arbetsförmedlingen. Syftet med ekosystemet är bidra till en effektivare digital matchning och vägledning mellan arbetssökande och arbetsgivare på den Svenska arbetsmarknaden.

Aktörerna inom ekosystemet kan generellt kategoriseras inom ett av områdena: rekryterings- och bemanningsföretag, utbildnings- och vägledningsleverantörer, nationella, regionala eller lokala myndigheter, fackföreningar, arbetsgivarföreningar, jobbannonser och arbetssökande. Beroende på kategori kan en medlems intresse för plattformen vara begränsat till vissa delar av plattformen.

Om plattformen

Plattformen består av tre huvuddelar: arbetsmarknadsdata i form av jobbannonser, språkdata i form av taxonomi för kompetenser och yrkestitlar, samt individdata i form av ex. CV och arbetsstatus.

- Arbetsmarknadsdata i form av jobbannonser som samlas in genom ett samarbete med flera annonsleverantörer, varefter de berikas och bearbetas för att sedan tillgängliggörs via Platsbanken och ett för ekosystemet gemensamt API.
- Språkdata i form av en taxonomi för kompetenser, färdigheter och jobbtitlar och relationer mellan dem. Datamängden utvecklas och underhålls inom Arbetsförmedlingen. Datamängden tillgängliggörs via

⁶ <https://jobtechdev.se/sv>

API:er möjliggör för aktörerna i ekosystemet att ”tala samma språk” vilket bland annat medför en förbättrad rapportering och statistik samt bättre matchning mellan jobbannonser och arbetssökande.

- Individdata i form av en arbetssökandes personliga uppgifter och erfarenheter som den enskilde individen äger kontroll över. Individen kan exempelvis genom plattformen låta rekryteringsverktyg få tillgång till denna information för att inte behöva fylla i sitt CV om och om igen. Ett annat exempel är att låta försäkringsbolag ta del av sin arbetslöshetsstatus för att få arbetslöshetsersättning utbetald automatiskt.

Källkoden för API och algoritmer är tillgängliga som öppen programvara⁷. För att möjliggöra och främja tillämpning av datan data utvecklas och tillgängliggörs även exempelapplikationer som öppen programvara.

Trafiklab

Om ekosystemet

Trafiklab⁸ är ett ekosystem, skapat 2011, som samlar aktörer inom den svenska kollektivtrafiksektorn för att samarbeta genom en plattform med öppen trafikdata, med tillhörande API:er och kompletterande öppna programvaruprojekt. Ekosystemets vision är att underlätta skapandet av nya tjänster som gör det enklare och mer attraktivt att resa med kollektivtrafiken. Ekosystemet och dess plattform utvecklas och faciliteras av Samtrafiken, ett företag som samägs av alla regionala kollektivtrafikmyndigheter och de flesta kommersiella trafikoperatörer i Sverige. Operatörer har även möjlighet att vara partner till Samtrafiken utan att vara delägare.

Generellt kan aktörerna inom ekosystemet kategoriseras inom något av områdena: regionala kollektivtrafikbolag, privata och

⁷ <https://gitlab.com/arbetsformedlingen>

⁸ <https://www.trafiklab.se>

offentligt ägda tågoperatörer, nationella, regionala och lokala myndigheter, privata bussoperatörer, samt privata produkt- och tjänsteleverantörer. Planerna omfattar integration med relaterade aktörer, såsom taxi- och mobilitetsoperatörer.

Om plattformen

Plattformen består av datamängder och API:er, antingen underhållna av Samtrafiken eller oberoende av aktörer inom ekosystemet. All data som lagras på Trafiklab-plattformen släpps med en anpassad licens baserad på principerna för CC0 1.0-licensen.

Fyra API:er ger statisk och realtidsdata om kollektivtrafiken, exempelvis kopplat till tidtabeller och avbrott. Dessa uppgifter görs idag tillgängliga i två typer av standardformat, underhållna av Samtrafiken och inhämtade från de regionala kollektivtrafikbolagen och privata operatörer enligt myndighetsdirektiv. Ytterligare två API:er tillhandahåller tidtabellsdata för en reseplanerare, en externt upphandlad produkt som erbjuds som en gratistjänst till ekosystemet.

Vissa API:er underhålls av andra organisationer, både offentliga och privata, och görs tillgängliga på Trafiklab-plattformen. Data omfattar tidtabells- och servicedata från Stockholms kollektivtrafik samt trafikinformation från Trafikverket. Plattformen länkar också till relaterade API:er som underhålls och är tillgängliggörs hos andra organisationer.

När det gäller öppen programvara har de för närvarande ett antal utvecklingskits och exempelapplikationer tillgängliga⁹. De har för avsikt att utveckla en ny reseplanerare som öppen programvara och dela sina API:er som öppet tillsammans med de olika delarna av plattformen.

⁹ <https://github.com/trafiklab>

HSL Developer Community

Om ekosystemet

HSL Developer Community¹⁰ (HSL DevCom) är ett ekosystem, skapat 2009, som samlar aktörer med intresse för öppna data och öppen programvara kopplat till Helsingforsregionens kollektivtrafik (jämfört med den nationella inriktningen hos Trafiklab). Ekosystemet och dess plattform utvecklas och faciliteras av Helsingfors Regionaltrafik (HSL enligt finska översättningen), ett bolag som samägs av nio kommuner i Helsingforsregionen i Finland. HSL ansvarar bland annat för planering och organisering av regionens kollektivtrafik.

I likhet med Trafiklab kan aktörerna inom ekosystemet i termer av verksamhet generellt kategoriseras som: nationella, regionala och lokala kollektivtrafikbolag, myndigheter, men också privata buss- och tågoperatörer, samt privata produkt- och tjänsteleverantörer. Eftersom användningen av reseplaneraren sträcker sig bortom Finland täcker medlemmarna av ekosystemet även norska, svenska, tyska och italienska marknader (i olika utsträckning).

Om plattformen

Plattformen består av flera datamängder och anslutna API:er samt ett antal relaterade öppna programvaruprojekt som är centrala för ekosystemaktörerna, både vad gäller intern verksamhet och som kundnära applikationer. All data är tillgänglig under CC-BY 4.0-licensen.

Huvuddelen av plattformen är relaterad till en reseplanerare som är utvecklad som öppen programvara¹¹. Reseplaneraren används av resenärer som planerar sina resor inom regionens kollektivtrafiksystem. Reseplaneraren var HSL:s första öppna programvaruprojekt och släpptes 2013. Kärnan av reseplaneraren

¹⁰ <https://dev.hsl.fi/>

¹¹ <https://github.com/HSLdevcom>

utgörs av ett annat öppet programvaruprojekt genom Open Trip Planner som utför ruttplanering baserat på ingående data.

Reseplaneraren använder i sig av underliggande öppna API:er som bl.a. tillhandahåller resvägs- och tidtabellsdata, samt geokodningsdata med adresser och koordinater. Därtill tillhandahålls kartdata som HSL hämtar från OpenStreetMap (OSM) och bearbetar, ex. genom skapande av ruttmönster och visualiseringar med hjälp av HSLs egen kartstil.

Andra API:er som tillhandahålls via plattformen omfattar:

- Bluetooth-fyrar installerade på bussar, busshållplatser och terminaler för tåg och tunnelbanor. Ett användningsfall är att de ger en överblick över antalet passagerare som väntar på eller för närvarande genomgår i skjuts. Ett annat användningsfall är att ge bättre information om den aktuella resan, och tillhandahålla trafikinformation och potentiella alternativa resealternativ. Uppgifterna är framtagna och publicerade av HSL.
- Resor gjorda med stadscyklar i Helsingfors. Uppgifterna innehåller information om resans start- och slutpunkter och används huvudsakligen för analys. Uppgifterna produceras av City Bike Finland och laddas upp till HSL en gång i månaden som publicerar informationen på plattformen.
- Parkeringsinformation för pendlare. Uppgifterna publiceras av kommuner och operatörer via ett webbläsarbaserat användargränssnitt på HSL:s webbplats. Uppgifterna publiceras sedan via API och på plattformen av HSL.

City of Chicago Open Data program

Om ekosystemet

City of Chicago startade sitt ekosystem 2011 genom lanseringen av stadens öppna data-program och dataportal¹². Ekosystemet i stort täcker hela staden, inklusive dess medborgare, civilsamhälle och

¹² <https://data.cityofchicago.org/>

näringsliv. Huvudmålet är att göra stadens interna data användbaroch konsumerbar av allmänheten. Den första drivkraften för initiativet var att öka transparensen, vilket med tiden utvecklades till ett högre fokus på att möjliggöra social och ekonomisk utveckling. Ett annat viktigt motiv är att engagera allmänheten för att förbättra sitt samhälle.

Aktörerna inom ekosystemet är tvärsektoriella och involverar forskningsinstitutioner, företag, civilsamhälle, journalister och medborgare. Eftersom det finns en stor variation i datamängderna som tillhandahålls finns det ingen specifik bransch eller område att klassificera dessa medlemmar i, som med de andra exemplen som presenteras i denna rapport, till exempel aktörer som är kopplade till den svenska arbetsmarknaden som i exemplet med JobTech Dev.

Om plattformen

Plattformen består av cirka 1400 olika datamängder tillgängliga via API:er. Uppgifterna kommer ur administrativa processer som utförs av Chicago City. Det finns därför en stor variation i typen av data som sträcker sig från varje tillstånd som utfärdas av staden, varje brott som rapporteras, till varje resa som tas via mobilitets- och samåkningsoperatörer. Den mest populära datamängden var dock listan över löner och alla anställda som arbetade för staden. Uppgifterna släpps under en anpassad licens.

Vad gäller öppen programvara har staden ett 80-tal publicerade öppna programvaruprojekt på sitt GitHub-konto¹³, där ett mindre antal är mer aktiva och använda än andra. Ett genomgående tema är att det finns en koppling till den data som publiceras på stadens dataportal. Två av de mer populära projekten består av algoritmer, som tillsammans med data från portalen stödjer beslut om vart man ska skicka livsmedelsinspektörer i staden, respektive förutsäga höga nivåer av E.coli-bakterier i Lake Michigan. Ett annat projekt gäller OpenGrid, en webbaserad karta som låter användaren utforska flera datamängder samtidigt.

¹³ <https://github.com/Chicago>

Road Data Lab (RoDL)

Om ekosystemet

Road Data Lab¹⁴ (RoDL) är ett Vinnova-finansierat innovationsprojekt som syftar till att etablera ett öppna data-ekosystem där aktörer på och i anslutning till vägar kan dela data för att möjliggöra studier av maskininlärning, till exempel sensordata från fordon eller väglagsdata insamlad från Trafikverket. I utvecklingen mot autonom körning har kostnaderna för att samla in och annotera data för maskininlärningsändamål identifierats som en betydande utmaning, och därför undersöker aktörer möjligheterna att minska kostnaderna, samtidigt som de behåller sin konkurrensfördel av data.

Kärnan i ekosystemet består av AI Sweden (Sveriges nationella centrum för tillämpad artificiell intelligens) som plattformsledare, och de två nyckelaktörerna Zenseact (ett mjukvaruföretag för autonom körning och säkerhet, ägt av Volvo Cars) och Univrses (ett nystartat företag med expertis inom datorseende i smarta stadsapplikationer). Andra projektmedlemmar inkluderar RISE Research Institutes of Sweden och Lunds universitet, som utökar kompetensen inom öppen innovation, maskininlärning och data-ekosystem. Det pågår diskussioner med organisationer inom trafik och transportdomänen för att delta i samarbetet, t.ex. Trafikverket, som redan driver ett datadelningsprojekt med billeverantörer för att samla in friktionsdata i säkerhetssyfte.

Om plattformen

Projektet utforskar tekniska, juridiska och organisatoriska frågor i relation till att dela data mellan partners, samt att lägga till dess värde. Till exempel kan en partner tillhandahålla råvideodata och en annan partner kan lägga till kommentarer till videodata. Denna samverkan faciliteras och samordnas via en teknisk plattform som tillhandahålls av AI Sweden. Zenseact har delat datamängder som kan användas fritt

¹⁴ <https://www.vinnova.se/p/road-data-lab/>

för att utforska tillämpningar. Lärdomar från detta, såväl tekniska som juridiska, är en viktig del av resultaten.

ESS-CSDL

Om ekosystemet

European Spallation Source (ESS) är ett Europeiskt infrastrukturkonsortium för forskning, som för närvarande bygger en multidisciplinär forskningsanläggning baserad på världens mest kraftfulla neutronkälla. ESS är ett paneuropeiskt projekt med 13 europeiska nationer som medlemmar, inklusive värdnationerna Sverige och Danmark, vilket i sig är en utmanande samarbetsutmaning.

ESS Control System Data Lab (ESS-CSDL) fokuserar på data från det integrerade styrsystemet (ICS) som är designat för att övervaka och kontrollera cirka 100 000 enheter i ESS-anläggningen med ett styrsystemrelaterat dataflöde på cirka 50 Gigabyte per sekund. ESS, även om det är en forskningsanläggning, är i nivå med, eller till och med i framkanten av, Industri 4.0-anläggningar när det gäller komplexitet och tekniska lösningar. ESS är för närvarande under uppbyggnad, men redan idag finns betydande data tillgänglig från den successivt startande anläggningen. ESS-CSDL¹⁵ har till syfte att etablera ett öppna data-ekosystem för larndata, för att katalysera innovation genom samverkan kring data med externa partners, särskilt genom att dela erfarenheter, men också för att göra larndata med hög kvalitet tillgänglig för forskning och utveckling inom maskininlärning.

Kärnan i ekosystemet utgörs av ESS ICS, kompletterat med larmsystemexperter från GoalArt. Forskare från Lunds universitet observerar och ger råd om projektet med avseende på datateknik och samverkanspraxis. ESS-CSDL-projektet har en referensgrupp, sammansatt av industrirepresentanter från domänerna automation, robotteknik, kemisk bearbetning, förpackning och datavetenskap. Det

¹⁵ <https://portal.research.lu.se/en/projects/ess-data-lab>

finns också en medlem från en systemanläggning i Tyskland, DESY (Deutsches Elektronen-Synchrotron) etablerad 1959, villig att dela med sig av sina erfarenheter och lära sig tillbaka, samt representanter från universitetets innovationskontor. Referensgruppsmedlemmarna kan ses som passiva aktörer i ekosystemet. Det långsiktiga målet för gemenskapen är att utveckla passiva medlemmar till att bli aktiva slutanvändare av datan men även nyckelaktörer som kan bidra till det gemensamma värdeskapandet.

Om plattformen

Plattform för delning och samverkan saknas idag men är utpekat som ett behov för att möjliggöra en framtida och hållbar delning och samverkan kring data inom ekosystemet.

Rekommendationer för att skapa och facilitera öppna data-ekosystem

Styrningsstrukturer för ekosystem och ägande av plattform

Med ett ekosystems styrningsstruktur avses fördelningen av rättigheter och ansvar bland ekosystemets aktörer, samt regler och processer som behövs för att kunna fatta beslut angående datan som delas och den samverkan som främjas. Styrningsstrukturen bestäms vanligtvis av plattformsledaren som är den aktör som äger och underhåller plattformen via vilken datan delas, och därmed utgör grunden för ekosystemet. En differentiering av ägarskapet till plattformsledaren kan i sin tur göras mellan enskilt ägande och samägande, samt konstellationen mellan offentliga och privata aktörer. Utifrån de fall där det finns minst en offentlig aktör inblandad i ägarskapet, kan tre situationer förekomma (se Tabell 1: Översikt över olika ägandeformer av plattformsledaren inom myndighetsdrivna öppna data-ekosystem.).

Tabell 1: Översikt över olika ägandeformer av plattformsledaren inom myndighetsdrivna öppna data-ekosystem.

	Offentlig sektor	Offentlig och privat sektor
Enskilt ägande	JobTech Dev, City of Chicago Open Data program, ESS-CSDL	–
Samägande	HSL Developer Community	Trafiklab, RoDL

I ekosystem med offentligt enskilt ägande är det en enskild myndighet som äger eller utgör plattformsledaren som i fallet med Arbetsförmedlingen och deras ekosystem JobTech Dev, samt City of Chicago med deras Open Data program.

I ekosystem med ett offentligt samägande är det två eller fler myndigheter som äger eller utgör plattformsägare i ett konsortium. Ett

exempel går att finna i HSL Developer. Plattformsledaren HSL/HRT är ett företag som är samägt av kommunerna inom regionen.

I ekosystem med offentligt-privat ägande är det två eller fler offentliga och privata (eller civilsamhälles-) aktörer som äger eller utgör plattformsägare i ett konsortium. Ett exempel utgörs av Trafiklab som är ett data-ekosystem som är fokuserat på den svenska kollektivtrafiken. Plattformsledaren Samtrafiken är ett företag som är samägt av de regionala kollektivtrafikbolagen tillsammans med privata operatörer.

Val kring styrningsstruktur och ägande bör särskilt beakta vilka nyckelaktörerna är och hur deras inflytande kan säkras och främjas. I fallen med samägarskap av plattformsledaren kan detta inflytande säkras på formell väg. I ekosystem med enskilt ägarskap måste detta inflytande säkras genom en mer informell struktur där plattformsägaren (som Arbetsförmedlingen inom JobTech Dev) aktivt måste möjliggöra dialog och samverkan med nyckelaktörerna. Sett från exemplen ovan kan nyckelaktörer utgöras av ägarna själva, formella eller informella partners, producenter till betydelsefulla datamängder, organisationer med stora användarbaser, tillika organisationer och individer med ett aktivt engagemang inom ekosystemet.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R1:** Överväga hur ägarskapet till av den organisation som utgör plattformsledare bäst kan anpassas för att främja samverkan och engagemang med och mellan nyckelaktörer inom ett ekosystem.
 - Om ägarskapet innehas av en enskild offentlig organisation behöver en aktiv dialog skapas med nyckelaktörer samt att dessa får möjlighet att utöva inflytande på plattformens innehåll och utveckling genom en öppen samverkan.
 - Om ägarskapet delas mellan ett antal offentliga och/eller privata aktörer behövs samverkan

möjliggörs genom ägarskapet, men även mot de aktörer som ej ingår i ägarskapet men som fortfarande är med i ekosystemet.

- **R2:** Identifiera nyckelaktörer inom ekosystemet och möjliggöra för dessa att ha ett informellt eller formellt inflytande över ekosystemets inriktning, plattformens utveckling och den data som delas. Vilka aktörer som utgör nyckelaktörer kan variera men kan inkludera plattformsledarens ägare, dess partners, organisationer med t.ex. stora mängder användare eller andra värdefulla tillgångar (som data och programvara), men även medborgare och slutanvändare.

Medling och neutralitet som grund för ett sammanhållet ekosystem

Att lita på datakällor och lita på att andra organisationer använder delad data på ett korrekt sätt kan upplevas som ett bekymmer (Runeson et al., 2021; Alorwu et al., 2022). Att bygga förtroende i ett distribuerat och mångfacetterat nätverk är svårt, och därför har vissa hypoteser att det kan finnas en roll för en centraliserad funktion för att säkerställa kvaliteten och tillförlitligheten hos data. För öppna data från myndigheter är det själva myndigheten som utgör denna garanti (som i fallet med JobTech Dev), medan vi i peer-to-peer-delning har sett få alternativ till de stora teknikaktörerna. ”Data trusts” – ”en juridisk struktur som tillhandahåller oberoende förvaltning av data” – föreslås som en lösning av Open Data Institute i deras rapport ”The Value of Data” (2020).

Tillit handlar vidare bara inte om datan i sig utan även i relation till dess användare och övriga aktörer inom ett ekosystem. Oron för att ge bort affärsfördelar nämndes flera gånger i fokusgrupperna, det vill säga att skapa mer affärsvärde för andra. Till exempel kan andra organisationer vara snabbare på att utveckla sina produkter och tjänster eller att andra organisationer kan hitta affärsvärde i data som du inte hittade. Därför är de mer benägna att arbeta med organisationer som inte är konkurrenter. Även i detta fall framstår behovet av en neutral aktör som övriga kan känna tillit till och som

aktivt kan medla och facilitera diskussioner kring datadelning och samverkan kan realiseras.

Som konkret exempel kan nämnas JobTech Dev där det krävdes en längre tid av förhandlingar från Arbetsförmedlingens sida för att få med sig ett kritiskt antal jobbanonsleverantörer till att dela sin annonsdata. Bland aktörerna rådde dels en skepsis till Arbetsförmedlingens intentioner med datan, dels mot att samverka och dela data med direkta konkurrenter. Med tiden kunde dock ett förtroendekapital byggas upp och en kompromiss nås gällande vilken data som delas mellan aktörerna. Kompromissen bestod praktiskt av att endast tillhandahålla en vissa delar av annonserna, vilket innebär att endast befattningen, metadata och en länk till den ursprungliga annonsen skulle tillhandahållas via plattformens API.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R3:** Se till att det finns en neutral plattformsledare inom ekosystemet som kan bygga förtroende och tillit hos ekosystemets övriga aktörer, samt möjliggöra samverkan mellan dessa trots ev. utmaningar, som konkurrensförhållanden mellan vissa (nyckel)aktörer.
- **R4:** Ta hänsyn till, diskutera och leda arbetet med att hantera ev. bekymmer och risker som aktörer ser med delning och användning av data, teknik och annan kunskap inom ekosystemet.

Ekosystemets syfte och incitament för samverkan och datadelning

Ekosystemets syfte kan ses som anledningen till att ekosystemets aktörer väljer att samverka och i förlängning hur detta anses skapa ett mervärde kontra att arbeta själv. Syftet definieras därav både av plattformsledaren som faciliterar ekosystemet och tillhandahåller dess underliggande plattform, och av de aktörer som ingår och tillkommer ekosystemet. Sker dessa beslut isolerat hos plattformsledaren och denne representerar en minoritet av ekosystemet, finns risk för att

syftet ej ligger i linjer med majoriteten och att dessa lämnar ekosystemet som konsekvens. Ekosystemets syfte och omfattning bör därför vara något som utvecklas organiskt med tiden och i dialog mellan dess aktörer.

Från de studerade ekosystemet har vi kunnat se olika bakomliggande drivkrafter som format dess syften. För RoDL och ESS-CSDL finns det ett huvudsakligt affärsdrivet syfte där det övergripande målet är att skapa kommersiellt gångbara tillämpningar på den delade datan. Detta trots att ekosystemen i olika grad är offentligt stödda. I kontrast kan nämnas t.ex. Trafiklab och JobTech Dev som har ett mer samhällsdrivet syfte med de övergripande målen att skapa en mer attraktiv och tillgänglig kollektivtrafik, resp. effektivare digital matchning och vägledning på den svenska arbetsmarknaden. Det ska dock nämnas att det finns kommersiella inslag då båda ekosystem vill främja skapandet av nya och förbättrade produkter och tjänster inom respektive område som kan bidra till det övergripande målet.

Oavsett bakomliggande drivkrafter är det viktigt att aktörer ser ett värde med att ingå i ekosystemet. Ett värde som inte nödvändigtvis behöver ligga i datan som delas som vi observerat inom ekosystemet ESS-CSDL. ESS sätter här värdet av samarbetet före det rena utbytet av data. De ansåg att deras egna data var av tillräcklig kvalitet och kvantitet för det primära syftet, samtidigt som de söker erfarenheter och kunskap om liknande anläggningar och data. Inom RoDL ser aktörerna åtminstone på kort sikt, att det finns ett värde i att få tillgång till mer data genom att delta i ekosystemet. Långsiktigt kan samarbetet också vara ett mål, inte bara ett medel.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R5:** Upprätthålla en tydlig dialog med ekosystemet angående dess övergripande syfte och främja ett värdeskapande hos aktörerna, både sett till att dela data och att delta i själva samverkan och kunskapsutbytet däromkring.

- **R6:** Skapa gemensamma incitament trots olika affärsmodeller som tillämpas av ekosystemets aktörer, oavsett om ekosystemet primärt kan ses som affärs- eller samhällsdrivet, och hur datan eller samverkan utgör en komponent där i.

Ekosystemets geografiska och demografiska omfattning

Med ekosystemets omfattning menar vi den geografiska och demografiska täckning som den delade datan avser. Likt syftet, definieras omfattningen av plattformsledaren antingen i direkt eller indirekt dialog och samförstånd med övriga ekosystemet.

Om vi tittar på geografisk täckning kan vi notera att JobTech Dev och Trafiklab har nationell täckning av Sverige, medan HSL DevCom och City of Chicago har en mer regional och kommunal täckning. Men även om huvudaktörerna kan vara begränsade geografiskt, behöver inte samma begränsningar gälla alla aktörer inom ekosystemet. När det gäller HSL DevCom är internationella företag med kart- och reseplaneringstjänster som Google och Apple till exempel intresserade av kollektivtrafiken på global nivå, inte bara i Helsingforsregionen i Finland.

Med tanke på den demografiska täckningen har Trafiklab, HSL DevCom och JobTech Dev ett sektorsspecifikt fokus på kollektivtrafiken respektive den svenska arbetsmarknaden. Men även om uppgifterna kan vara begränsade behöver användningsfall inte vara det. Dessa kan vara generiska eller placerade i en separat domän där data från den berörda plattformen bara är en pusselbit. Däremot har City of Chicago ett mycket allmänt fokus när det gäller dess ekosystem eftersom dess data och öppna programvaruprojekt kan relatera till ett brett spektrum av tjänster och funktioner som staden levererar.

Utöver aktörer och användningsfall, kan det även förekomma att nya dataproducenter tillkommer (eller faller ifrån) vilket kan påverka både den demografiska och geografiska omfattningen. Som exempel kan nämnas HSL DevCom som idag består av nio kommuner men där

förhoppningen är att ytterligare kommuner inom regionen ska tillkomma.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R7:** Upprätthålla en tydlig dialog med ekosystemet angående dess omfattning och framtida utveckling. I detta arbete bör plattformsledaren beakta och ev. inkludera aktörer vars användningsfall kan sträcka sig utanför ekosystemets avsedda geografiska och demografiska täckning.

Funktioner för delning och samverkan inom ekosystemet

På en generell nivå kan aktörer i ett data-ekosystem delas in i de två grupperna dataproducenter respektive datakonsumenter. Dock kan ofta en mer funktionell indelning göras av aktörerna som nämnts ovan, exempelvis inom rollerna dataproducenter, dataintermediärer, tjänsteleverantörer, applikationsutvecklare, verktygs- och infrastrukturleverantörer samt slutanvändare.

Dataproducenterna är de aktörer som producerar datan och delar den vidare ekosystemet. Vanligtvis inom dessa ekosystem utgörs dataproducenterna av offentlig aktörer då datadelning bland privata aktörer fortfarande är under tidig utveckling. Dataintermediärer agerar vanligen som en länk mellan dataproducenter och datakonsumenter genom att samla in och dela vidare datan. Tjänsteleverantörer kan t.ex. erbjuda tjänster kopplade till berikning, processering, analys och visualisering av datan för andra aktörer inom ekosystemet, inklusive slutanvändarna. Applikationsutvecklarna utvecklar tillämpningar som använder sig av datan för att sedan säljas till slutanvändare. Verktygs- och infrastrukturleverantörer tillhandahåller medel som kan hjälpa övriga aktörer inom ekosystemet att använda datan och utveckla sina respektive tillämpningar med koppling till datan.

Värt att notera är att en aktör kan inneha flera av dessa roller samtidigt. Ett exempel utgörs av Arbetsförmedlingens data-ekosystem

JobTech Dev. Arbetsförmedlingen tillgängliggör här både egenproducerade data (t.ex. taxonomi för yrken och kompetenser) och insamlad data (t.ex. jobbannonser) som de sedan tillgängliggör via en plattform med tillhörande API:er. Vissa API:er utgör i sig tjänster som berikar och processar datan som bland annat består av jobbannonser. De utvecklar och tillhandahåller även verktyg och applikationsexempel som öppen programvara för att hjälpa aktörerna inom ekosystemet att arbeta med datan och utveckla sina respektive tillämpningar på den. Slutligen utvecklar och tillhandahåller de även själva produkter och tillämpningar som använder sig av den insamlade och egenproducerade datan, som ex. Platsbanken.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem rekommenderar vi att:

- **R8:** Överväga behoven hos ekosystemets aktörer och utifrån detta ta sig an funktioner för att bemöta dessa för att främja delning och samverkan inom ekosystem. Funktioner kan inkludera insamling, bearbetning och publicering av data, samt tillhandahållande av verktyg och infrastruktur för att kunna använda och bearbeta datan, och tjänster och applikationer som kan möjliggöra olika användningsfall baserade på datan.

Verktyg och processer för öppen samverkan

För att möjliggöra en öppen samverkan inom ekosystemet måste plattformsledare själv öppna upp sina processer och kommunikation för ekosystemet.

Öppna kommunikationskanaler är ett sådant sätt genom vilka plattformsledaren och övriga aktörer inom ekosystemet kan samarbeta och kommunicera med varandra. Alla de fyra studerade plattformsledarna har sådana kanaler, antingen synkrona eller asynkrona, kompletterade med alternativa kanaler som e-post, sociala medier och supportdesk. Dessa kanaler fungerar vanligtvis som platser för att lägga upp frågor, be om hjälp men också rapportera buggar och begära ny funktionalitet eller data.

Öppen kravhantering är ytterligare ett sätt som innebär att plattformsledaren öppet kommunicerar vilka krav som den arbetar med. Det innebär också att göra det möjligt för ekosystemets aktörer att kommentera och begära funktioner mer formellt, till exempel genom öppna ärendehanteringssystem. Inget av de studerade ekosystemen har ett centralt sådant utan använder sig framförallt sådana som finns tillgängliga via externa samverkansplattformar som GitHub eller GitLab i relation till aktuella öppna programvaruprojekt. När det gäller att kommunicera mer långsiktiga planer är det bara staden Chicago som publicerar dessa för sina öppna programvaruprojekt.

Aktivt engagemang är ett tredje sätt för att främja en öppen samverkan inom ekosystemet vilket innebär att plattformsledaren behöver kommunicera aktivt med övriga aktörer, och även engagera, initiera och underlätta samarbeten inom ekosystemet. Utöver att använda de öppna kommunikationskanalerna används fysiska och virtuella möten som hackathons, meetups och konferenser ofta bland alla de studerade plattformsledarna. Denna aspekt är särskilt viktig innan en ny datamängd eller ett öppet programvaruprojekt släpps. Till exempel rapporterade HSL att de hade problem med att få externa bidrag då de varit för fokuserade på utvecklingen och mindre på externt engagemang. Som ett motexempel kan man betrakta det arbete som Arbetsförmedlingen har lagt ner på att förbereda jobbbannonssamarbetet.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R9:** Möjliggöra en öppen samverkan inom ekosystemet genom att tillhandahålla öppna kommunikationskanaler, facilitera en öppen kravhanteringsprocess och aktivt engagera aktörer inom ekosystemet i samarbetet. Öppna och vedertagna verktyg bör väljas i största möjliga mån.

Öppen programvara som komplement och katalysator

Bland de studerade fallen har HSL ett starkare fokus på öppen programvaruutveckling då vissa öppna programvaruprojekt har en mer central roll i HSL DevComs ekosystem och riktar sig till slutanvändarna. I de andra ekosystemen fungerar de utvecklade öppna programvaruprojekten främst som hjälpmedel för utvecklarna i ekosystemen. I de studerade fallen har vi identifierat sex kategorier av öppna programvaruprojekt som plattformsledaren och ekosystemet aktörer kan överväga att initiera och samarbeta kring:

- **Datainsamling:** För JobTech Dev använder Arbetsförmedlingen öppen programvarubaserad teknik (utvecklad och underhållen av en tredjepartsleverantör) för att samla in platsannonser från de olika platsannonsleverantörerna.
- **Databerikning:** Öppen programvara kan också användas för att berika tillhandahållen data med metadata eller på andra sätt tillföra värde till den. När det gäller JobTech Dev tillhandahåller Arbetsförmedlingen internt framtagen funktionalitet, dels som ett API och dels som öppen programvara, som hjälper aktörer att berika jobbannonser med metadata.
- **Inspiration och dokumentation:** Alla fyra studerade plattformsledare utvecklar och underhåller applikationsexempel som kan hjälpa användare att förstå hur data kan användas. I denna mening kan exemplen ses som en källa till dokumentation. Men de kan också inspirera till nya användningsområden som kan vara av värde både för nya och etablerade företag.
- **Utvecklingsstöd:** Alla fyra plattformsledare ser också värde i att tillhandahålla gemensamma verktyg, bibliotek och ramverk som kan göra det möjligt för aktörer att skapa nya produkter och tjänster baserat på data.
- **Algoritmer:** City of Chicago publicerar algoritmer som används inom staden, dels för transparens, men också för att

få hjälp med att göra algoritmerna mer effektiva och korrekta. Arbetsförmedlingen publicerar sina algoritmer för externa aktörer att testa och validera de resultat de producerar.

- Produkter och tjänster: HSL tar ett steg längre och utvecklar även öppna programvaruprodukter som kan användas både av tjänsteleverantörer, trafikoperatörer och slutanvändare (dvs. pendlarna).

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem rekommenderar vi att:

- **R10:** Utveckla och samarbeta med ekosystemets aktörer om öppen programvara som kan förbättra användning och innovation baserat på data. Användningsfall kan innefatta insamling och berikning av data, inspiration och dokumentation, eller allmänt stöd till utvecklare genom verktyg, bibliotek och infrastruktur. Plattformsledaren kan också överväga att dela sina applikationer som använder data som input, vilket kan inkludera algoritmer, såväl som verktyg och konsumentinriktade produkter.

Insamling och delning av data inom ekosystemet

För att samverka kring data måste plattformsledaren överväga vilken typ av data som delas (eller skulle kunna) delas och tillhandahålls genom plattformen. Särskilt gäller att se till hur datan kan samlas in där vi särskiljer på fyra olika fall med internt producerad data, aktiva databidrag, passiva databidrag och extern ekosystemdata.

- Internt producerade data avser data producerad av plattformsledaren. I det här fallet blir plattformsledaren den som underhåller datan och behöver överväga hur de vill att aktörer ska kunna bidra till det. I fallet med JobTech Dev och deras taxonomidata kan aktörer främst bidra med feedback men det finns exempel där aktörer har bidragit

med datamängder som införlivas i den delade datamängden manuellt.

- Aktiva databidrag innebär data som aktörer inom ekosystemet aktivt bidrar till plattformen via verktyg eller infrastruktur som tillhandahålls av plattformsledaren. I det här fallet underhåller de bidragande aktörerna sin data och kan själva välja när de ska uppdatera vad som finns på plattformen. Som exempel från de studerade fallen är Career-initiativet på JobTech Dev-plattformen, där individer lagrar sina CV-data, och stadscykel- och bilparkeringsdata, som bidrar till HSL DevCom-plattformen.
- Passiva databidrag innebär data som plattformsledaren samlar in från ekosystemaktörerna. Originaldata underhålls av ekosystemaktören, men vad som finns tillgängligt på plattformen är beroende av plattformsledarens rutiner för insamling och uppdatering. Ett exempel är trafikdata från de regionala trafikhuvudmännen som samlas in av Samtrafiken. Ett andra exempel är platsannonserna från jobbannonsleverantörerna som samlas in av Arbetsförmedlingen. Ett tredje exempel är de nationella uppgifterna som samlas in och bryts ner på samhällsnivå av Chicago City.
- Externa ekosystemdata avser data som underhålls i ett annat data-ekosystem och samlas in av plattformsledaren och görs tillgänglig på plattformen, på ett direkt eller berikat sätt. Ett exempel gäller OpenStreetMaps kartdata, som importeras av HSL och används som input till deras andra datamängder och API:s. Om HSL eller någon av ekosystemaktörerna vill bidra till kartdata kan de vända sig direkt till OpenStreetMap-projektet.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R11:** Underlätta och möjliggöra delning av data, utöver den egna. Detta kan göras antingen genom att hjälpa aktörerna att aktivt eller passivt bidra med sin data. I det aktiva fallet använder aktörerna en teknisk infrastruktur tillhandahållen av plattformsledaren. I det passiva fallet samlar plattformsledaren in data från aktören för att sedan publicera den på plattformen. En ytterligare aspekt är att också överväga hur de kan använda externa ekosystem för att samarbeta kring data.

Balansering av värde och risk genom form på delade data

Det kan finnas många anledningar till att data ej kan delas öppet inom ett ekosystem. Risker kan t.ex. innefatta att ge upp konkurrensfördelar, möjliggörande av oetiska användningsfall, samt skada för integritet och sekretess på individ-, organisations- och samhällsnivå. För att data ska kunna delas måste därför sådana risker hanteras samtidigt som möjligheter för värdeskapande inte får gå förlorade.

I studierna har vi sett tre aspekter hos data som kan hjälpa aktörer inom ett ekosystem att kompromissa och anpassa den data som delas. Här skiljer vi mellan datans granularitet, grad av bearbetning och aktualitet.

- Granularitet avser abstraktionsnivån hos datan. Data som abstraherar från individer är mindre känsliga. Värdet kan fortfarande behållas om känsliga aspekter anonymiseras. Som exempel kan nämnas JobTech Dev som vid insamling av annonser, innan fortsatt bearbetning och publicering filtrerar bort känsliga personuppgifter) och enbart behåller nödvändiga fält (som yrkesnamn och placeringsort) för att kunna lista annonserna och hänvisa tillbaka till originalannonserna.
- Grad av bearbetning avser om datan på något sätt har blivit bearbetad från dess råa form. Inom JobTech Dev filtreras dubletter av jobbannonser bort, samtidigt som annonser

berikas med statistiska nycklar för akutella yrken och kompetenser. Inom RoDL har videodata bearbetats för att anonymisera personer och fordon i bildmaterialet, för att skydda individers integritet och säkerställa att GDPR följs.

- Aktualitet avser hur aktuell datan är. Äldre data kan vara mindre känslig att dela än realtidsdata. Viss typ av data åldras relativt snabbt vilket gör det till en fråga om timmar eller dagar innan de kan delas med mindre affärsrisk. Inom RoDL där konkurrenter samverkar delas äldre data just av denna anledning. Här ligger dock fokus på att skapa träningsset för maskininlärning vilket i aktuellt fall ej kräver realtidsdata. Gällande de jobbannonser som delas inom JobTech Dev är dock aktualitet kritiskt för att datamängden ska vara av värde från ett jobbmatchningsperspektiv.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R12:** Överväga och anpassa datans form för att möjliggöra delning med hänsyn till aktuella integritets-, sekretess- och affärsrisker samtidigt som potential för värdeskapande bibehålls. Granularitet, grad av bearbetning och aktualitet kan från detta perspektiv användas som redskap av en plattformsledare och ekosystemets aktörer för att hitta en lämplig form på datan som alla eller en majoritet av ekosystemets aktörer är bekväma med.

Standarder och format för datadelning inom och utom ekosystemet

Ekosystem kan genom utveckling och val av gemensamma och öppna standarder bidra till att främja interoperabilitet mellan dess aktörer men även mot omgivande eller närliggande domäner. I de studerade fallen har standarder huvudsakligen varit domänspecifika såsom GTFS-formatet som refererar till trafikdata. I fallen med Trafiklab och HSL DevCom kan det förekomma att

dataproducenterna använder ett annat format. Här tar i stället plattformsledarna Samtrafiken och HSL på sig ansvaret och omvandlar datamängderna till de gemensamt valda och föredragna formaten.

På samma sätt har Arbetsförmedlingen valt att anta HR Open-standarden för deras individdata-initiativ för att individers CV ska vara portabla till olika rekryteringsplattformar internationellt. När det inte finns några befintliga standarder som är lämpliga erbjuder ekosystemet en potentiell mötesplats för att samarbeta kring utveckling av nya, som i fallet med Trafiklab som utvecklade en ny standard för biljetter och betalningstransaktioner när ett behov uppstod.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem att rekommenderar vi att:

- **R13:** Anta och främja öppna standarder där de finns för att möjliggöra interoperabilitet mellan tjänster och portabilitet av data inom ekosystemet och utåt. Vid behov bör plattformsledaren också ta en aktiv roll för att underlätta utvecklingen av nya standarder för att uppnå samma mål.

Mätning och uppföljning av ekosystemets hälsa och utveckling

Ett ekosystem är ett känsligt system där aktörerna lever i symbios och är beroende av varandra i olika utsträckning. För att beskriva och följa upp det rådande tillståndet inom ekosystem är det vanligt att referera till en metafor gällande ekosystemets hälsa. Det finns ingen enhällig definition av detta begrepp. Dock brukar tre olika attribut refereras till i sammanhanget vilka avser ekosystemets produktivitet, robusthet och diversitet (Iansiti & Levien, 2004).

Med produktivitet avses ekosystemets förmåga att använda sig och förädla råvaran genom skapande av nya eller förbättringar av befintliga tillämpningar. Utifrån fallet med JobTech Dev kan detta t.ex. avse produktiviteten bland matchningsaktörer i att skapa nya och

förbättrade tjänster för digital matchning och vägledning baserat på den data som delas och tillhandahålls inom ekosystemet.

Med robusthet avses ekosystemets motståndskraft mot och förmåga att återhämta sig från interna och externa störningar. I fallet med JobTech Dev kan detta avse att betydelsefulla aktörer väljer lämna ekosystemet av olika anledningar, alternativt att Arbetsförmedlingen får ändrat uppdrag och därmed ej längre kan bistå med den underliggande plattformen och delningen av datan.

Med diversitet avses ekosystemets förmåga att underhålla och attrahera en diversifierad samling av aktörer från olika områden med varierande användningsfall. Återigen gällande fallet med JobTech Dev kan detta exemplifieras med diversitet avseende aktörer från olika områden, t.ex. matchning, rekrytering, utbildning, fack, arbetsgivarorganisationer och andra myndigheter, men även diversitet gällande tekniska tillämpningar baserat på den delade datan.

Inom JobTech Dev har ovan nämnda kategorier förfinas till sex områden där nyckeltal definieras och följs upp på månatlig basis.

Dessa områden är:

- Synlighet: Hur synligt och uppsökt är ekosystemet för omvärlden den senaste månaden?
- Tillväxt: Hur ser ekosystemets tillväxt ut i förhållande till antalet aktörer över den senaste månaden?
- Samverkan: Hur stor har samverkan med aktörerna varit den senaste månaden? Får vi återkoppling och engagemang från ekosystemet?
- Storlek: Hur stor är ekosystemets momentana population av aktörer? Finns det en kritisk massa? Är det en hållbar mängd utifrån ev. frånfall?
- Kvalitet: Är plattformen med datan och dess API:er tillgängliga och av en kvalité som är hållbar för att användas i produktion av ekosystemets aktörer?
- Diversitet: Vilka möjligheter erbjuds ekosystemets aktörer? Är det diversifierat eller fokuserat till vissa områden?

Det är viktigt att notera att dessa aspekter och områden av ett ekosystems hälsa är begränsat till att reflektera just ekosystemet

välstånd och hållbara utveckling genom användning och samverkan kring den underliggande plattformen (datamängder och tillhörande API:er) och tillhörande resurser (t.ex. relaterad öppen programvara och teknisk dokumentation). För JobTech Dev är de tre aspekterna därav lämpliga för att på kort sikt följa hur den data som tillhandahålls kommer i användning och hur samverkan utvecklas.

Gällande mer långsiktig och övergripande effektmätning av ekosystemet avseende dess vision (för JobTech Dev – bidra till en förbättrad digital matchning och vägledning på den svenska arbetsmarknaden) är ett ekosystems hälsa som koncept och verktyg begränsat. För att undersöka och stärka denna koppling bör kvalitativa undersökningar göras bland ekosystemets aktörer gällande deras upplevda effekt av från att vara en del av ekosystemet och av att nyttja de gemensamma resurser som tillhandahålls i form av ex. data, källkod och kunskap.

En generellt viktig aspekt dock är betydelsen som nyckelaktörerna inom ett ekosystem har på dess hälsa då dessa t.ex. kan ha en avgörande påverkan gällande ekosystemets inriktning, tillhandahåller betydelsefulla datamängder eller resurser för ekosystemets övriga aktörer, alternativt på andra sätt samverkar och hjälper övriga aktörer inom ekosystemet att (för)bli framgångsrika.

Rekommendationer

För myndigheter som har, eller planerar att ta en roll som plattformsledare i ett öppna data-ekosystem rekommenderar vi att:

- **R14:** Utarbeta en plan för kontinuerlig mätning och uppföljning av ekosystemets hälsa som används och visas upp internt och externt för att hjälpa styrning av ekosystemet men också attrahera nya aktörer.

Referenser

- Alorwu, A., Visuri, A., van Berkel, N., & Hosio, S. J. (2022). (Re)using Crowdsourced Health Data: Perceptions of Data Contributors. *IEEE Software*, 39(01), 36-42.
- Alves, C., de Oliveira, J. A. P., & Jansen, S. (2017). Software Ecosystems Governance-A Systematic Literature Review and Research Agenda. In *ICEIS* (3) (pp. 215-226).
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government information quarterly*, 32(4), 399-418.
- Enders, T., Wolff, C., & Satzger, G. (2020). Knowing What to Share: Selective Revealing in Open Data. In *ECIS*.
- Iansiti, M., & Levien, R. (2004). *The keystone advantage: what the new dynamics of business ecosystems mean for strategy, innovation, and sustainability*. Harvard Business Press.
- Immonen, A., Palviainen, M., & Ovaska, E. (2014). Requirements of an open data based business ecosystem. *IEEE access*, 2, 88-103.
- Lindman, J., Kinnari, T., & Rossi, M. (2015). Business roles in the emerging open-data ecosystem. *IEEE Software*, 33(5), 54-59.
- Linåker, J., & Regnell, B. (2020). What to share, when, and where: balancing the objectives and complexities of open source software contributions. *Empirical Software Engineering*, 25(5), 3799-3840.
- Linåker, J., & Runeson, P. (2020, August). Collaboration in Open Government Data Ecosystems: Open Cross-sector Sharing and Co-development of Data and Software. In *International Conference on Electronic Government* (pp. 290-303). Springer, Cham.

Linåker, J., & Runeson, P. (2021). How to enable collaboration in open government data ecosystems: A public platform provider's perspective. *JeDEM-eJournal of eDemocracy and Open Government*, 13(1), 1-30.

Oliveira, M. I. S., Lima, G. D. F. B., & Lóscio, B. F. (2019). Investigations into Data Ecosystems: a systematic mapping study. *Knowledge and Information Systems*, 61(2), 589-630.

Rudmark, D., & Andersson, M. (2022). Feedback Loops in Open Data Ecosystems. *IEEE Software*, 39(01), 43-37.

Runeson, P., Olsson, T., & Linåker, J. (2021). Open Data Ecosystems—An empirical investigation into an emerging industry collaboration concept. *Journal of Systems and Software*, 182, 111088.

Worthington, S., Patel, H., & Charemza, M. (2022). What Else Is New? Open Data Users Need to Know What's Changed. *IEEE Software*, 39(01), 48-54.

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. *Information polity*, 19(1, 2), 17-33.

Zuiderwijk, A., & Janssen, M. (2014, June). The negative effects of open government data—investigating the dark side of open data. In *Proceedings of the 15th Annual International Conference on Digital Government Research* (pp. 147-152).