

# Testa ny teknik för automatisering inom offentlig förvaltning

I2019/03237/DF

# Sammanfattning

Lantmäteriet och Myndigheten för digital förvaltning (DIGG) har i samverkan utfört regeringsuppdraget om att testa ny teknik för automatisering inom offentlig förvaltning. Målen med uppdraget var att utforska hur det offentligas processer och rutiner kan effektiviseras genom automatisering, med bibehållen eller ökad kvalitet, samtidigt som transparensen bakom beslutsfattande bibehålls eller ökar.

Samtidigt med detta uppdrag pågår en accelererad digitalisering driven av en pandemi. Den offentliga förvaltningen har på mycket kort tid genomgått en mycket stor förändring avseende arbetssätt där digitala stöd, arbeten och möten används utanför traditionella kontorsmiljöer och med stora inslag av innovation och ny teknik. Denna stora utveckling har än mer visat på behov att stärka och vidmakthålla förtroende i den digitala världen. Uppdragets syfte har därmed blivit än mer aktuellt.

För uppdraget har särskilt utpekats att nya tekniker och tillväxtteknologier (*emerging technologies*) som artificiell intelligens (AI) och blockkedjeteknik är teknikområden som förutses vara drivande för automatisering. Inom AI finns tillämpningar i form av produkter och tjänster som kan användas för att automatisera och skapa kundnytta genom effektivisering och kvalitetshöjning även för offentlig förvaltning. Detta har bland annat redovisats i DIGG:s delrapport "Främja den offentliga förvaltningens förmåga att använda AI". Även om det också finns mycket att fortsätta utforska och utveckla inom AI-området så är det en vedertagen och mycket sannolik uppfattning att AI kan vara en avgörande faktor som bidrar till avsevärda förbättringar, genom bland annat automatisering i offentlig förvaltning.

Blockkedjetekniken (ibland benämnt *blockchain* eller *distributed ledger technology*) är som ny teknik relativt väl etablerad och känd, tillgänglig och möjlig att tekniskt utveckla, implementera och förvalta i offentlig förvaltning. Det finns dock utmaningar och frågeställningar kring ansvar och ägarskap, förvaltning och utveckling och annat som relaterar till att det handlar om distribuerade systemlösningar. Regelverk, organisationer, administration och verksamheter är vanligtvis styrda utifrån ett centraliserat perspektiv. Detta har inneburit att vi ännu inte sett så många tillämpningar i samhället vid sidan om kryptovalutor eller digitala valutor samt vissa exempel från logistikindustrin.

Det finns flera faktorer som kan verka begränsande för en snabbare digitalisering och automatisering. Att juridiken inte utvecklas i takt med tekniken, att det saknas resurser och kompetens att utveckla och att politiken inte bidrar är några argument som nämns inom myndigheter när digitalisering diskuteras. I detta uppdrag har resultatet tagit sikte på att främst möta ett behov av att skapa förtroende för offentlig förvaltnings användning av ny teknik, och då främst vid automatisering. Ny teknik är av naturen oprövad och med den komplexitet som viss teknik har så kan otydlighet och okunskap skapa misstroende och rädsla. Just AI och svårigheten att förklara hur det fungerar har skapat debatt och diskussion kring transparens, förklarbarhet och etik.

Resultatet av uppdraget är kompetens, insikter, erfarenheter samt ett koncept på en förtroendemodell för automatisering i offentlig förvaltning. Modellen bygger på insyn och förståelse för att säkerställa tillit och förtroende vid automatisering inom den offentliga förvaltningen med hjälp av bland annat AI. Konceptet har testats på en verksamhet i Lantmäteriet och är generiskt för att kunna användas inom hela den offentliga förvaltningen. Modellen förklaras och illustreras med exempel i rapporten och dess bilagor. I denna första version är den framtagen främst för användningen av AI för automatisering. Modellen är dock möjlig att anpassa för även andra digitala stöd för automatisering. Förtroendemodellen bygger på ett antal förtroendemekanismer, eller förtroendeingivande faktorer som kan sammanfattas till:

- transparens till hur ett system för automatiserade åtgärder är uppbyggt genom
- tydlighet i form av en deklARATION av det automatiserade systemets beståndsdelar och dess förmåga att utföra uppgifter, eller kompetens, på ett riktigt, rättssäkert och effektivt sätt och
- en strukturerad, säker och öppen logg över systemets och komponenternas identiteter och versioner, innehåll med koppling till specifika ärenden eller åtgärder som utförts.

Lantmäteriet och DIGG ser att förtroendemodellen för automatiserat beslutsfattande med AI kan vara en viktig del för att bibehålla förtroende och tillit till den svenska offentliga förvaltningen men också till den europeiska/globala digitaliserade offentliga förvaltningen. Lantmäteriet och DIGG rekommenderar regeringen att ge DIGG i uppdrag att utveckla, skala upp och etablera modellen i Sverige och EU.

# Innehållsförteckning

<b>Sammanfattning</b> .....	<b>2</b>
<b>1 Inledning</b> .....	<b>7</b>
1.1 Uppdraget.....	7
1.2 Metod och ansats.....	7
1.3 Avgränsning.....	8
1.4 Bakgrund.....	9
<b>2 Förtroendemekanismer som bygger en förtroendemodell</b> .....	<b>10</b>
2.1 Förtroendet för offentlig förvaltning.....	10
2.2 Förtroende för digitalisering och särskilt om automatisering i offentlig förvaltning.....	12
2.3 Transparens.....	13
2.4 Rättssäkerhet.....	15
2.5 Kunskap.....	16
2.6 Resultat och utfall.....	17
2.7 Etik.....	18
2.8 AI och förtroende.....	19
2.9 AI-ID och version.....	21
2.10 Algoritm och ursprung.....	22
2.11 AI-modell och process, systematik för träning, validering och återföring av prestation, resultat och kunskap.....	22
<b>3 Omvärld</b> .....	<b>23</b>
<b>4 Resultat och rekommendationer</b> .....	<b>24</b>
4.1 Förtroendemodell för automatisering i offentlig förvaltning.....	24
4.1.1 Förtroendemodellen och dess komponenter.....	25
4.1.2 Ett exempel, en fallbeskrivning.....	27
4.1.3 Om förtroendemodellkort.....	28
4.1.4 Versionslogg.....	30
4.2 Resultat och reflektioner.....	30
4.3 Rekommendationer.....	31
<b>5 Bilaga 1 - Litteratur-, käll- och inspirationsförteckning</b> .....	<b>32</b>

<b>6</b>	<b>Bilaga 2 - Mall Förtroendemodellkort</b>	<b>34</b>
6.1	<i>Översiktlig Sammanfattning</i>	34
6.1.1	Tänkt användning	35
6.1.2	Prestanda	35
6.1.3	Begränsningar	35
6.1.4	Processöversikt	35
6.2	<i>Modellinformation</i>	35
6.2.1	Arkitektur	35
6.2.2	Teknisk plattform	35
6.2.3	Träning	35
6.2.4	Begränsningar	35
6.2.5	Implementationsdetaljer	36
6.2.6	Artiklar eller andra resurser för mer information	36
6.2.7	Licens	36
6.2.8	Kontaktperson	36
6.2.9	Avsedd användning	36
6.2.10	Användningsfall utanför tillämpningsområdet	36
6.3	<i>Faktorer</i>	36
6.3.1	Relevanta faktorer	36
6.3.2	Utvärderingsfaktorer	36
6.4	<i>Utvärdering av modellen</i>	36
6.4.1	Resultat	37
6.4.2	Användarresultat	37
6.4.3	Unitära resultat	37
6.4.4	Trösklar och jämförelser	37
6.4.5	Utvärderingsmetod	37
6.5	<i>Data</i>	37
6.5.1	Utvärderingsdata	37
6.5.2	Träningsdata	37
6.6	<i>Övervakning, loggning och profilering</i>	38
6.7	<i>Etiska aspekter</i>	38
6.8	<i>Juridiska aspekter</i>	38
6.8.1	Integritetslagstiftning	38
6.8.2	Sekretess eller annan skyddslagstiftning	38
6.8.3	Omprovning och överklagande	38
6.8.4	Tillsyn	39
6.9	<i>Betänkanden och rekommendationer</i>	39
<b>7</b>	<b>Bilaga 3 – Användningsfall på Lantmäteriet</b>	<b>40</b>
7.1	<i>Förtroendemodellkort - Förändringsdetektering - Processen</i>	40

7.1.1	Översiktlig Sammanfattning .....	40
7.1.2	Processinformation .....	42
7.1.3	Avsedd användning .....	44
7.1.4	Faktorer .....	45
7.1.5	Utvärdering av processen .....	45
7.1.6	Data .....	46
7.1.7	Övervakning, loggning och profilering .....	48
7.1.8	Etiska aspekter .....	48
7.1.9	Juridiska aspekter .....	49
7.1.10	Betänkanden och rekommendationer .....	49
7.2	<i>Förtroendemodellkort - Förändringsdetektering - Detekteringsmodell</i> .....	50
7.2.1	Översiktlig Sammanfattning .....	50
7.2.2	Modellinformation .....	51
7.2.3	Avsedd användning .....	53
7.2.4	Faktorer .....	54
7.2.5	Utvärdering av modellen .....	54
7.2.6	Data .....	57
7.2.7	Etiska aspekter .....	59
7.2.8	Juridiska aspekter .....	60
7.2.9	Betänkanden och rekommendationer .....	60

# 1 Inledning

## 1.1 Uppdraget

Lantmäteriet och Myndigheten för digital förvaltning (DIGG) har fått i uppdrag av regeringen att undersöka om det är möjligt att använda ny teknik som artificiell intelligens (AI) och blockkedjeteknik för att i högre grad automatisera rutiner och processer för ärendehantering och beslutsfattande i offentlig förvaltning<sup>1</sup>.

Målsättningen med uppdraget är att processer och rutiner ska effektiviseras med bibehållen eller ökad kvalitet samtidigt som transparensen bakom beslutsfattande också bibehålls eller ökar genom ett koncept som kan appliceras i delar av Lantmäteriets verksamhet och vidareanvändas i den offentliga förvaltningen.

## 1.2 Metod och ansats

Uppdraget är till sin karaktär mer av ett innovativt och utforskande uppdrag än ett uppdrag om att ta fram evidensbaserade resultat. Uppdraget har utgått ifrån en idé om ett behov att stärka förtroende kring automatisering och har därför utvecklat ett koncept för en teknisk lösning som ska vara möjlig att applicera på delar av Lantmäteriets verksamhet och sedan kunna vidareanvändas inom den offentliga förvaltningen.

Arbetet har handlat om att utveckla och pröva tesen om att skapa en förtroendemodell för AI och sedan prova om resultatet fungerar. Resultatet har till vissa delar blivit väldigt nära idén samtidigt som andra delar har anpassats till den verklighet som råder och de kunskaper och erfarenheter som finns på området.

Uppdraget rör sig på ett område som fortfarande saknar tillämpningar och erfarenheter i tillräcklig omfattning för att verkligen ge möjlighet att komma till ett färdigt resultat. Resultatet är ändå i så pass långt utvecklat skick att det redovisas och rekommenderas för vidareutveckling och användning i offentlig sektor.

Det är en allmän uppfattning att automatisering med ny teknik, bland annat AI, redan är en av digitaliseringens stora möjligheter och att AI är en teknik som kommer att användas i allt högre grad även hos myndigheter. Istället för att

---

<sup>1</sup> Uppdrag I2019/03237/DF om att testa ny teknik vid automatisering inom offentlig förvaltning

använda AI för att skapa lösningar för automatisering har detta uppdrag utgått ifrån att skapa förutsättningar som gör det möjligt för offentlig förvaltning att göra automatiseringen med fortsatt förtroende.

Utvecklingen av resultatet har beaktat gällande nationell rätt och EU-rätt, särskilt Europaparlamentets och rådets förordning 2016/679 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av sådana uppgifter (dataskyddsförordningen). Resultatet vilar på de riksdagsbundna målen för digitaliseringen av offentlig förvaltning och den nationella inriktningen för AI.

Inriktning, syn och uppfattning om automatisering i offentlig förvaltning har varit särskilt intressant att beakta i utvecklingen och ett citat från betänkandet Juridik som stöd för förvaltningens digitalisering (SOU 2018:25) kan inleda en beskrivning av resultatet i detta uppdrag.

*"I takt med utvecklingen av ny teknik och nya arbets sätt skapas också mervärden genom nya möjligheter till uppföljning, styrning, forskning och kunskap. Samtidigt behöver de centrala värden som under lång tid burit upp den svenska förvaltningen fortsatt vara en bas för den digitala förvaltningen. Det är angeläget att även det digitala samhället genomsyras av ett demokratiskt synsätt och att alla känner en grundtrygghet i den digitala samhällsutvecklingen. För att enskildas tillit till den digitala förvaltningen ska kunna upprätthållas är inledningsvis transparens i myndigheternas verksamhet genom möjlighet till insyn fortfarande av grundläggande betydelse. En öppen förvaltning med ordning och reda på uppgifter och informationssamlingar är därtill en nödvändig förutsättning för att den offentliga förvaltningens uppgifter ska kunna vidareutnyttjas på ett sätt som skapar både ekonomiska och samhällsnyttiga värden."*

### **1.3 Avgränsning**

Uppdraget fokuserar primärt på automatisering med stöd av nya tekniker som AI och blockkedjor. Omfattningen för uppdraget har varit för begränsad för att ge utrymme att redovisa kunskaper om exempelvis AI, blockkedjeteknik, transparens och förtroendeskapande mekanismer. Rapporten innehåller ingen redovisning av djupare fakta. Det som redovisas om resultatet är förhållandevis kortfattat. I vissa fall ges en källhänvisning i texten men den käll- eller inspirationshänvisning som finns bilagd ger en fingervisning om vad som tagits i beaktande, inspirerat eller även gett direkt input till resultatet. Den förtroendemodell beskrivs har några komponenter som både är text och teknik. Hela konceptet beskrivs i denna



rapport men den tekniska delen redovisas dock inte med kod eller någon form av systemdokumentation då den ännu inte finns utvecklad.

## 1.4 Bakgrund

I uppdraget framgår det att skälen till regeringsuppdraget är de riksdagsbundna målen för digitalisering. De handlar om att Sverige ska vara bäst i världen på att använda digitaliseringens möjligheter. Men också om en innovativ och samverkande statsförvaltning som är rättssäker, effektiv, har hög kvalitet, service och tillgänglighet och som därigenom bidrar till Sveriges och EU:s utveckling.<sup>2</sup>

Svensk offentlig förvaltning är präglad av transparens och tydlighet för verksamheter och beslut samt av att det finns en hög grad av rättssäkerhet. Det finns en del utmaningar kopplat till användningen av AI. En är att det är svårt att redovisa exakt hur AI kommer fram till slutsatser. Det finns heller inte många AI-tillämpningar, särskilt i offentlig förvaltning, som gett erfarenhet eller visat på om sådana automatiserar med ett förtroende från användare. Automatisering med stöd av AI kan komma att behöva legitimeras i någon mening. Idén var att det skulle kunna gå att ange ett mått på en AI:s tillstånd, kompetens och förmåga och logga den på en blockkedjelösning och därmed ge transparens och förmedla trygghet i systemet och säkerhet till informationen.

---

<sup>2</sup> Prop 2011/12:1, utg.omr.22, bet.2011/12:TUI, rskr.2011/12:87, prop.2009/10:175, bet.2009/10:FiU:38, riskr.2009/10:315

## 2 Förtroendemekanismer som bygger en förtroendemodell

För att skapa ett koncept för förtroende för automatisering, en förtroendemodell, behövs kunskap om vilka typer av mekanismer som bygger förtroende.

Svensk offentlig förvaltning bygger på en lång tids utveckling under förhållandevis stabila, demokratiska och ekonomiska förhållanden. Detta har gett en mycket hög grad av ordning och reda, en generellt sett väl fungerande administration och en demokratisk, rättssäker offentlig förvaltning. Den kan lätt tas för given men det är en rad olika förutsättningar och komponenter som tillit och förtroende vilar på.

För förståelse för förtroendemodellens uppbyggnad ges nedan en beskrivning av några förtroendemekanismer som varit relevanta för uppdragets resultat.

### 2.1 Förtroendet för offentlig förvaltning

Samhällets förtroende för offentlig förvaltning och dess myndigheter är generellt högt i Sverige. Det finns en lång tradition av god administration och offentlig förvaltning av samhällsfunktioner på olika nivåer. Det finns en rad faktorer som spelar roll för tilliten till det offentliga och de redovisas och följs upp på flera sätt och perspektiv.

Ett exempel är en undersökning som årligen genomförs av Kantar Sifo, Anseendeindex myndigheter 2020<sup>3</sup> mäter dels allmänhetens tilltro och anseende av en viss myndighet och dels bedömning av vilken kvalitet och framgång myndigheten tillskrivs. Undersökningen låter ett större antal intervjuade bedöma ett slumpmässigt urval, 2/3 av alla myndigheter i Sverige, utifrån en anseendemodell som mäter både det rationella och emotionella omdömet om myndigheterna ifråga. Undersökningen tittar också närmare på vilka faktorer som driver anseendet för en myndighet.

---

<sup>3</sup> <https://www.kantarsifo.se/rapporter-undersokningar/anseendeindex-myndigheter-2020>

Där återfinns bland annat följande faktorer:

- upprätthåller etik och moral,
- ger en likvärdig behandling åt alla,
- arbetar för jämställdhet mellan män och kvinnor,
- är trovärdiga i media,
- känt och tydligt vad de håller på med,
- använder sociala medier, appar och webben för att synas och bli mer tillgängliga,
- uttrycker sig på ett sätt som är enkelt att förstå,
- levererar tydliga resultat,
- är effektiva, god servicenivå,
- håller vad de lovar,
- har en förtroendeingivande ledning och duktiga medarbetare.

Undersökningen visar att förtroende kan stiga och dala ganska dramatiskt för en myndighet beroende på hur den uppfattas. En viktig del för att skapa förtroende och tillit är öppenhet, kommunikation och information. Det är vedertaget att insyn motverkar exempelvis korrupcion och att det i sin tur leder till ökat förtroende för exempelvis en myndighet. Detta faller under vad som i det följande redovisas under rubriken Transparens.

En annan faktor som spelar stor roll för förtroendet är upplevelsen att myndigheter vet vad de gör, är handlingskraftiga, har god och effektiv service och att de är experter på sitt område. Ett gott aktuellt exempel är det förtroende som Folkhälsomyndigheten och, i synnerhet, några av myndighetens representanter, erhåller under pågående pandemi. Ett exempel på att förtroende kan minska då kunskap eftersätts är när skolor haft obehöriga lärare i för stor omfattning och då skolors förmåga att lära ut ifrågasätts. Denna typ av förtroendemekanism redovisas under Kunskap nedan.

Rättssäkerhet är ett ord som förmedlar ganska starka intryck om myndigheter i medias rubriksättning. Rättsmedvetandet och rättssäkerheten är något som är starkt förankrat i det svenska demokratiska samhällsbygget. Därför är också rättssäkerheten något som behandlas i det följande.

Slutligen finns ett område som har blivit särskilt aktualiserat med digitalisering, automatisering och AI. Det är etik, som egentligen omfattar en rad olika

delområden som exempelvis värderingar, likabehandling och rättvisa, moral och juridik.

I det följande redovisas således fyra områden som särskilt betydelsefulla i konceptet för en förtroendemodell för automatiserade åtgärder i offentlig förvaltning: transparens, rättssäkerhet, kunskap och etik.

## **2.2 Förtroende för digitalisering och särskilt om automatisering i offentlig förvaltning**

Det har fram tills alldeles nyligen inte funnits någon samlad kunskap om vilka myndigheter som använder automatiserade beslutsmetoder och i vilken omfattning så sker. Riksrevisionen har nyligen genomfört en granskning av automatiserat beslutsfattande i statsförvaltningen och i den rapport som lämnades finns en inventering och ett underlag som är värdefullt för både kunskap och fortsatt utveckling<sup>4</sup>. Exempel på beslut som fattas automatiskt idag är vissa om föräldrappening hos Försäkringskassan, om trängselskatt hos Transportstyrelsen och en andel beslut om bl. a. in-teckningar hos Lantmäteriet.

Förutom beslut så används teknik för automatisering av flera funktioner i processer och myndighetsflöden, ibland även för sådant som leder fram till beslut. Det kan vara automatiserade funktioner för att ta emot och diarieföra ärenden eller att fördela information, skicka information och fakturera.

Det saknas också en djupare kunskap om hur digitalisering, automatiserade beslut eller automatisering i övrigt hos myndigheter påverkar förtroendet för myndigheterna. En svensk undersökning från 2019, Medborgarna och automatiserat beslutsfattande<sup>5</sup>, påtalar att det finns förvånansvärt få studier som ägnats åt medborgares perspektiv på automatiserat beslutsfattande i offentlig sektor. Det finns studier av tekniska aspekter och om hur handläggares roller förändras. Men kunskap om hur förtroende och attityder till offentlig sektors digitalisering och automatisering påverkas och utvecklas behöver ökas. Den nämnda studien undersöker om medborgarna är medvetna om automatiserat

---

<sup>4</sup> Automatiserat beslutsfattande i statsförvaltningen – effektivt, men kontroll och uppföljning brister (RiR 2020:22)

<sup>5</sup> Denk, Hedström, Karlsson, SOM-institutet, Göteborgs universitet, 2019,  
[https://som.gu.se/digitalAssets/1734/1734139\\_denk-hedstr--m-karlsson.pdf](https://som.gu.se/digitalAssets/1734/1734139_denk-hedstr--m-karlsson.pdf)

beslutsfattande och hur de förväntar sig hur automatisering påverkar beslutsfattande i offentlig sektor.

## 2.3 Transparens

Transparens, som egentligen betyder genomskinlighet, innefattar många aspekter i form av öppenhet, kommunikation och information. Transparens är både ett stort forsknings- och kunskapsområde. Transparens saknar givna ramverk och det är svårt för såväl privata som offentliga organisationer att hitta rätt nivåer och grader av transparens för att erhålla de effekter man eftersträvar. I uppdraget har en insats av ett litet forskarteam från RISE<sup>6</sup> avseende just transparens belyst vikten och utmaningarna med transparens.

Det är vedertaget att transparens motverkar korruption och att det i sin tur leder till ökat förtroende för exempelvis en myndighet. Men det är inte självklart att transparens leder till bra effekter ifråga om förtroende för en verksamhet.

Transparensen måste visa på att verksamheten är av bra kvalitet för att effekten förtroende ska uppstå hos kunder, allmänhet och medborgare. En dåligt designad transparens kan leda till att minska på anseende och förtroende för offentlig förvaltning.

Undersökningen Medborgarna och automatiserat beslutsfattande, som nämns ovan, visade att det var endast 20 % som hade kännedom om att det fanns automatiserat beslutsfattande i offentlig sektor. Samtidigt ansåg en majoritet av de tillfrågade att beslut blir mer opartiska när datorer beslutar istället för handläggare. Men majoriteten ansåg också att de automatiska besluten inte blir mer tillförlitliga, att de tar mindre hänsyn till människors situation och minskar insynen i beslutsfattandet.

De som hade kännedom om automatiska beslut sedan tidigare var också mer positiva till de automatiska beslutens opartiskhet och tillförlitlighet men ansåg, i samma utsträckning som de utan kännedom om automatiska beslut sedan tidigare, att automatiska beslut är mindre hänsynstagande till människors situation och att det blir minskad insyn i beslutsfattandet.

---

<sup>6</sup> Se bilaga 1

Detta visar på en del av komplexiteten att ha ”rätt” transparens. Ifråga om AI och transparensen till beslutsfattande och andra åtgärder med stöd av AI blir transparensfrågan än mer utmanande. Den så kallade black box-problematiken som AI förknippas med; att AI-teknik är komplicerad och att det vid maskininlärning blir svårt att mer exakt förklara hur AI kommer till slutsatser och lösningar; gör att tydlighet och insyn till beslut och åtgärder blir komplicerat att lösa.

Det pågår både debatt och utveckling som syftar till att möta denna problematik och öka transparensen till AI-åtgärder. Det finns lösningar och pågående utveckling som syftar till att öppna och ge tydlighet till processen om hur AI utvecklas och tränas, till innehållet och annat kring data som används i AI samt till hur AI faktiskt fungerar i olika avseenden. Detta uppdrag startade med en idé om att det skulle gå att redovisa ett mått på kompetens eller förmåga hos en AI. Det har under arbetet inte framkommit något användbart mått på kompetens, förmåga eller rättvisa hos AI som kan användas eller redovisas. Men det har i övrigt inte visats på några tillämpningar av mått för AI och att dessa genom transparens skapar förtroende. Här finns arbete kvar att göra avseende den tekniska utvecklingen.

Det resultat som presenteras av detta uppdrag har designats utifrån ett tänkt framtida behov av att redovisa allt om en AI-förmåga som kan vara av intresse att granska, att följa upp och att ta i beaktande för fortsatt utveckling. Behovet att ha denna insyn kan variera både i detaljer och i ändamål. I den bästa av världar behöver ingen ta del av AI-förmågans egenskaper eller de automatiska beslutens underlag. I vissa fall behövs ingående och djupa revisioner.

Parallellt med detta uppdrag har, som nämnts ovan, en granskning av automatiserat beslutsfattande i statsförvaltningen utförts av Riksrevisionen<sup>7</sup>. Granskningen avsåg att utreda om det automatiserade beslutsfattandet är effektivt utan att rättssäkerheten eftersätts. Som bakgrund utgår granskningen från att digitaliseringen kan bidra till ökad produktivitet, lägre kostnader, snabbare beslut och ökad rättssäkerhet, men om systemet inte fungerar så kan systematiska fel orsaka stora mängder felaktiga beslut som i sin tur kan leda till minskad tillit till

---

<sup>7</sup> <https://www.riksrevisionen.se/rapporter/granskningsrapporter/2020/automatiserat-beslutsfattande-i-statsforvaltningen--effektivt-men-kontroll-och-uppfoljning-brister.html>

den statliga verksamheten. Riksrevisionens granskning omfattar inte några automatiska beslut tagna med AI men i deras enkät till myndigheterna fanns det dock frågor som berör användningen av AI.

Traditionell systemutveckling och kod kräver tekniska kunskaper och ställer också krav på en täckande systemdokumentation i en god offentlighetsstruktur. Att redovisa AI som utför åtgärder som annars kan utföras av människor kan inte ske genom traditionell systemdokumentation. Konsekvenserna av koder och algoritmer kan inte dokumenteras på samma sätt som den logiska och traditionella programmeringen.

Den förtroendemodell som presenteras i detta uppdrag är i någon mån en kombination av dokumentation om process, information, data, system, mjuka värden och transparensdesign. Innehållet kan svara mot en rad olika behov och syften som granskning, bevisföring och information. Förtroendemodellen är ett första steg till att uppfylla en del av det som ses som grundläggande för automatisering i offentlig förvaltning, enligt utredningen Juridik som stöd för förvaltningens digitalisering SOU 2018:25.<sup>8</sup>

*”Säkerställande av möjligheter till insyn är dock enligt vår bedömning, ur rättsligt perspektiv, en av de grundläggande förutsättningarna för att förvaltningen ska kunna hantera rättsliga utmaningar i samband med automationsåtgärder och också kunna visa för allmänheten att de risker som finns tas om hand. På det sättet utgör god offentlighetsstruktur genom säkerställande av insynsmöjligheter en grundläggande förutsättning för att säkerställa rättssäkra automatiserade förfaranden inom förvaltningen.”*

Insyn är ett begrepp som tolkats som att det ingår i det vidare begreppet transparens.

## 2.4 Rättssäkerhet

Digitaliseringen har i allmänhet en positiv effekt avseende rättssäkerhet i den offentliga förvaltningen. Det beror främst på att digitala lösningar ger snabbare, enklare och mer kostnadseffektiva processer, ökad likabehandling, högre grad av service och en kvalitativt bättre handläggning av frågor och ärenden.<sup>9</sup>

---

<sup>8</sup> <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2018/03/sou-201825/>

<sup>9</sup> <https://www.regeringen.se/rattsliga-dokument/statens-offentliga-utredningar/2018/03/sou-201825/>

Samtidigt med att digitaliseringen ger stora positiva effekter så ökar också förväntningarna på offentlig förvaltning. Det är sannolikt att förväntan på och kraven på rättssäkerhet blir högre när processer och beslut blir automatiserade och då det går snabbare. Att det ska finnas möjlighet att överklaga och ompröva automatiserade beslut till en mänsklig bedömning har varit uppe i diskussioner utan att det fastställts i någon regel eller policy för offentlig förvaltning. Om en överprövningsmöjlighet saknas kan tilltron för AI minska, särskilt om det blir negativa beslut. Med detta kan förtroendet för automatisering och digitalisering sjunka.

## 2.5 Kunskap

En viktig faktor för att bygga förtroende och tillit till offentlig förvaltning är att den utförs med god kunskap om de verksamhetsområden respektive myndighet ansvarar för. Det är också viktigt att utvecklingen av offentlig verksamhet sker utifrån denna kunskapsbas och den erfarenhet som finns samlad i offentlig förvaltning.

Automatisering innebär att arbetsuppgifter överförs från människor till maskiner. Traditionellt har dessa uppgifter ofta varit av enkel karaktär. Även när datorer började användas inom tjänstesektorn har automatiseringen omfattat enklare arbetsuppgifter. Den datorkraft och de programmeringsmöjligheter som funnits har inte medgivit att utföra komplexa uppgifter. Det har också saknats datamängder och beräkningskraft.

AI har under de senaste åren utvecklats snabbt och kan redan idag lära sig, hantera och utföra komplexa uppgifter som tidigare bara kunde utföras av människor. Det är framför allt de ökade datamängderna, utvecklingen av algoritmer och en kraftigt ökad beräkningskraft som möjliggjort den här utvecklingen. Det går att överföra kunskap som praktiskt tillämpas och verkställer sådant som bara människor tidigare klarat.

Den stora skillnaden mellan automatisering med AI och automatisering genom annan digitalisering är överföringen av kunskap till systemet. Vid traditionell systemutveckling och programmering överför vi människor kunskap i form av explicita regler, ett enkelt exempel på sådan regel är "ålder >18". För mera komplexa uppgifter, som exempelvis att avgöra om en bild visar en trafikskylt, är det svårt att lösa det med traditionell programmering. För den typ av uppgifter lämpar sig datadrivna metoder som maskininlärning för att överföra kunskaper till



systemet implicit genom att bearbeta data för att skapa reglerna. Det kan exempelvis ske genom att träna systemet på stora mängder bilder på trafikskyltar för att lära sig vad som är karaktäristiskt för dessa. Fördelen med en datadriven ansats är att systemet kan tränas om på nytt data till skillnad mot traditionell programmering där systemet måste programmeras om. En datadriven ansats kan därför öka en komplex lösnings robusthet, anpassningsbarhet och förvaltningsbarhet.

Idén till det koncept som presenteras i rapporten var att det initialt skulle kunna gå att ange en AI:s kunskap och förmåga i någon form av mått. Det skulle kunna vara ett mätetal likt ett betyg, en gradering eller en omfattning av den träningsdata som AI-tillämpningen fått. Det har inte gått att finna någon sådan utvecklad lösning. Det har heller inte gått att nå fram till en sådan egenutvecklad lösning som kan presentera ett mått på kunskap.

Istället har det setts som möjligt att ganska långt redovisa olika delar i det som bygger kunskap och förmåga hos AI att prestera olika åtgärder. Att ge transparens till vad som bygger kunskap kan vara lika värdefullt som att kunna skylta med att man exempelvis har behöriga lärare, jurister som handläggare eller lantmätariingenjörer som lantmätare. En deklaration av en AI och vad den består av kan likt hur ett CV redogör för en människas samlade erfarenheter beskriva vad en AI baserat åtgärder och beslut på.

## **2.6 Resultat och utfall**

Utfall, effekter och konsekvenser i form av exempelvis beslut från en AI i produktion är resultatet från den samlade kompetensen, val och hantering av data och algoritmer samt träningen av AI-modellen. De producerade åtgärder som, förutom beslut, kan vara utförande av processer, förslag och underlag till beslut kommer med träning och utveckling att uppnå en grad av träffsäkerhet eller riktighet. Ett exempel som ofta används är förmågan hos en AI att diagnostisera eller finna cancer utifrån röntgenbilder. Där används ibland procentsatser för att beskriva hur bra, eller dåligt, en AI presterar. Det går också att jämföra med hur väl vi människor presterar i jämförelse.

En utgångspunkt i uppdraget har varit att det kunde finnas ett mått på AI-förmåga eller intelligens som i sin tur kunde borga för att AI gör det bästa möjliga i varje given situation och därmed inger förtroende. Något sådant mått har ju inte hittats

och det är nog fler parametrar än de som idag används för att mäta resultat och utfall från AI-tillämpningar som bör beaktas.

Idag handlar utfall och resultatmätt mycket om teknisksystemtekniska mått som exempelvis *accuracy*, *precision* och *recall* (att måtten anges med de engelska begreppen är för att vi kan tappa en del av betydelsen genom att översätta orden till svenska). Mått för AI-förmåga handlar ännu mycket om leverantörers beskrivningar av de AI som utvecklats. Men utfall och resultat är ändå något som bör vara med i en transparent redovisning av förmågan hos en AI-lösning.

## 2.7 Etik

Det finns inget samlat etiskt ramverk för offentlig sektor. Det närmaste etiska förhållningsregler för offentlig sektor man kan se i författning är möjligen förvaltningslagen. I förvaltningslagen stadgas bland annat om legalitet, objektivitet och proportionalitet, saklighet, opartiskhet, smidighet och enkelhet. Samtidigt finns inte ordet etik eller moral ens nämnt i vare sig proposition eller konstitutionsutskottets betänkande över förslaget till förvaltningslag.

Etiska förhållningsregler för offentlig sektor och myndigheter kan annars hittas i organisationers värderingsgrunder, i styrande politik och i övrigt i juridisk normgivning och tolkning.

Med digitaliseringen, och främst det ökande intresset för AI, har behoven av etiska förhållningsregler och etiska riktlinjer för AI aktualiserats. Inte minst har frågor om opartiskhet, saklighet och risker för att AI matas med data som leder till diskriminerande eller odemokratiska slutsatser skapat ett stort intresse för etik- och moralfrågorna. Det har också lett till starka krav på att hantera just etik för AI.

Det finns redan en mängd policydokument på området, där inte minst Etiska riktlinjer för tillförlitlig AI<sup>10</sup> som EU-kommissionens expertgrupp på hög nivå publicerat, sätter en etisk agenda för offentlig sektor. Dessa riktlinjer och bedömningslistan för tillförlitlig AI som följer med dessa, har beaktats i utvecklingen av förtroendemodellen.

---

<sup>10</sup> <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

När det gäller offentlig förvaltnings automatisering så kan det antas att den utvecklingen görs eller bör göras med beaktande av gällande värdegrunder, exempelvis den gemensamma värdegrunden för statsanställda<sup>11</sup>, och etiska riktlinjer på EU-, nationell eller lokal nivå. Värdegrunden för statsanställda har också beaktats i förtroendemodellen.

Eftersom etiken bör beaktas ifråga om AI och automatisering så blir en slutsats att de som använder AI också bör vara transparenta med hur deras etik ser ut. Det finns inte någon sådan medveten redovisning för traditionell verksamhet så det är något som behövs utvecklas när det gäller exempelvis träning och utveckling av AI-modeller. Detta är också något som beaktats i konceptet som detta uppdrag redovisar. Det är dock med en viss reservation, då det saknas erfarenhet av att beskriva etiska förmågor som överförs från människa till maskin eller som ligger till grund för utveckling av AI.

Etik handlar om en reflektion över moral. Moral är människans uppfattning om vad som är rätt och fel. Det diskuteras om etik för AI ska handla om att människan hanterar AI på ett etiskt godtagbart sätt eller om AI självt ska utrustas med förmåga att bete sig etiskt korrekt. Förmodligen är det både och. Åtminstone i det längre perspektivet om man ser att AI utvecklas till så kallade ”*artificiell generell intelligens*”<sup>12</sup> eller autonoma AI-system.

## 2.8 AI och förtroende

Tanken på att digitaliseringen, främst utifrån automatiseringen med AI, riskerar att leda till minskat förtroende har sin utgångspunkt i att dels AI har en problematik kring vad man kallar ”*black box*”, där det inte riktigt går att förklara exakt hur en algoritm kommer fram till en slutsats och att det kan skapa misstänksamhet. Men det handlar också om att det finns utmaningar att få AI att agera på det sätt som man normalt ställer som krav på människor i offentlig förvaltning att förhålla sig till. Det är till exempel krav på rättssäkerhet, hur tjänstemän agerar med iakttagande av etik och moral, likabehandling, kunskap och annat.

---

<sup>11</sup> <https://www.regeringen.se/informationsmaterial/2013/10/s2013.011/>

<sup>12</sup> Wikipedia: Artificiell generell intelligens (*artificial general intelligence*, AGI) är en hypotetisk AI som uppvisar människolik *intelligens*, det vill säga, som klarar av att utföra vilken intellektuell uppgift som helst som en människa kan utföra

Eftersom AI, per definition, är teknik som efterliknar mänskligt beteende, så kommer också utvecklingen mot alltmer självständig förmåga hos AI att göra maskinen till en mer människolik resurs. Det kan i sin tur innebära att kraven på resursen blir mer lik de krav som ställs på människor. Dessa krav kan till och med bli starkare på maskinen än på människor utifrån förtroendeperspektivet.

Det är förmodligen långt kvar, enligt de flesta källor inom vetenskapen på området AI, till att vi har fullständigt autonoma eller generella AI-system som utvecklar sig själva och tar beslut helt fristående från mänsklig styrning och kontroll. Idag är den tillämpade AI-tekniken främst en teknisk komponent som utvecklas och programmeras för att hantera stora mängder data genom beräkningar så att AI-modeller kan hantera komplexa uppgifter. Träningen av modellerna är oftast helt separerat från produktionen. Det är människor som hanterar processen från lärandet till att exempelvis de färdiga AI-modellerna sätts in i ett produktionssystem.

Denna utveckling ställer ändå helt andra krav än den tidigare IT-utvecklingen eftersom den handlar om att processa och bedöma data. Tekniken möjliggör redan automatisering utan AI. Men att automatisera hela flöden där både bedömningar och beslut med rättsverkan för människor sker utan mänsklig kontroll och styrning, det vill säga automation, ligger längre fram i tiden. Vi har ännu inte den mognaden och det skulle kräva rigida och omfattande kontrollmekanismer, förmodligen reglerat i framtida regelverk, för att säkerställa kvalitet, säkerhet och etik.

Arbetsförmedlingens ”beslutsrobot” hade ett systemfel som orsakade att ett större antal beslut inte skickades ut och att det kunde finnas ett antal felbeslut bland dessa. Besluten rörde aktivitetsstöd till arbetslösa och nyhetsbevakningen innebar negativ publicitet för Arbetsförmedlingen i medier.

Utan att det någonstans framgått att Arbetsförmedlingens automatisering gjorts med AI så följde senare ett debattinlägg som nämner detta som exempel. Debattinlägget gick under rubriken ”Hur ska konsumenter få rätt mot ett AI-system?”<sup>13</sup>. Inlägget, som kan sammanfattas med att vinster i form av förenklingar i samhället kan gå förlorade om det inte finns transparens och tillit, visar på

---

<sup>13</sup> <https://www.nyteknik.se/opinion/hur-ska-konsumenter-fa-ratt-mot-ett-ai-system-6995773>

betydelsen av att ny teknik redan från början präglas av öppenhet och tydlighet – transparens – för att inge förtroende samt att brister kan leda till förtroendekris.

De justeringar som vidtogs av Arbetsförmedlingen, förutom att programmera om, var att tydligare upplysa om att det var ett automatiserat beslutsförfarande och att det finns ett kontrollsystem.

## 2.9 AI-ID och version

Rättssäkerhet kräver att det är möjligt att identifiera upphov och ansvar för bland annat beslut. Det är speciellt viktigt vid myndighetsutövning eftersom den ska vara möjlig att genomlysas, förklaras och att den är kopplad till ett ansvar för konsekvenserna. Detta är grundläggande för offentlig förvaltning och rättssäkerhet. Att automatiserade beslut kan kopplas till någon yttersta beslutsfattare och ansvarig för besluten är en viktig princip. Det kan vara lika viktigt att veta ”vem” som bidragit med underlag till besluten. Frågan om det ska gå att identifiera en AI i en beslutsprocess kan därför vara av intresse när det handlar om förtroende för automatisering.

En AI som utför handläggning och beslutsfattande, eller kundservice, utvecklas ständigt med ny och återförd kunskap och träning (jämför med kompetensutveckling för en människa). Det innebär att denna utveckling också leder till uppgraderade AI-modeller. Såväl AI-modell som version kan behöva anges för god spårbarhet och revision. Sedan om det är att betrakta som en ID eller inte kommer säkert diskuteras vidare. Innan AI fungerar helt autonomt kanske det inte ska göras alltför nära liknelser med människor. Maskiner har nummer som visar vilken exakt instans det är fråga om. Även komponenter måste vara spårbara och detta bör gälla för AI-komponenter i handläggningsstöd och liknande som kan finnas på myndigheter. Att myndigheter har en god dokumentation av de IT-stöd som används bör redan idag vara gällande. Den förtroendemodell som föreslås i denna rapport tar denna dokumentation lite längre och föreslår att en löpande logg av version och identifiering av algoritmer och modeller för AI sker.

Det skulle också vara möjligt med ett nationellt öppet register med alla AI-ID som underlag för hur AI och automatiskt beslutsfattande används inom offentlig förvaltning. Det finns exempel på sådana register från Helsingfors stad.<sup>14</sup> Men ett

---

<sup>14</sup> <https://ai.hel.fi/sv/ai-registret/>

sådant register behöver utredas vidare då det kan komma att innehålla väldigt mycket som är poänglöst att föra register över. Det kan bli en övertransparens som omfattar både system som inte ska redovisas av säkerhetsskäl och system som uppenbart inte har någon koppling till rättsverkande åtgärder.

## **2.10 Algoritm och ursprung**

AI-algoritmer kan vara relativt generella. Förutom att utveckla egna algoritmer finns det en marknad med mer eller mindre fritt tillgängliga algoritmer som kan vidareutvecklas och tränas till AI-modeller som är specifika för de ändamål man vill använda dem till. När algoritmer utvecklas och tränas för specifika ändamål, med ändamålsspecifika data och stöd av AI-tränare skapas mer verksamhets- och organisationsspecifika modeller. Det är viktigt att kunna redogöra för vilken eller vilka algoritmer som används och vilket ursprung de har.

## **2.11 AI-modell och process, systematik för träning, validering och återföring av prestation, resultat och kunskap**

När en AI-modell utvecklas och tränas med data från start och fortsatt under tid, med återföring av resultat och nya data och när den valideras och utvecklas ytterligare, så grundläggs förutsättningarna för de åtgärder modellen hanterar och utför. Vid varje uppdatering skapas en ny version av modellen som baseras på ett unikt tillstånd, det vill säga de omständigheter som rådde vid den tidpunkten.

Genom att beskriva dessa processer och hanteringen av data samt den träning och återföring av resultat som sker, ger man en stor förståelse för hur systemet fungerar. Detta är en viktig del i transparensen och det skapar förtroende.

### 3 Omvärld

Omfattande diskussioner och utveckling av såväl förtroendeskapande mekanismer kring AI som etik- och moralfrågor för AI pågår i hela världen. Det har från flera håll, under uppdragstiden, presenterats lösningar och idéer som möter delar av vad denna rapport behandlar.

Något som inspirerat till delar av resultatet i detta uppdrag är något som benämns *Model Cards for Model Reporting*<sup>15</sup> och som sedan utvecklats av Google<sup>16</sup>. Att dokumentera och göra ett AI-system transparent inom myndigheters verksamheter genom en form av innehållsdeklaration liknar Model Cards, men syfte och föremål för redovisningen är annat för den offentliga förvaltningen än den privata och kommersiella marknaden. I förtroendemodellen har innehållsdeklarationen utvecklats utifrån offentlig förvaltnings särskilda förutsättningar.

Flera initiativ att sätta upp etiska och moraliska riktlinjer och policys för AI-utveckling och tillämpning har också gjorts. Förutom EU-kommissionens riktlinjer och andra etiska förhållningsregler för offentlig sektor finns också motsvarande på den privata och kommersiella sektorn. IBM har exempelvis givit ut *Everyday Ethics for Artificial Intelligence* som haft positivt genomslag.

Även andra förtroendegrundande områden vid sidan om etik och moral utvecklas till ramverk och förhållningsregler. Ett aktuellt arbete som pågår och som definitivt kommer att möta en eventuellt fortsatt utveckling av det här presenterade förtroendekonceptet är EU-kommissionens strategi och utveckling för att skapa en europeisk AI-inriktning som bland annat omfattar ett ekosystem för förtroende för AI<sup>17</sup>. Detta förtroendeekosystem som utvecklas av DG CNECT ska enligt uppgift presenteras under inledningen av 2021.

---

<sup>15</sup> <https://arxiv.org/abs/1810.03993>

<sup>16</sup> <https://modelcards.withgoogle.com/about>

<sup>17</sup> WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust, European Commission, 2020-02-19

## 4 Resultat och rekommendationer

### 4.1 Förtroendemodell för automatisering i offentlig förvaltning

Automatisering av uppgifter och processer i offentlig verksamhet kan ses som en självklarhet som minskar kostnader och därmed avgifter och anslag.

Automatisering av mer kvalificerade uppgifter, exempelvis beslut, finns redan införda.

Riksrevisionen har nyligen utfört en granskning av automatiserade beslut i statliga myndigheter och redogjort för en del slutsatser<sup>18</sup>. I rapporten framgår av en inventering att tre myndigheter använder AI för automatisering av delar av beslutsprocesser även om inte själva beslutet tas av en AI. En slutsats Riksrevisionen gör är att det visserligen finns etablerade strukturer för styrning av de automatiserade beslutsprocesserna, men att kunskapsnivån om automatiserat beslutsfattande behöver öka. Det konstateras ett behov av att stödja myndigheterna med kunskapsunderlag och stöd för att ge bättre förutsättningar för automatiseringen. Det ges exempel om att kunskapsunderlaget skulle kunna bestå av modeller för hur myndigheter kan strukturera och styra sina automatiserade processer.

Den här presenterade förtroendemodellen kan uppfylla en del av detta behov.

Automatiserade åtgärder, inklusive beslut, med positivt utfall för kunder och sökande, kommer inte att bestridas i någon större omfattning och de kommer därmed heller inte att verka negativt avseende förtroende för offentlig förvaltnings automatisering. Men när utvecklingen kommit så långt att mer kvalificerade åtgärder vidtas, även inom ramen för ärendeberedningen, så ökar risken att negativa konsekvenser också leder till minskat förtroende.

Det resultat som redovisas i denna rapport avser att skapa en modell för att stärka förtroendet för automatiserade åtgärder med AI och tekniska system som mer självständigt kan efterlikna mänskligt beteende. Förtroendemodellen bygger på att så långt som möjligt redovisa och deklarerar de olika komponenter som AI-

---

<sup>18</sup> Automatiserat beslutsfattande i statsförvaltningen – effektivt, men kontroll och uppföljning brister (RiR 2020:22)



förmågan vilar på och som gör det mer förståeligt och transparent hur en AI-automatiserad funktion kommer fram till en åtgärd.

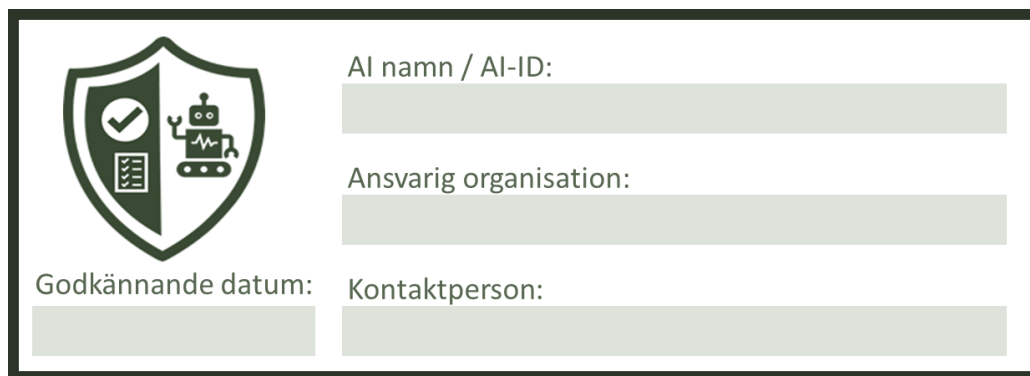
I det följande redovisas en förtroendemodell för automatisering i offentlig förvaltning, dels i en kortare beskrivande text av komponenterna i modellen och dels genom att visa exempel på hur modellen och dess innehåll ser ut. Det redovisas även exempel från en av Lantmäteriets utforskande AI-satsningar där modellen provats.

#### 4.1.1 Förtroendemodellen och dess komponenter

Konceptet med en förtroendemodell för automatisering i offentlig förvaltning består följande komponenter:

##### 4.1.1.1 Nivå 1 - förtroendemärkning

En symbol eller märkning tillsammans med en kort beskrivning av att det finns en automatiserad åtgärd i den process som en kund eller sökande möter hos myndigheten. Märkningen ska kunna innebära en igenkänning att myndigheten har en helt eller delvis automatiserad process med stöd av exempelvis AI och att den är utformad i enlighet med förtroendemodellen för offentlig förvaltnings automatisering. Denna nivå ska normalt sett vara tillräcklig för de flesta att känna förtroende.




	AI namn / AI-ID:	<input type="text"/>
	Ansvarig organisation:	<input type="text"/>
	Godkännande datum:	<input type="text"/>
	Kontaktperson:	<input type="text"/>

Bild. En illustration av hur en förtroendemärkning kan se ut.

##### 4.1.1.2 Nivå 2 - Övergripande informationsvy

En övergripande bild som beskriver hur den automatiserade tjänsten är uppbyggd och vilken information som finns att tillgå om man är intresserad att fördjupa sig i information om den automatiserade tjänsten och egenskaperna i funktionen hos

exempelvis en AI. En enklare schematisk bild över de olika komponenterna och vilket innehåll som finns att tillgå, oavsett vem som är intresserad. Den kan finnas länkad från märkningen eller motsvarande.

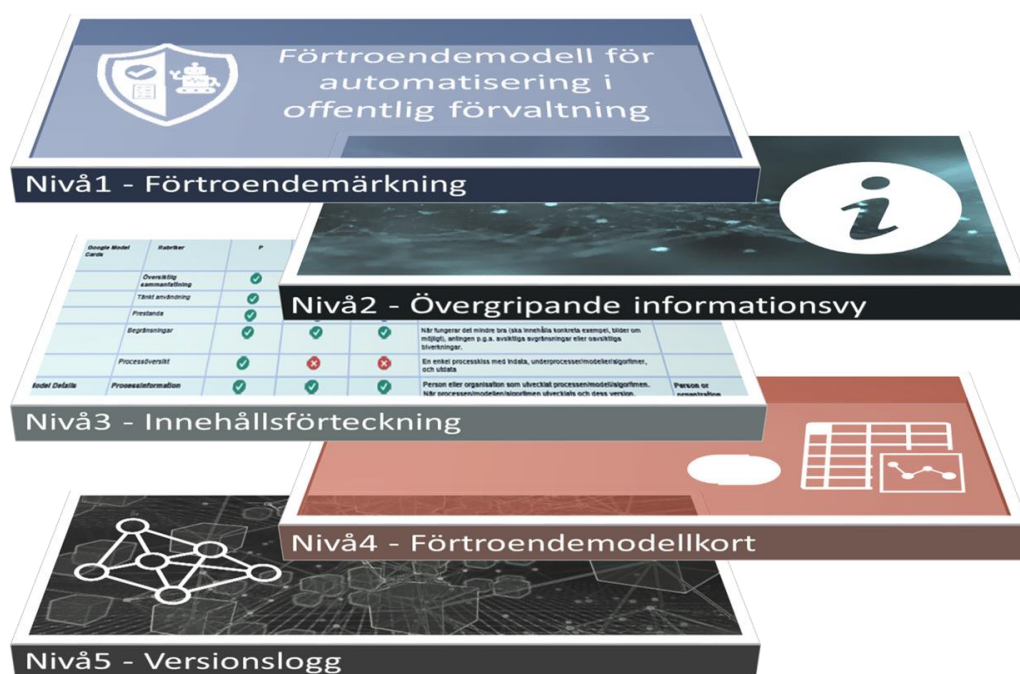


Bild. Övergripande informationsvy av förtroendemodellen

#### 4.1.1.3 Nivå 3 – innehållsförteckning, detaljerad information

En bred och något djupare beskrivning av det innehåll som kan studeras i detalj på nästa nivå för att bilda sig en uppfattning om vilket område som är mest relevant för det intresse någon kan ha.

#### 4.1.1.4 Nivå 4 – förtroendemodellkort

Detaljerade deklARATION och beskrivningar, en innehållsdeklARATION med detaljer, av de olika komponenterna i innehållsförteckningen i nivå 3.

I denna nivå finns troligen information så utförlig att den kräver ingående kunskaper och mer analys för det syfte den granskas. Exempelvis kan denna nivå vara eftertraktad vid revision eller som bevis för en rättslig bedömning.

#### 4.1.1.5 Nivå 5 - versionslogg

En säker och transparent logg, i detta fall med en blockkedjelösning, av de viktigaste delarna i det automatiserade systemet. Det kan vara referenser, exempelvis ärendenummer, signaturer (som krävs för att registrera på loggen),

data som använts i den automatiserade åtgärden, version av AI och de förtroendemodellkort som gällt samt ID för en AI eller annat system.

#### 4.1.2 Ett exempel, en fallbeskrivning

Detta exempel beskriver en möjlig ärendeprocess och hur förtroendemodellen kan upplevas eller ge tillit till det automatiserade åtgärderna.

X har köpt en fastighet med ett bostadshus och en stor byggbar tomt där det tidigare stått ett äldre hus. Köpet har gjorts helt digitalt, genom en digital fastighetsmäklare. X ansöker genast om lagfart och en inteckning. X noterar att ansökningstjänsten delvis kommer att behandlas automatiskt och känner en trygghet med detta då X ser förtroendemärkningen som är typisk för svenska myndigheters digitala automatiserade tjänster.

*Förtroendemärkningen har en länk till en övergripande bild av vad förtroendemodellen innehåller. Utifrån den kan man gå vidare till både den visuella redovisningen av loggen av alla beslut och kopplade informationsslag på en blockkedja och till de beskrivningar som finns i förtroendemodellen och de förtroendemodellkort som har all relevant information.*

Lantmäteriet handlägger ärendet. Beslutet om lagfart tas av en handläggare och beslutet om inteckningen fattas automatiskt genom en AI i handläggningsstödet. X får ett digitalt beslut där det framgår att inteckningen är automatiskt beviljad samt en länk till Lantmäteriets förtroendemodell för automatisering i offentlig sektor, men X behöver inte titta på detta då X är helt trygg med beslutet.

*Beredningen och besluten i Lantmäteriets handläggningsstöd loggas alltid och de delar som innebär automatiska åtgärder genererar loggar till blockkedjan som knyts till relevant version av förtroendemodellkort med mera och den vid beslutet gällande versionen av AI-modellen som utfört åtgärden.*

Då lagfarten är beviljad betalar banken ut lånet som X tagit med fastigheten som pant. När handläggaren gör en efterkontroll ser hen, i fastighetsregistret, att det egentligen är två fastigheter. Men registerkartan och mäklarens drönarvideo ser ut att visa en fastighet med två hus, ett uppenbart med ett stort slitet tak och ett modernt funkishus. Registerkartan är dagsaktuell och visar en stor fastighet. Även registerkartan har en märkning att den delvis uppdateras med automatiska tjänster och AI. Med stöd av Lantmäteriet och besiktning på plats uppnås visshet om att kartan justerats utifrån flygbilder som tolkats av AI att visa en fastighet med två byggnader trots att det var två fastigheter där den ena hade ett fallfärdigt ruckel.

Uppgifter från taxering och annat som tagits in i bedömningen ledde till en felaktig justering i kartan. AI-modellen har sedan dess uppgraderats så att denna typ av misstag inte kan ske.

*Genom att ta del av informationen i deklARATIONEN av AI-modellen fås underlag för den fortsatta hanteringen. Fullständig tillgång till träningsdata och den data som bedömts automatiskt ges till utredare, eventuell domstol och andra som behöver insyn. Lantmäteriet kan återskapa exakt den situation som gav upphov till felet och kan därmed justera systemet och ha en bra grund för rättning och eventuell kompensation. Full rättssäkerhet uppnås.*

#### 4.1.3 Om förtroendemodellkort

En förutsättning för förtroende är att system genomför åtgärder med förväntad kvalitet. För att få det behöver arbetet med utvecklingen av system ske systematiskt för varje del, från design till användning och utvärdering. Genom att dokumentera de olika delarna från framtagande ända till enskilda beslut så kan kvalitet visas och bedömas.

I förtroendemodellen som föreslås finns en viktig komponent, förtroendemodellkort. Förtroendemodellkort är inspirerat av rapporten ”*Model Cards for Model Reporting*”. Förtroendemodellkortet täcker dock in hela system, inom offentlig förvaltning, och inte enbart AI/ML-delar. Detta för att AI/ML är oftast bara en del i att lösa en uppgift.

Förtroendemodellkortet ger ett stöd i vad som bör dokumenteras. Avsikten är att förtroendemodellkort har en bredd och ett innehåll så att det uppfyller det som inger förtroende för tjänsterna som utförs. I den version som ligger i denna rapport finns fortfarande saker utvecklas. Ytterligare tillämpning och utveckling rekommenderas i denna del. Exempelvis har redan konstaterats att en viktig redovisning i ett förtroendemodellkort är energiförbrukning och hållbarhetsaspekter med en AI-förmåga som används.

AI- och ML-modeller belyses särskilt eftersom de kan lösa avancerade uppgifter samtidigt som kvalitet påverkas av hela framtagandeprocessen och användningen bör ske inom de begränsade förutsättningar som systemet har.

Förtroendemodellkortet kan bland annat användas för att:

- få förståelse för enskilda beslut

- kunskap om kvalitet och förutsättningar för system
- beslutsunderlag för införande av system
- utveckling, för att systematiskt få med relevanta aspekter
- granskning av resultat och egenskaper

För att kunna ta till sig informationen så finns det förtroendemodellkort för övergripande beskrivning av systemet, samt detaljerade beskrivningar av systemdelar och detaljerade beskrivningar av ML-delar. Även de detaljerade beskrivningarna innehåller övergripande beskrivningar som en introduktion. Även på modellkorts nivå beskrivs dessa delar och på så vis separeras process, modell och algoritm i separata förtroendemodellkort. En universell mall med tänkt innehåll för respektive del kan ses i *Mall förtroendemodellkort*, bilaga 2. Som komplement och illustration av mallen finns även ett hypotetiskt förtroendemodellkort framtaget för Lantmäterisystemet förändringsdetektering, process (bilaga3), modell (bilaga3). Nedan beskrivs och sammanfattas innehållet av ett förtroendemodellkort för ML.

Innehåll i Förtroendemodellkort för ML:

- **Översiktlig sammanfattning**  
*Tänkt användning, prestanda och begränsningar*
- **Information**  
*Beskrivning av process/modell/algoritm, träning, systeminformation, och begränsningar*
- **Avsedd användning**  
*Tänkta användningsfall och även möjliga användningsfall*
- **Faktorer**  
*Faktorer i data som påverkar resultat*
- **Utvärdering**  
*Beskrivning av utvärderingar och resultat*
- **Data**  
*Beskrivning av utvärderings och träningsdata*
- **Övervakning, loggning och profilering**  
*Vilken form av övervakning finns*
- **Etiska aspekter**  
*Beskrivning av etiska ställningstaganden*

- **Juridiska aspekter**

*GDPR, sekretess och annan skyddslagstiftning, tillsyn samt överklagande*

- Betänkanden och rekommendationer

#### 4.1.4 Versionslogg

Eftersom åtgärder, exempelvis beslut i enskilda ärenden, utförs av en specifik version av den automatiserade lösningen som förklaras av en viss version av förtroendemodellkort är det viktigt att knyta versionerna till varandra och den åtgärd som utförts. Detta sker i en logg som förs löpande av förtroendemodellen.

Loggen utgörs, i det föreslagna konceptet, av en blockkedjekomponent då teknikens egenskaper ger en säker och hållbar registrering av data samtidigt med transparens och effektivitet. Det finns flera typer av blockkedjeteknik och designen av en versionslogg av den typ som beskrivs här bör inledningsvis vara av den typ som benämns "privat" och inte "publik". Med det menas att det är en begränsad krets som kan uppdatera informationen på blockkedjan. Däremot kan den göras publikt tillgänglig för att läsa, men inte uppdatera. Därmed kan versionsloggen ha en hybridkaraktär mellan att vara privat och publik.

Man kan välja vilka uppgifter som ska registreras på loggen. Det är dock viktigt att man knyter samman de mest relevanta uppgifterna om AI-förmågan som utför automatiserade åtgärder eftersom det ger en form av verifiering, eller bevis, om det som utförts.

Eftersom det är möjligt att koppla flera olika typer av data och information om den automatiserade tjänsten så kan det uppstå fler positiva effekter med denna logg än att se den som ett led i transparensen. Versionsloggen kan också fungera som ett verktyg att återskapa den situation som förelåg vid en tidpunkt, om det uppstår ett behov av att göra så.

Versionsloggen har inte utvecklats tekniskt. I teorin så ses det som möjligt att ha en mer automatiserad registrering av de uppgifter som ska loggas, från de systemkomponenter som finns i exempelvis hanteringen av ett myndighetsärende.

## 4.2 Resultat och reflektioner

Resultatet av uppdraget är kompetens, insikter, erfarenheter samt ett koncept på en förtroendemodell för automatisering i offentlig förvaltning enligt tidigare beskrivning. Modellen bygger på insyn och förståelse för att säkerställa tillit och

förtroende vid automatisering inom den offentliga förvaltningen med hjälp av bland annat AI. Konceptet har testats och validerats på en verksamhet inom Lantmäteriet med goda resultat. Lantmäteriet och DIGG anser att modellen är generiskt att kunna börja användas inom hela den offentliga förvaltningen och fortsätta att utvecklas vid olika praktiska tillämpningar.

En viktig del av resultatet är också själva lärandet och utforskandet i ett sådant här uppdrag. Uppdraget rör sig inom outforskat område, där till och med techbolag som Google forskar, där det saknas tillämpningar och erfarenheter i tillräcklig omfattning för att verkligen ge möjlighet att komma till ett färdigt resultat.

### **4.3 Rekommendationer**

Lantmäteriet och DIGG ser att förtroendemodellen för automatiserat beslutsfattande inom offentlig förvaltning kan vara en viktig del för att bibehålla förtroende och tillit till den svenska offentliga förvaltningen. Förtroende och tillit kan ses som en förutsättning för att exempelvis lyckas med effektivisering som är ett viktigt område i de riksdagsbundna målen för digitalisering.

Förtroendemodellen kan också vara ett svenskt bidrag till att skapa god ordning och reda i den europeiska/globala digitaliserade offentliga förvaltningen.

Lantmäteriet och DIGG rekommenderar att regeringen ger DIGG ett uppdrag att utveckla, skala upp och etablera modellen i Sverige och EU. Uppdraget kan också ske i samverkan.

# 5 Bilaga 1 - Litteratur-, käll- och inspirationsförteckning

Följande förteckning avser kontakter och möten, källor, litteratur och artikelstudier samt inspirationskällor.

## **Personer i uppdragets arbetsgrupp**

- Anders Lagerqvist, Lantmäteriet
- Anders Enmark, Lantmäteriet
- Johan Stang, Lantmäteriet
- Anders Ekholm, Lantmäteriet
- Jan Dalheimer, Lantmäteriet
- Patrick Eckemo, Myndigheten för digital förvaltning (DIGG)
- Mats Snäll, Lantmäteriet (t o m 6/12 2020), DIGG

## **Forskarteam från RISE som medverkat i uppdraget:**

- Ulrik Franke, Senior Researcher, RISE
- Alexander Rad, Post doc, RISE
- Jakob Dexe, Industrial PhD-student, RISE och KTH

## **Chromeaway, blockkedjeteknikutvecklare**

- Jörgen Modin
- Tobias Rådeskog

## **Personer som i olika sammanhang bidragit till resultat och kunskap**

- Daniel Akenine, nationell teknikchef, Microsoft
- Therese Svensson, Data Science Consultant, IBM Svenska AB
- Mikael Haglund, Chief Technology Officer, IBM Svenska AB
- Tim Paydos, Global General Manager, Government Industry, IBM
- Charley Rich, Sr Director Analyst, Gartner Research Engagement Services
- Pieter den Hamer, Sr Director Analyst, Gartner Research Engagement Services



- Mark Oost, Global CTO Analytics & AI Services | Netherlands

### **Riksrevisionen**

- Linda Talme
- Caroline Zakrisson

### **Delar av DIGG:s AI referensgrupp för offentlig förvaltning:**

- Christian Guttman, Dr, Global Head of AI & Data Science, TietoEVERY och NAI
- Fredrik Heintz, Dr, Associate Professor of Computer Science, Linköpings universitet
- Kristinn R. Thórisson, Ph.D. Professor, Dept. Comp. Sci. Reykjavik University. Managing Director, Icelandic Institute for Intelligent Machines
- Stefan Larsson, samhällsvetenskaplig AI-forskare, jurist och docent i teknik och social förändring, LTH, Lunds universitet
- Jeanette Nilsson, AI Ecosystem Driver, AI Agendan, RISE
- Göran Lindsjö, Senior AI Advisor
- Cecilia Magnusson Sjöberg, jur. dr och professor i rättsinformatik vid Stockholms universitet
- Ingrid af Sandeberg, senior projektledare, Arthur D. Little, styrelseledamot i Stockholm AI

### **KOMET, kommittén för teknologisk innovation & etik**

- Jon Simonsson, ordförande, KOMET
- Charlotte Hall, utredningssekreterare, KOMET

# 6 Bilaga 2 - Mall

## Förtroendemodellkort

### Mall Förtroendemodellkort

- Mallen används i samband med framtagande av Modellkort för automatiserade åtgärder i offentlig förvaltning
- Ett modellkort skapas för respektive lager (Process, Modell samt Algoritm)
- Text inom <> är stödtext för respektive rubrik
- Exempel på Modellkort (modell) finns i bilaga 3

### 6.1 Översiktlig Sammanfattning

<En sammanfattning av innehållet i övriga kortet. Ska gå att förstå av en "okunnig".>

Typ	L
In	
Ut	
Ingår i	
Arkitektur	
Dokument Version	
Algoritm Version	

### 6.1.1 Tänkt användning

<Varför har det utvecklats? Vad har målen varit? Användningsområden?>

### 6.1.2 Prestanda

Hur väl fungerar det? (beskrivet i enkla termer)>

### 6.1.3 Begränsningar

<När fungerar det mindre bra (ska innehålla konkreta exempel, bilder om möjligt), antingen p.g.a. avsiktliga avgränsningar eller oavsiktliga biverkningar.>

### 6.1.4 Processöversikt

Denna rubrik är endast tillämpad för beskrivning av Processen

<En enkel processkiss med indata, underprocesser/modeller/algorithm, och utdata>

## 6.2 Modellinformation

<Person eller organisation som utvecklat processen/modell/algorithmen.

När processen/modellen/algorithmen utvecklats och dess version.>

### 6.2.1 Arkitektur

Gäller enbart för modeller & algoritmer.

<Grundläggande om modellens/algorithmens arkitektur. Vid en välkänd modell/algorithm kan det räcka att ange dess namn och konfiguration samt länk till artikel med mer information, annars behövs även en kort sammanfattning. **M:** Ska ge personer med kunskap om AI förståelse för begränsningar och betänkan den inbyggda i själva modellarkitekturen.>

### 6.2.2 Teknisk plattform

Gäller enbart för processer.

<Beskriv i ord eller bild vald teknisk plattform, ta särskilt hänsen de till och uppmärksamma ev. molnlösningar>

### 6.2.3 Träning

<Vilka parametrar som använts vid träning, t.ex. antal epoker, batchstorlek, loss-funktion, learning rate, etc.>

### 6.2.4 Begränsningar

<Mer ingående information om begränsningar och avgränsningar.>

### 6.2.5 Implementationsdetaljer

<T.ex. användning av ramverk, programvara, m.m.>

### 6.2.6 Artiklar eller andra resurser för mer information

<Länkar till externa resurser, t.ex. artikel där processen/modellen/algoritmen introduceras, eller andra som haft liknande användning eller som av annan anledning är relevant.>

### 6.2.7 Licens

<Hur är processen/modellen/algoritmen licenserad? Vad krävs för att få använda den?>

### 6.2.8 Kontaktperson

<Vem ska man kontakta för mer information>

### 6.2.9 Avsedd användning

<Primära avsedda användningar och användare.>

### 6.2.10 Användningsfall utanför tillämpningsområdet

<Alternativa tänkbara eller faktiska användningsfall, eventuellt med specifika betänkanden eller kommentarer.>

## 6.3 Faktorer

### 6.3.1 Relevanta faktorer

<De faktorer som kan påverka processen/modellen/algoritmen. Det kan handla om 3 huvudsakliga typer av faktorer: Olika **grupper** av objekt, t.ex. ålder och kön i en modell för ansiktsgenkänning, byggnadsändamål och byggnadsstorlek i en modell för att hitta byggnader eller olika längd på meningar i en NLP-modell. Påverkan av olika **instrument**, t.ex. kameramodell och bearbetningsmetoder.

**Miljön**, t.ex. belysning, plats, m.m.>

### 6.3.2 Utvärderingsfaktorer

<De faktorer (av dem som nämns i 4.1.) som har utvärderats. Varför har de utvärderats eller inte (t.ex. för att det inte fanns tillgängligt eller ansetts vara relevant)?>

## 6.4 Utvärdering av modellen

<Kort (1-2 meningar) beskrivning om vilka mått som använts.>

#### 6.4.1 Resultat

<Processens/modellens/algorithmens prestanda beskriven med siffror och grafer för helheten.>

#### 6.4.2 Användarresultat

Endast för processer.

<Resultat från användning i verkligheten. Kan vara under en testperiod eller i produktion (vilket det är ska anges). Ska ge en indikation om processens prestanda, exempel på mätvärden: Användarbetyg, antal omprövningar och andel som ändrade resultatet, stickprovskontroller>

#### 6.4.3 Unitära resultat

<Vad är modellens prestanda för de enskilda faktorerna (t.ex. olika kön, byggnadsändamål eller meningslängder) beskriven med siffror och grafer för helheten?>

#### 6.4.4 Trösklar och jämförelser

<Trösklar för när processen/modellen/algorithmen (enligt oss) är tillämpbar (t.ex. ett krav på precision/recall från "högre instans"), jämförelser med andra processer/modeller/algoritmer och metoder, m.m.>

#### 6.4.5 Utvärderingsmetod

<Hur har utvärderingsvärdena (t.ex. precision/recall) räknats fram? Kan hänvisa till extern information om oförändrade standardmetoder används och en egen beskrivning inte tillför något.>

### 6.5 Data

#### 6.5.1 Utvärderingsdata

<Information om det data som använts för att genomföra utvärderingen (källa, utbredning, förbehandling, m.m.). uppmärksamma särskilt om det finns personuppgifter eller skyddsvärd information i utvärderingsdata.>

#### 6.5.2 Träningsdata

<Information om det data som använts för att träna modellen (källa, utbredning, förbehandling, m.m.). uppmärksamma särskilt om det finns personuppgifter eller skyddsvärd information i utvärderingsdata.>

## 6.6 Övervakning, loggning och profilering

<Beskriv om systemet genererar loggar samt någon form av övervakning. Ta särskild hänsyn till mänsklig interaktion med dessa loggar eller övervakning>

## 6.7 Etiska aspekter

<EU-kommissionens Etiska riktlinjer för tillförlitlig AI ställer upp tre komponenter i sitt ramverk för tillförlitlig AI: Inkluderar, men är inte begränsat till: Laglig AI, Etisk AI och Robust AI.

Etisk AI omfattar i sin tur fyra principer:

- Respekt för människans autonomi
- Förebyggande av skada
- Rättvisa
- Förklarbarhet.

Dessa principer finns förklarade i detalj i riktlinjerna. De kan med fördel vara vägledande för den redovisning som görs under denna rubrik.>

## 6.8 Juridiska aspekter

<Krav som ställs på processen/modellen/algoritmen utifrån gällande lagar och förordningar. Kan ha delvis överlapp med Etiska aspekter, t.ex. gällande personuppgifter. Uppdelningen som gäller är att kapitlet Etiska ställningstaganden ska hantera den etiska sidan ("bör man") och Juridiska aspekter ska hantera den juridiska sidan ("får man").>

### 6.8.1 Integritetslagstiftning

<Uppmärksamma särskilt den data som används och då i förhållande till t.ex. GDPR>

### 6.8.2 Sekretess eller annan skyddslagstiftning

<Överväg vilken information gällande sekretess som är relevant att uppmärksamma>

### 6.8.3 Omprövning och överklagande

Enbart för processer.

<Kan man få beslutet omprövat (av en mänsklig handläggare) eller överklagat? Hur?>

#### 6.8.4 Tillsyn

<Finns det någon som har tillsyn? Externt eller internt? Myndighet, företag eller annan?>

#### 6.9 Betänkanden och rekommendationer

<Finns det några ytterligare aspekter att ta hänsyn till, som inte hanterats i någon annan rubrik? Indikerar t.ex. resultaten på att det behövs mer tester? Finns det några faktorer/grupper som inte täckts av det använda data? Finns det några andra betänkanden eller rekommendationer som en användare eller vidareutvecklare av modellen bör tänka på?>

# 7 Bilaga 3 – Användningsfall på Lantmäteriet

## 7.1 Förtroendemodellkort - Förändringsdetektering - Processen

### Exempel på Förtroendemodellkort, process

#### 7.1.1 Översiktlig Sammanfattning

<b>Typ</b>	Process
<b>In</b>	Sant ortofoto (skalriktig flygbild) Höjdmodell (DSM) Existerande byggnader
<b>Ut</b>	Förändrade byggnader
<b>Består av</b>	<a href="#">Detekteringsmodell</a> <a href="#">Förändringsalgoritm</a>

##### 7.1.1.1 Tänkt användning

Processen är utvecklad för att användas i Lantmäteriets ajourhållning av byggnadsgeometrier, som sedan bl.a. används i våra kartprodukter.

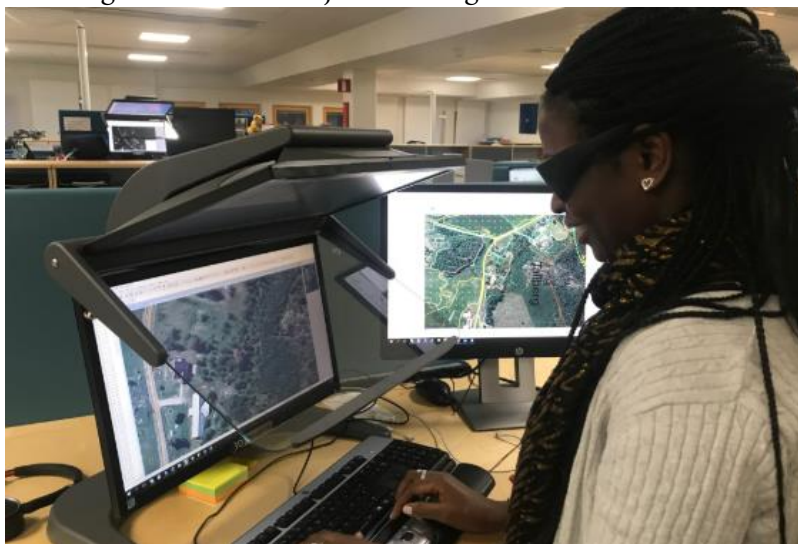
De ortofoton som används är framställda baserade på våra egna flygfoton och höjdmodellen baserat våra egna flygfoton samt laserskanning. De existerande byggnaderna hämtas ur våra kartprodukter.

Informationen om de förändrade byggnaderna används sedan av en mänsklig operatör för att uppdatera våra kartprodukter.

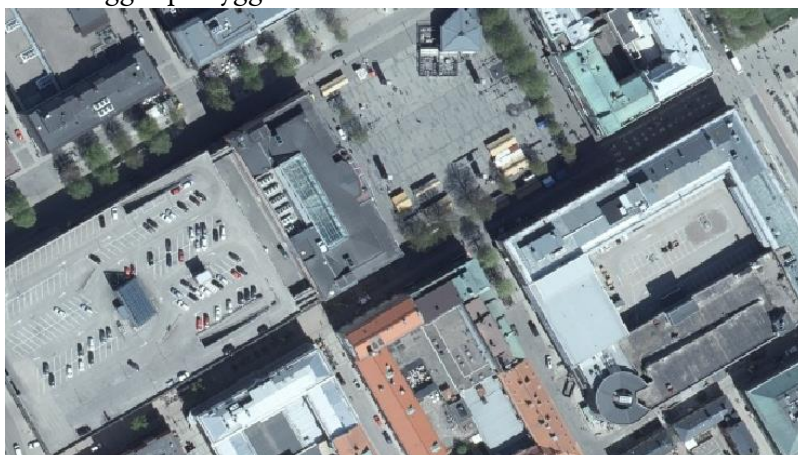


### 7.1.1.2 Begränsningar

- Ingen information uppdateras automatiskt, resultatet används bara som underlag i den manuella ajourhållningen



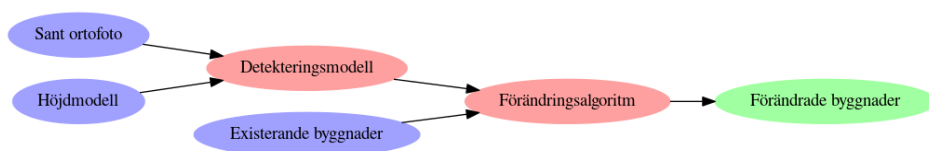
- Fokus ligger på byggnader utanför tätorterna



- Processen hittar inte täckta byggnader, t.ex. under tät skog



### 7.1.1.3 *Processöversikt*



### 7.1.2 **Processinformation**

Processen har utvecklats av personal på Lantmäteriet under 2019.

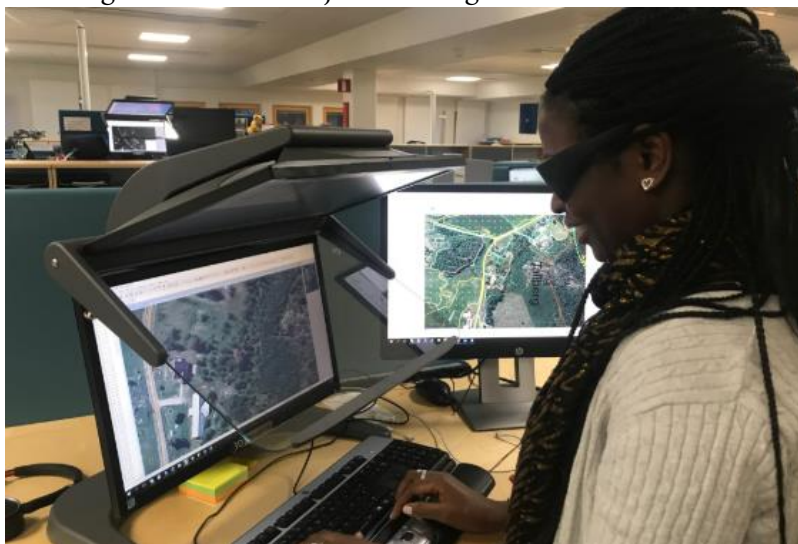
Detta är första versionen av processen.

#### 7.1.2.1 *Teknisk plattform*

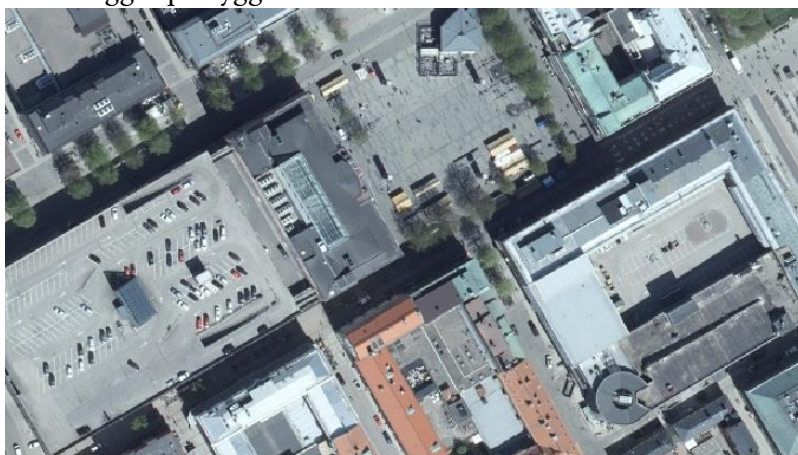
Processen körs uppdelad, maskininlärningsbiten körs på igång manuellt på en lokal Linux-server. Förändringsdetekteringen körs av ajourhållare på egen dator.

### 7.1.2.2 Begränsningar

- Ingen information uppdateras automatiskt, resultatet används bara som underlag i den manuella ajourhållningen



- Fokus ligger på byggnader utanför tätorterna



- Processen hittar inte täckta byggnader, t.ex. under tät skog



#### 7.1.2.3 *Implementationsdetaljer*

-

#### 7.1.2.4 *Artiklar eller annan resurs för mer information*

-

#### 7.1.2.5 *Licens*

-

#### 7.1.2.6 *Kontaktperson*

Jan Dalheimer, Mattias Petterson och Anders Ekholm, Lantmäteriet.

### 7.1.3 **Avsedd användning.**

Processen är avsedd att användas vid ajourhållning av Lantmäteriets topografiska kartor.

#### 7.1.3.1 *Användningsfall utanför tillämpningsområdet*

Liknande process skulle kunna användas för andra ytoobjekt än byggnader genom att träna om modellen och anpassningar av detekteringsalgoritmen.

## 7.1.4 Faktorer

### 7.1.4.1 *Relevanta faktorer*

Modellens resultat påverkas i hög grad av de indata som används. Några av de viktigare faktorerna som kan påverka:

- Plats i Sverige
  - Både grövre indelning som norr/söder och olika län, men också lokala variationer
  - Typ av område, landsbygd eller skog t.ex.
- Fototidpunkt, t.ex. innan eller efter lövsprickning
- Kamera och efterbearbetning
- Existerande byggnader

### 7.1.4.2 *Utvärderingsfaktorer*

För vår användning kan kamera och efterbearbetning anses vara konstant, därav har den faktorn inte utvärderats något vidare.

För att minimera påverkan av vart i Sverige ett område ligger liksom fototidpunkt så har tränings- och testdata valts som ska ge en så representativ bild som möjligt, med en god spridning både spatialt och temporalt (se kapitel 6 och 7). För att påvisa resultatet av detta har även resultat utbrutna i dessa faktorer presenterats.

Påverkan av kvalitén av existerande byggnader har inspekterats manuellt.

## 7.1.5 Utvärdering av processen

### 7.1.5.1 *Resultat*

Vår process får **60 % precision** (d.v.s. 60 % av de förändringar vi hittar behöver åtgärdas) och **73 % recall** (d.v.s. hittar vi 73 % av de förändringar som behöver åtgärdas) med tröskelvärdet för när en pixel är en byggnad X.

### 7.1.5.2 *Användarresultat*

Samtliga testanvändare ansåg att mängden felträffar gjorde resultatet oanvändbart.

### 7.1.5.3 *Unitära resultat*

...

### 7.1.5.4 *Intersektionella resultat*

...

### 7.1.5.5 Trösklar och jämförelser

...

### 7.1.5.6 Utvärderingsmetod

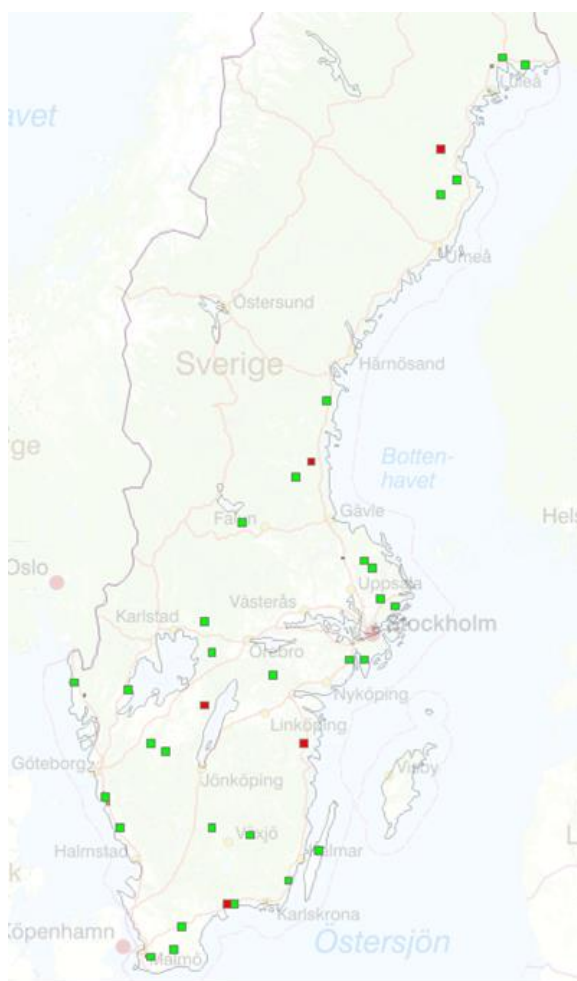
...

## 7.1.6 Data

### 7.1.6.1 Utvärderingsdata

Data från 5 områden om 10x10 km vardera med en upplösning på 25 cm/pixel användes. Områdena, som ligger geografiskt utspridda över hela Sverige, valdes för att ge en jämn representation av olika delar av landet och olika fototidpunkter.

Förbehandlingen bestod av normalisering och utklipp i 224x224 pixlar stora rutor med överlapp.

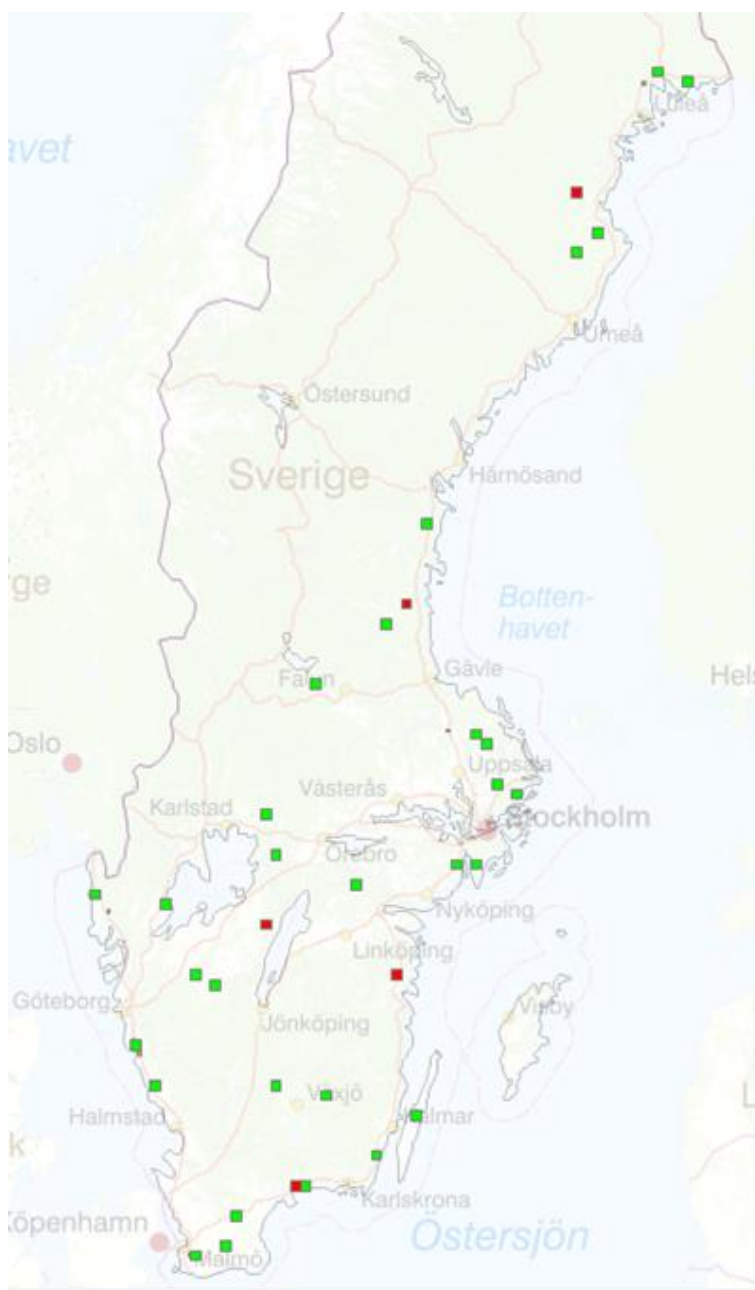


Röda rutor är utvärderingsdata.

### 7.1.6.2 Träningsdata

Data från 30 områden om 10x10 km vardera med en upplösning på 25 cm/pixel användes. Områdena, som ligger geografiskt utspridda över hela Sverige, valdes för att ge en jämn representation av olika delar av landet och olika fototidpunkter.

Förbehandlingen bestod av normalisering och utklipp i 224x224 pixlar stora rutor med överlapp.



Gröna rutor är träningsdata.

### 7.1.7 Övervakning, loggning och profilering

Inte tillämpligt.

### 7.1.8 Etiska aspekter

För denna tillämpning har den statliga värdegrundens sex principer gällt och dessa uttrycks här i den för AI omtolkade formen:

**Demokrati:** AI verkar och verkställer under offentligt anställda personers ansvar utifrån de beslut som riksdag och regering fattar. AI ska också bidra till att bygga och bevara förtroendet för myndigheterna och staten.

**Legalitet:** AI följer de lagar och regler som gäller. Det finns fungerande rutiner och processer för att upprätthålla rättssäkerhet i de åtgärder AI vidtar så att myndighetens verksamhet är rättssäker. Automatiska beslut som inverkar menligt på enskilds rätt kan, på begäran, alltid prövas om av en offentligt anställd person.

**Objektivitet:** AI saknar, i grunden, personliga relationer som ger upphov till intressekonflikter, jävssituationer och korruption. Automatiska åtgärder och beslut vilar på saklig och opartisk preparering, behandling och träning av AI-algoritmer och modeller.

**Fri åsiktsbildning:** Öppenhet är en grundprincip i demokratin. Myndigheten har en god offentlighetsstruktur och AI verkar utifrån offentlighetsprincipen med den grundläggande insyn i myndigheters verksamhet som den ger. Myndighetens AI är transparent och tillgängligt för granskning och kontroll. AI används inte för att efterforska källor eller personer som använder rätten till yttrande- och meddelarfrihet

**Respekt:** AI ska utföra åtgärder för människor lika och respektfullt utan att mänskliga fri- och rättigheter kränks och utan att någon diskrimineras på grund av exempelvis kön, könsöverskridande identitet eller uttryck, etnisk tillhörighet, religion eller annan trosuppfattning, funktionsnedsättning, sexuell läggning eller ålder.

**Effektivitet och service:** AI tillämpas för effektivitet med service och tillgänglighet. Myndigheten informerar, vägleder och ger information om AI och



dess tillämpning på ett enkelt och begripligt sätt i de sammanhang där det är aktuellt. Tillämpningen av AI sker ett effektivt sätt som hushållar med resurserna.

För Lantmäteriets tillämpningar och system, inkl. AI, gäller också Lantmäteriets värderingsgrund.

**Service:** Vi eftersträvar enkelhet och ser vår roll i ett större sammanhang. Vi är lyhörda och ger alltid ett vänligt och professionellt bemötande. Vi bidrar till varandras framgång och arbetsglädje

**Öppenhet:** Vi litar på varandra och andra kan lita på oss. Vi visar respekt för varandras åsikter och olikheter. Vi är öppna, tydliga och begripliga.

**Handlingskraft:** Vi skapar resultat genom att ta initiativ och vara öppna för nya lösningar. Vi genomför det vi har bestämt. Vi använder återkoppling för att utveckla verksamheten.

### 7.1.9 Juridiska aspekter

#### 7.1.9.1 *Integritetslagstiftning*

Ingen del av processen omfattas av integritetslagstiftning.

#### 7.1.9.2 *Sekretess eller annan skyddslagstiftning*

Processen använder icke-sekretessgranskade ortofoton och underliggör därmed samma krav som annan hantering av ortofoton som ej genomgått sekretessgranskning. Resultatet (byggnadsförändringar) anses dock inte omfattas av detta.

#### 7.1.9.3 *Omprövning och överklagande*

Inte tillämpligt.

#### 7.1.9.4 *Tillsyn*

Ingen.

### 7.1.10 Betänkanden och rekommendationer

-

## 7.2 Förtroendemodellkort - Förändringsdetektering - Detekteringsmodell

### Exempel på Förtroendemodellkort, modell

#### 7.2.1 Översiktlig Sammanfattning

<b>Typ</b>	ML Modell
<b>In</b>	Sant ortofoto (skalriktig flygbild) Höjdmodell (DSM)
<b>Ut</b>	Byggnader
<b>Ingår i</b>	<a href="#">Förändringsdetektering</a>
<b>Arkitektur</b>	U-Net med ResNet50

##### 7.2.1.1 Tänkt användning

Modellen är utvecklad för att användas för att hitta förändringar på byggnadsgeometrier.

De ortofoton som används är framställda baserade på våra egna flygfoton och höjdmodellen baserat våra egna flygfoton samt laserskanning.

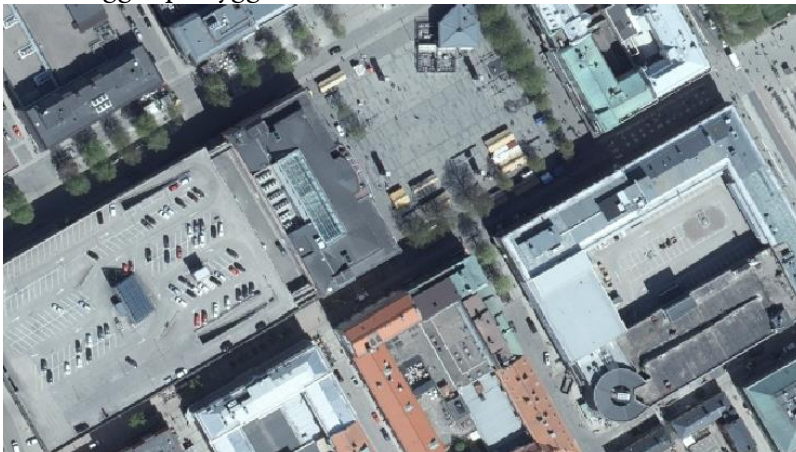
De hittade byggnaderna jämförs sedan mot existerande kartprodukter för att hitta förändringar.

##### 7.2.1.2 Prestanda

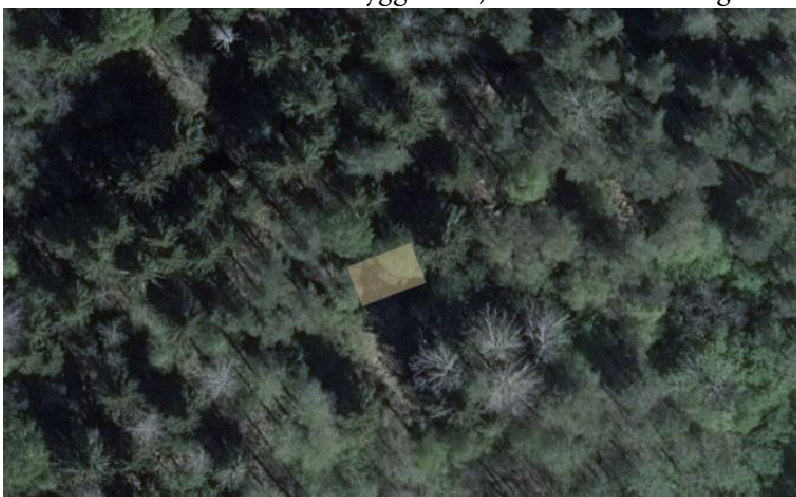
Modellen hittar 96 % av de byggnader som finns och 94 % av de byggnader som hittas finns.

### 7.2.1.3 Begränsningar

- Fokus ligger på byggnader utanför tätorterna



- Processen hittar inte täckta byggnader, t.ex. under tät skog



- De byggnadsgeometrier som modellen skapar är relativt grova och inte rektangulära som de flesta byggnader är



### 7.2.2 Modellinformation

Modellen har utvecklats och tränats av personal på Lantmäteriet under 2019.

Detta är första versionen av modellen.

### 7.2.2.1 Arkitektur

Vi använder oss av U-Net arkitekturen som tagits fram av forskare vid Freiburgs Universitet i Tyskland. Den utvecklades ursprungligen för biomedicinsk forskning, bl.a. för att markera celler i bilder, men den har fått tillämpning inom många andra områden för att hitta saker i bilder.

### 7.2.2.2 Träning

Modellen är tränad på bilder med storleken 224x224 pixlar med en upplösning om 25 cm/pixel. Följande träningskonfiguration användes:

<b>Loss-funktion</b>	Jaccard
<b>Optimizer</b>	Adam
<b>Learning rate</b>	0,002
<b>Epoker</b>	50
<b>Batch storlek</b>	32 bilder

### 7.2.2.3 Begränsningar

Modellen har problem att hitta byggnader som är delvis täckta av t.ex. skog:



Den får även visa felträffar på bl.a. större stenar, timmerupplag och grushögar:



#### 7.2.2.4 *Implementationsdetaljer*

Implementation använder sig av ramverket PyTorch.

#### 7.2.2.5 *Artiklar eller andra resurser för mer information*

U-Net: <https://arxiv.org/abs/1505.04597>

#### 7.2.2.6 *Licens*

U-Net är licensfri.

PyTorch är öppen källkod med en BSD-liknande licens.

#### 7.2.2.7 *Kontaktperson*

kundtjanst@lm.se

### 7.2.3 **Avsedd användning**

Modellen är framförallt avsedd att användas som ett delsteg för att hitta förändringar på byggnadsgeometrier med ett manuellt bearbetningssteg innan förändringarna läggs in i produkter.

Primärt avsedda användare är personer som jobbar med ajourhållning av byggnader i kartor.

### 7.2.3.1 *Användningsfall utanför tillämpningsområdet*

Man skulle även kunna skapa en karta med byggnader baserat direkt på resultatet av denna modell, men med relativt dålig kvalitet.

## 7.2.4 Faktorer

### 7.2.4.1 *Relevanta faktorer*

Modellens resultat påverkas i hög grad av de indata som används. Några av de viktigare faktorerna som kan påverka:

- Plats i Sverige
  - Både grövre indelning som norr/söder och olika län, men också lokala variationer
  - Typ av område, landsbygd eller skog t.ex.
- Fototidpunkt, t.ex. innan eller efter lövsprickning
- Kamera och efterbearbetning

### 7.2.4.2 *Utvärderingsfaktorer*

För vår användning kan kamera och efterbearbetning anses vara konstant, därav har den faktorn inte utvärderats något vidare.

För att minimera påverkan av vart i Sverige ett område ligger liksom fototidpunkt så har tränings- och testdata valts som ska ge en så representativ bild som möjligt, med en god spridning både spatialt och temporalt (se kapitel 6 och 7). För att påvisa resultatet av detta har även resultat utbrutna i dessa faktorer presenterats.

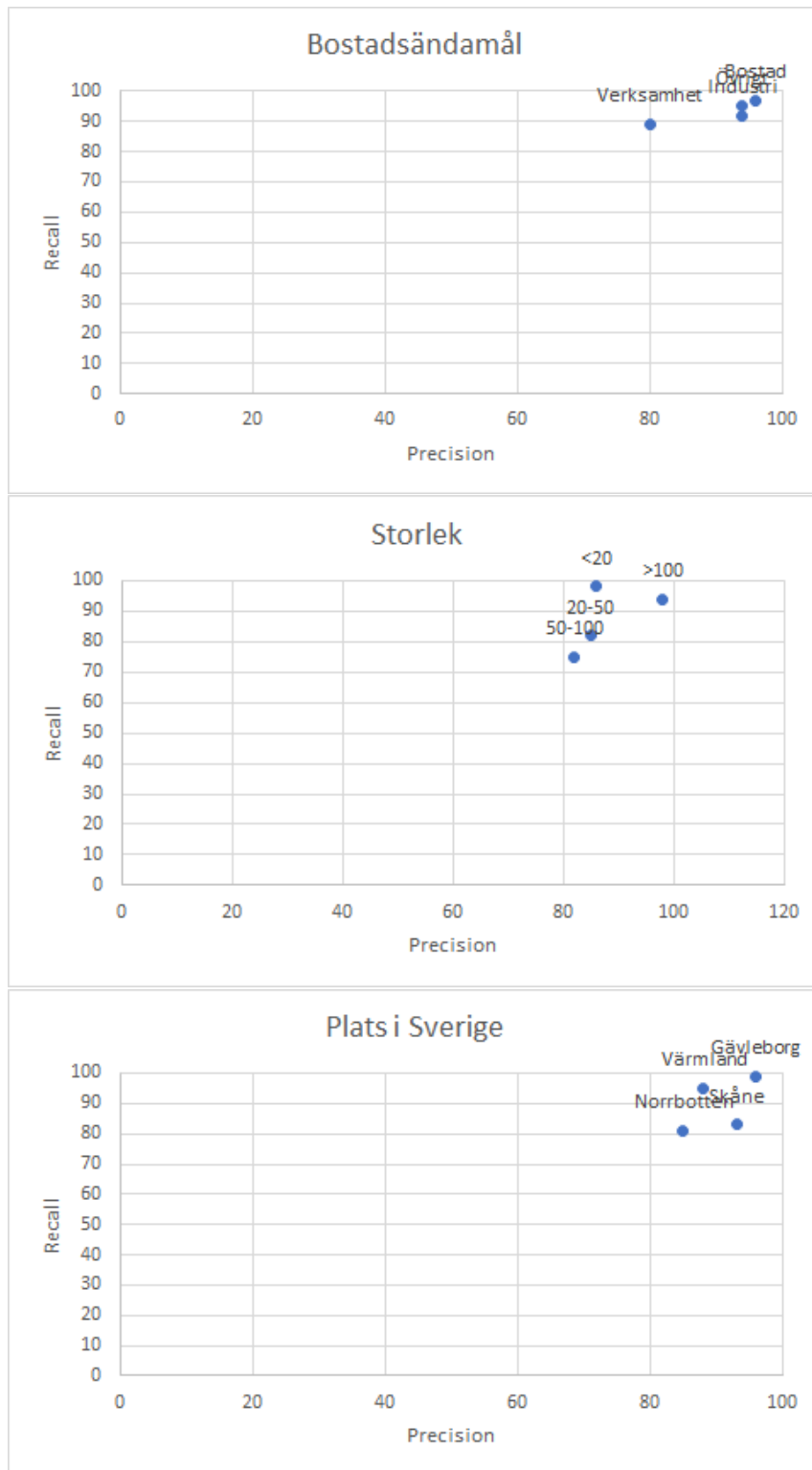
## 7.2.5 Utvärdering av modellen

Våra primära utvärderingsvärden vi använder oss av är Precision och Recall som är vanliga mått i utvärderingen av maskininlärningsmodeller.

### 7.2.5.1 *Resultat*

Vår modell får 94 % precision (d.v.s. 94 % av de byggnader vi hittar finns på riktigt) och 96 % recall (d.v.s. hittar vi 96 % av de byggnader som finns) med tröskelvärde för när en pixel är en byggnad X.

### 7.2.5.2 Unitära resultat



### 7.2.5.3 Intersektionella resultat

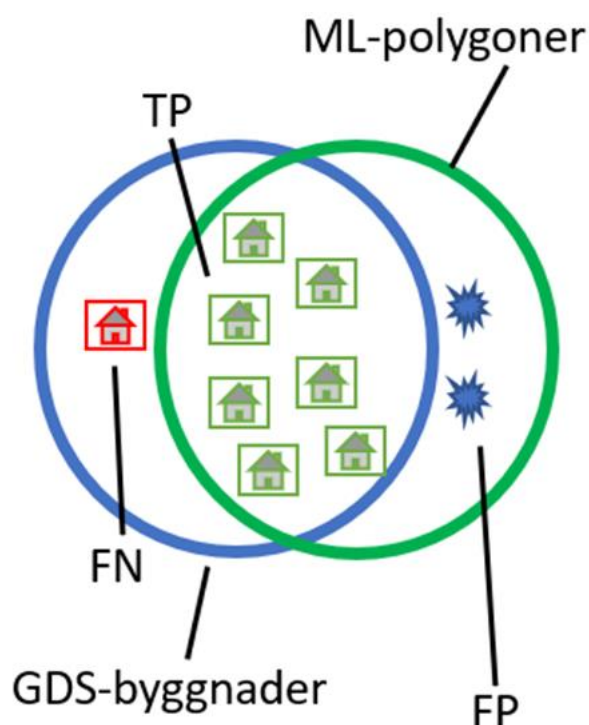
Inte tillämpligt.

### 7.2.5.4 Trösklar och jämförelser

Uppskattningsvis så hittar en människa under normal ajourhållning 80-90 % av alla byggnader (recall) och 99% av de byggnader som hittas är faktiska byggnader (precision).

### 7.2.5.5 Utvärderingsmetod

Precision och recall baserar sig på antal sanna träffar, antal falska träffar och antal missar.




$$\textit{Precision} = \frac{TP}{TP + FP}$$

$$\textit{Recall} = \frac{TP}{TP + FN}$$

För indelning i sanna träffar, falska träffar och missar överlades byggnadsgeometrierna från modellen och de vi har i vår databas för att få ett IoU-värde (intersection-over-union, se nedan).



$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$


**Sedan applicerades följande regler:**

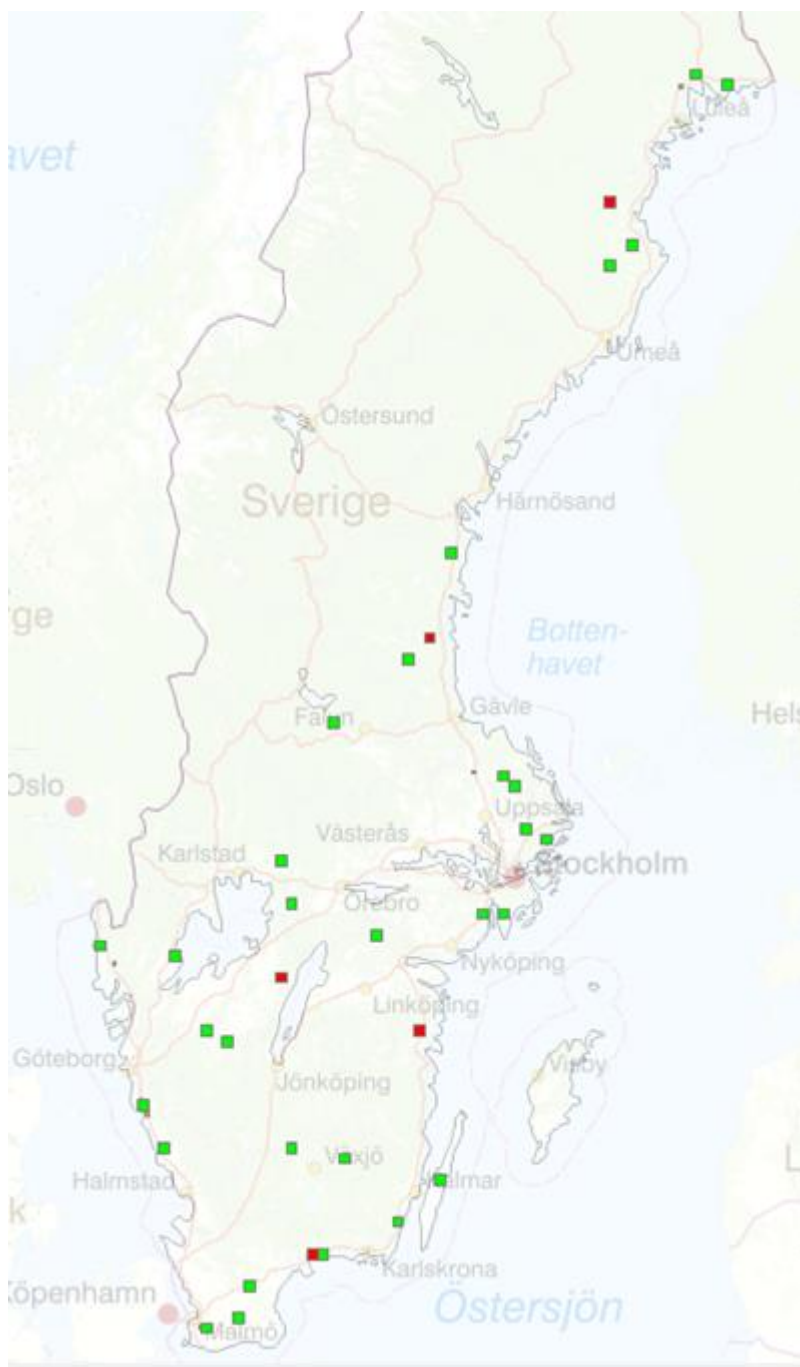
- Om IoU är större än 0,5 (d.v.s. mer än 50 % av ytorna sammanfaller) är det en sann träff
- Om IoU är mindre än 0,5 (d.v.s. mindre än 50 % av ytorna sammanfaller) och ytan för maskininlärnings-byggnaden är större är det en falsk träff
- Om IoU är mindre än 0,5 (d.v.s. mindre än 50 % av ytorna sammanfaller) och ytan för maskininlärnings-byggnaden är mindre är det en miss
- Maskininlärnings-byggnader som inte vidrör en byggnad i databasen är en falska träffar
- Byggnader i databasen som inte vidrör en maskininlärnings-byggnad är en miss

## 7.2.6 Data

### 7.2.6.1 Utvärderingsdata

Data från 5 områden om 10x10 km vardera med en upplösning på 25 cm/pixel användes. Områdena, som ligger geografiskt utspridda över hela Sverige, valdes för att ge en jämn representation av olika delar av landet och olika fototidpunkter.

Förbehandlingen bestod av normalisering och utklipp i 224x224 pixlar stora rutor med överlapp.

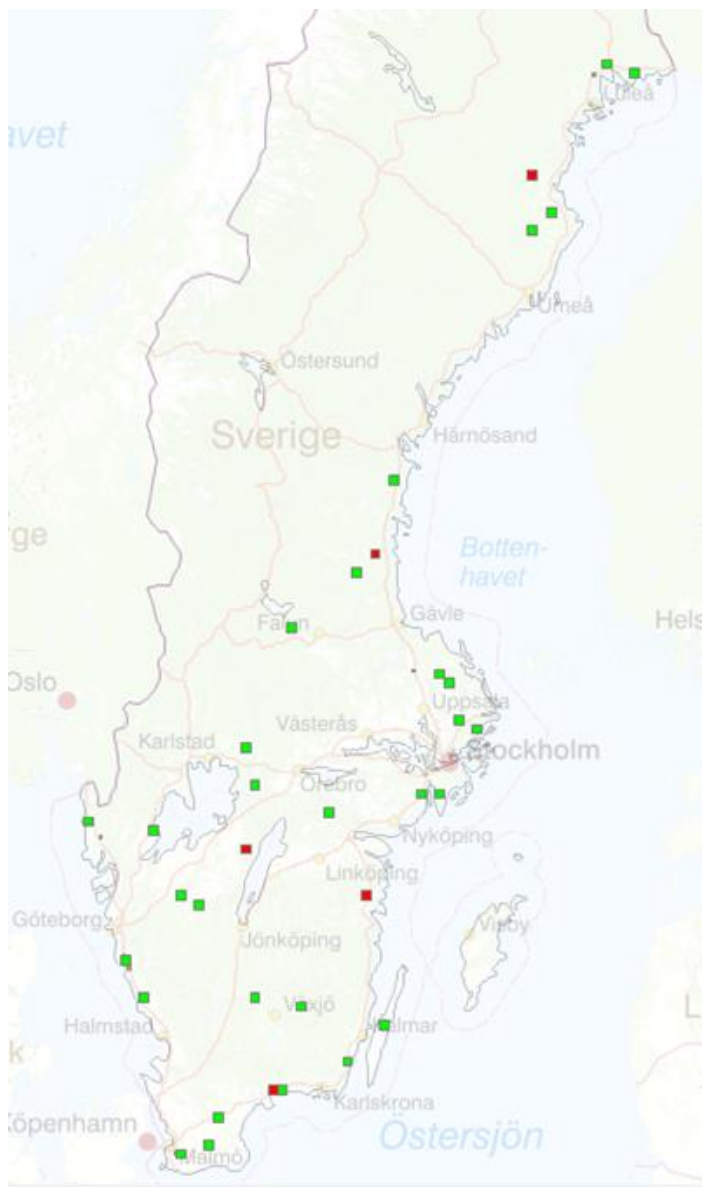


Röda rutor är utvärderingsdata.

### 7.2.6.2 Träningsdata

Data från 30 områden om 10x10 km vardera med en upplösning på 25 cm/pixel användes. Områdena, som ligger geografiskt utspridda över hela Sverige, valdes för att ge en jämn representation av olika delar av landet och olika fototidpunkter.

Förbehandlingen bestod av normalisering och utklipp i 224x224 pixlar stora rutor med överlapp.



Gröna rutor är träningsdata.

### 7.2.7 Etiska aspekter

Denna modell har potentialen att hitta byggnader som med avsikt dolts. Det anses dock inte som sannolikt att detta sker i större omfattning än vad som sker vid manuell ajourhållning.

## 7.2.8 Juridiska aspekter

### 7.2.8.1 *Integritetslagstiftning*

Inte tillämpligt.

### 7.2.8.2 *Sekretess eller annan skyddslagstiftning*

Processen använder icke-sekretessgranskade ortofoton och underliggjer därmed samma krav som annan hantering av ortofoton som ej genomgått sekretessgranskning. Resultatet (byggnadsförändringar) anses dock inte omfattas av detta.

## 7.2.9 Betänkanden och rekommendationer

- De byggnadsgeometrierna som skapas är i de flesta fall inte tillräckligt "rektangulära" för att direkt användas i en kartprodukt