

Stora data & öppna data

Exempel på policyinitiativ från Asien, Europa och USA

Hanteringen av stora datamängder och öppna data reser en lång rad frågor kring policy. Hur kan utvecklingen leda till ökad tillväxt, innovation och bättre tjänster? Vilken är statens eller politikens roll på området? Utifrån dessa frågor har Tillväxtanalys studerat hur några länder arbetar med policyinitiativ på områdena öppna data och stora data. Studien innehåller exempel från Indien, Japan, Storbritannien, Sydkorea och USA.

Dnr: 2013/289
Swedish Agency for Growth Policy Analysis
Studentplan 3, SE-831 40 Östersund, Sweden
Telephone: +46 (0)10 447 44 00
Fax: +46 (0)10 447 44 01
E-mail: info@growthanalysis.se
www.growthanalysis.se

For further information, please contact Carl Jeding
Telephone: +46 10-447 44 78
E-mail: carl.jeding@tillvaxtanalys.se

Förord

I takt med digitaliseringen av allt fler funktioner i samhället växer mängden lagrade data exponentiellt. Att på ett smart sätt utnyttja dessa datamängder ses av många som en viktig drivkraft för ökad tillväxt.

Sverige har i internationell jämförelse en mycket stark tradition av att både samla in offentliga data och av öppenhet. Hanteringen av stora datamängder och öppna data reser emellertid en lång rad policyutmaningar. Hur kan utvecklingen bidra till innovation och bättre tjänster? Vilken är politikens roll på området?

Inom Tillväxtanalys uppdrag kring löpande omvärldsbevakning har vi tagit fram en studie om hur några länder arbetar med dessa policyutmaningar på områdena öppna data och stora data. Studien innehåller exempel från Indien, Japan, Storbritannien, Sydkorea och USA. Utifrån dessa exempel lyfter rapporten fram några områden där Sverige kan hitta inspiration och lärdomar för en politik som tar tillvara utvecklingsmöjligheterna med öppna data och stora data.

Denna rapport är skriven av Carl Jeding (Storbritannien), Niklas Z Kviselius, Shigeyuki Naito, Yoonjin Cho och Emelie Gustafsson (Japan och Sydkorea), Peter Wennerholm och Andreas Muranyi-Scheutz (Indien) och Anna Ledin (USA). Projektet har letts av Sigrid Hedin vid Tillväxtanalys Stockholmskontor.

Stockholm, oktober 2013

Enrico Deiacò

Avdelningschef, Innovation och globala mötesplatser

Tillväxtanalys

Innehållsförteckning

| | |
|---|-----------|
| Sammanfattning | 7 |
| Summary | 9 |
| 1 Policyinitiativ inom stora data och öppna data för innovation inom näringsliv och förvaltning..... | 11 |
| 1.1 Vad är stora data och öppna data? | 11 |
| 1.2 Stora data och öppna data som tillväxtfråga | 11 |
| 1.3 Rapportens disposition..... | 13 |
| 2 Indien | 14 |
| 2.1 Sammanfattning | 14 |
| 2.2 Den indiska IT-industrin – en översikt | 14 |
| 2.3 Riktade policyer och strategier | 15 |
| 2.3.1 Policyer och strategier avseende stora data | 15 |
| 2.3.2 Planerade policyer och strategier avseende stora data | 17 |
| 2.3.3 Policyer och strategier avseende öppna data | 18 |
| 2.3.4 Planerade policyer och strategier avseende öppna data | 19 |
| 2.4 Framtida utmaningar | 20 |
| 3 Japan | 23 |
| 3.1 Sammanfattning | 23 |
| 3.2 Riktad policy och strategier | 23 |
| 3.3 Resultat inom innovation och värdeskapande..... | 25 |
| 3.3.1 Gruvindustri..... | 26 |
| 3.3.2 Hälso- och sjukvård..... | 26 |
| 3.3.3 Social innovation..... | 27 |
| 3.4 Utmaningar för vidare utveckling..... | 27 |
| 3.4.1 Lagstiftning..... | 27 |
| 3.4.2 Främjande..... | 28 |
| 3.4.3 Standardisering..... | 29 |
| 3.4.4 Forskning och utveckling..... | 29 |
| 3.4.5 Brist på experter..... | 30 |
| 4 Storbritannien | 31 |
| 4.1 Sammanfattning | 31 |
| 4.2 Selektiv politik och strategier..... | 31 |
| 4.2.1 Lagstiftning och översyn | 33 |
| 4.3 Exempel på användning inom innovation och värdeskapande..... | 34 |
| 4.4 Utmaningar för fortsatt utveckling..... | 35 |
| 5 Sydkorea..... | 36 |
| 5.1 Sammanfattning | 36 |
| 5.2 Inledning..... | 36 |
| 5.3 Riktad policy och strategier | 37 |
| 5.3.1 Lagstiftning och översyn | 39 |
| 5.4 Resultat inom innovation och värdeskapande..... | 40 |
| 5.4.1 Tillverkningsindustri..... | 42 |
| 5.4.2 Hälso- och sjukvård..... | 43 |
| 5.4.3 Social innovation..... | 43 |
| 5.5 Utmaningar för vidare utveckling..... | 44 |
| 5.5.1 Främjande..... | 44 |
| 5.5.2 Standardisering..... | 44 |
| 5.5.3 Forskning och utveckling..... | 45 |
| 5.5.4 Brist på experter..... | 46 |
| 6 USA..... | 47 |
| 6.1 Sammanfattning | 47 |
| 6.2 Inledning..... | 47 |
| 6.3 Riktade policyer och strategier | 48 |
| 6.3.1 Stora data | 48 |
| 6.3.2 Öppna data | 49 |
| 6.4 Resultat i form av innovationer och värdeskapande..... | 50 |

| | | |
|----------|---|-----------|
| 6.5 | Framtida utmaningar | 51 |
| 7 | Sammanfattande resultat..... | 52 |
| 7.1 | Forskningssatsningar | 52 |
| 7.2 | Infrastruktur | 52 |
| 7.3 | Kompetens | 53 |
| 7.4 | Insatser för att stödja myndigheter och andra aktörer | 53 |
| 7.5 | Kostnader för datatillgång | 53 |
| 7.6 | Integritet och säkerhet..... | 54 |
| 7.7 | Attraktionskraft, synlighet och tillgänglighet..... | 54 |
| 8 | Lärdomar för Sverige | 55 |

Sammanfattning

Rapporten beskriver vissa policyinitiativ inom områdena öppna data och stora data med exempel från Indien, Japan, Storbritannien, Sydkorea och USA. Rapporten riktar sig i första hand till svenska beslutsfattare inom politik och statsförvaltning. Området studeras främst ur ett tillväxtperspektiv.

Landsexemplen visar att flera olika politiska initiativ tagits för att stimulera eller underlätta framväxten och öka användningen av stora data och öppna data. Det primära motivet bakom initiativen är att stora och öppna data har potential att bidra till en ökad innovationskraft, ekonomisk tillväxt och jobb. Några av länderna anstränger sig för att synas och positionera sig inom fältet och bygga ett starkt nationellt varumärke.

Inom vissa näringslivssektorer finns stora förhoppningar om ökat värdeskapande genom att bättre använda stora data – rapporten innehåller exempel från bland annat sjuk- och hälsovård, gruvindustri, transporter, tillverkningsindustri och energiproduktion.

Det finns en stor potential för offentligt initierad forskning att förbättra förmågan att utvinna kunskaper från stora och komplexa datamängder. Genomgående finns en förväntan om att stora och öppna data ska kunna hjälpa till att sänka sjukvårdskostnaderna i samhället – även om patientdata framstår som en av de mest svårhanterliga typerna av data att göra tillgängligt samtidigt som den personliga integriteten säkras.

Bristen på experter framträder som en stor utmaning på medellång sikt. Samtliga länder har utvecklat riktade insatser för att säkra den framtida tillgången på kompetens inom områden. Det bör särskilt noteras att det finns behov av kompetenslyft bland statliga organisationer som genom ny lagstiftning och policy ska börja offentliggöra data på ett strukturerat och lättillgängligt sätt.

I länderna görs nu satsningar på infrastruktur av olika slag, som superdatorer och datacenter, men även utbyggnad av bredband för att möjliggöra överföring av stora datamängder. Här ingår även stödprogram för nya företag inom IT-sektorn där dessa erbjuds tillgång till inkubatorer och acceleratorer. I flera länder ses små och medelstora företag som nyckelspelare för att ta tillvara den innovationskraft som finns i en bättre användning av öppna data och stora data.

Öppna data medför i samtliga fall en intensiv policydiskussion kring integritet, säkerhet, ägandeskap och ansvar. Här återfinns allt från stor konservatism, baserad på tidigare lagrum, till principen att all data ska offentliggöras utan extra kostnader för användaren. För länder som tidigare inte har tillämpat offentlighetsprinciper handlar det om att ändra både den formella lagstiftningen och förvaltningskulturen när data ska släppas fritt och kommunikationen med medborgaren ska förbättras.

Lärdomar som kan dras från ett svenskt perspektiv är:

- I de studerade länderna förekommer en ”sense of urgency” kring området och olika policysatsningar har initierats.
- Små och medelstora företag ses som nyckelaktörer för att utveckla och ta tillvara innovationer relaterade till stora och öppna data.

- Sverige har goda förutsättningar att vara en ledande nation inom öppna data, till exempel en hög grad av digitalisering, en stor och transparent offentlig sektor och en hög grad av tillit hos medborgarna.

Summary

This report examines policy initiatives in the areas of big data and open data, with examples from India, Japan, Britain, South Korea and the United States. The report is aimed primarily at Swedish government decision-makers. The area is examined mainly from the perspective of economic growth.

The cases show that countries have taken policy initiatives to encourage or facilitate the development and increased use of big and open data. The primary motive behind these initiatives is that big and open data have the potential to contribute to positive outcomes such as increased innovation, economic growth and jobs. Some of the countries are making an effort to position themselves in the field and build a strong national brand.

In some industry sectors there are high expectations of value creation by making better use of big data – this report contains examples including health care, mining, transportation, manufacturing, and energy production.

There is certainly potential for public sector initiated research, to improve the ability to extract knowledge from large and complex data sets. Throughout the country cases there is anticipation that big and open data are to help lower health care costs in society – even as patient data appears as one of the most difficult types of data to make available whilst ensuring privacy.

Lack of experts emerges as a challenge from a mid-term perspective. All of the countries have directed efforts to secure the future supply of skills in areas relevant to the increased use of big and open data. Related to the question of competence is a need for skill development also inside governmental organizations that through new legislation and policy are to begin publishing data in a structured and accessible way.

The countries in the study are investing in infrastructure of various kinds, such as supercomputers and data centers, but also the expansion of broadband to allow the transfer of larger data sets. One aspect of infrastructure is also implementation of support programs for start-up companies in the IT sector where companies are offered access to incubators and accelerators. In several of the countries small and medium sized enterprises are seen as key players in taking advantage of the innovation surrounding a better use of big and open data.

Open data in all cases spurs a discussion of privacy, security, ownership and responsibility. There is a range among the countries studied from large conservatism based on existing perhaps outdated law, to following the principle that all data should be published at no extra cost to the user. For countries that are new to a philosophy of open government, necessary changes are to be done in the formal legislative but also informative administrative culture in order to make sure that data is released to the public.

Lessons-learned from a Swedish perspective are:

- In the studied countries a "sense of urgency" in relation to big and open data can be observed and policy initiatives have been implemented.
- Small and medium sized companies are regarded as key actors for developing and commercializing innovations emerging from the use of big and open data.

- Sweden has great possibilities to be a leading country within the field of open data, for instance a high degree of digitization, a large and transparent public sector and a high degree of trust among the citizens.

1 Policyinitiativ inom stora data och öppna data för innovation inom näringsliv och förvaltning

Hantering av stora datamängder och öppna data har på senare tid gett upphov till en lång rad frågor kring policyinsatser. Liksom annan IT-utveckling förespås exempelvis användningen av stora data få avtryck på ett flertal områden, såsom stadsplanering, bank och finans-, telekom- samt hälso- och sjukvårdssektorn. Hur kan utvecklingen leda till bättre tjänster och innovation samt ökad tillväxt? Vilken är statens eller politikens roll på området? Detta har varit utgångsfrågor för denna studie. Det är viktigt att understryka att denna rapport på inget sätt gör anspråk på att utvärdera svensk politik på området, eller föra fram konkreta policyrekommendationer. Istället är syftet att lyfta fram några intressanta exempel på policyinitiativ från ett antal länder, där Sverige skulle kunna dra lärdomar av hur man hanterar frågor kring stora och öppna data.

1.1 Vad är stora data och öppna data?

”Big Data” eller stora data är en beskrivning av datamängder som är för stora för att hantera med konventionella datahanteringsverktyg och metoder. Den massiva ökningen av digital information sätter även press på en snabbare utveckling av tekniker för att hantera och analysera samt ta tillvara informationen. Vad som anses vara ”stora data” varierar beroende på kapacitet hos olika organisationer och begreppet är därmed relativt till sin karaktär. I följande studie har vi förstått begreppet stora data som möjligheter att skapa en bättre användning av data för att förbättra organisatoriskt (privat och offentligt) beslutsfattande.

”Öppna data” ses som en ansats att viss typ av data, i första hand sådana som offentliga aktörer förfogar över, ska göras tillgängliga för användning av utomstående aktörer, oftast utan kostnad. De idéer som ligger bakom denna ansats har funnits under en längre tid, medan själva termen är relativt ny och har fått ökad betydelse i och med att internetanvändningen har brett ut sig, och i takt med att policyinitiativ för öppna data har tillkommit.

Genom att kombinera flera källor av (öppna) data, till exempel för att skapa en viss applikation, kan stora informationsmängder genereras, vilket illustrerar sambandet mellan begreppen öppna data och stora data.

1.2 Stora data och öppna data som tillväxtfråga

Stora data och öppna data är egentligen två i viss mån skilda företeelser, som ställer krav på olika typer av statliga insatser. På en mer grundläggande nivå handlar dock båda om utvecklingen av vad som ibland kallas informationsekonomin – alltså en ekonomi där information och kunskap är den främsta värdeskapande tillgången. Vi har därför valt att studera företeelserna tillsammans i denna rapport.

Ett antal positiva samhällseffekter förväntas av en ökad åtkomst och tillämpning av stora data och öppna data. Det talas också ofta om att användningen av stora data och öppna data ska förstärka varandra. Förhoppningen är bland annat att användningen ska främja ny kunskap och innovationer som ska bidra till värdeskapande i kommersiell verksamhet och därmed leda till tillväxt. Det finns också många exempel på hur offentliga data används för att förbättra eller utveckla smartare offentliga tjänster.

Beträffande användningsområden inom den privata sektorn, anses stora data komma till användning både inom tillverkningsindustrin och tjänsteproducerade företag. Syftet med användningen är att organisera, analysera och skapa värde genom att använda data som genereras dagligen. Inom alla sektorer har mängden lagrad data ökat kraftigt, även om vissa sektorer, som till exempel försäkring, bank och finans, har pekats ut att särskilt kunna dra nytta av stora datamängder för bättre beslutsfattande.

Konkreta exempel på användning är bland annat inom detaljhandel, teknologi för smarta elnät och utökad användning av sensorer för att skapa mer intelligenta system. I takt med att inte bara människor utan även maskiner och föremål i allt högre utsträckning börjar generera data, kommer stora data att få ytterligare relevans. Även verksamheter såsom hälso- och sjukvård, bank- och finansiell verksamhet samt utvinning av råvaror förväntas kunna nå ökad effektivitet och produktivitet genom användning av stora data. Även vetenskapliga framsteg förväntas kunna uppnås inom komplexa och tvärvetenskapliga forskningsfrågor.

Stora data förväntas också förbättra offentlig förvaltning och styrning samt bidra till att finna lösningar på stora samhällsutmaningar. Exempelvis genom att bidra med underlag inom hälso- och sjukvårdssektorn, inklusive förbyggande arbete, sysselsättning, ekonomisk produktivitet, trygghet, naturkatastrofer och resurshantering. Exempel är skattehantering och bekämpning av bedrägerier och korruption.

Utmaningar som brukar nämnas för att främja en produktiv användning av stora data och öppna data är kompetens- och kunskapsbrist beträffande till exempel nya analytiska metoder och standardiserade format. Det förespås exempelvis uppstå en brist på experter inom datahantering. Brist på anpassning av politik och reglering kopplad till integritet, säkerhet, immateriella rättigheter etc. inom nya användningsområden anses också hämma utvecklingen.

Som vi visar i våra landexempel i rapporten har många länder tagit politiska initiativ på området för att stimulera eller underlätta framväxten och ett ökat användande av stora data och öppna data. Motiven bakom dessa initiativ är dels att stora och öppna data har potentialen att bidra till positiva resultat såsom en ökad innovationskraft, ekonomisk tillväxt och jobb. Dels bygger initiativen på uppfattningen att politikens insatser, till exempel genom lagstiftning, reglering eller tillgänglighet till infrastruktur och kompetens, påverkar hur olika aktörer tar till sig och utnyttjar teknik. Det är också denna uppfattning som ligger bakom vårt val att studera policyinitiativ på området.

Politikområdet har utvecklats på senare tid, till exempel har insatser för att tillgängliggöra offentlig data genomförts. Här har ansatsen ofta varit att säkerställa att data ska bli mer transparenta och användbara för offentliga och kommersiella ändamål, samtidigt som missbruk av data måste hindras. Det finns således en spänning mellan öppenhet och slutenhet. Förändrad lagstiftning syftandes till att främja kommersiell verksamhet och innovationer genom att tillgängliggöra offentlig data kan bland annat noteras från USA, Sydkorea, och Storbritannien.

Även satsningar på kunskaps- och kompetensutveckling inom området kan observeras. Till exempel forskningsinsatser beträffande insamling, lagring, sökning, delning, analys och visualisering av data. Större forskningsinitiativ har rapporterats från USA och EU.

För svenskt vidkommande har Vinnova utvecklat en plattform för öppna data (öppnadata.se) kopplat till ett regeringsuppdrag.¹ Stockholms Stad har också utvecklat en plan för öppna data² för att gradvis publicera data inom geodata (väginformation och kartor), miljödata som en respons till en förändrad lagstiftning i Sverige och EU³.

1.3 Rapportens disposition

Tonvikten i rapporten är titta närmare på några policyinsatser relaterade till stora data och öppna data som berör innovations- och tillväxtfrågor i Indien, Japan, Storbritannien, Sydkorea och USA. Utgångspunkten för landexemplen är följande områden:

- Policyinsatser och strategier
- Användning
- Vilka är utmaningarna för att dra nytta av stora data och öppna data?

I följande avsnitt redovisas situationen i de studerade länderna. Varje landrapport inleds med en kort uppsummerande sammanfattning. Efter genomgången av de studerade länderna följer ett avsnitt där huvuddragen av de som återfunnits i länderrapporterna sammanfattas. Studien avslutas med att lyfta fram några områden där policyinsatser kan vara av intresse för en lärande svensk IT och innovationspolitik.

¹ www.regeringen.se/sb/d/15697/a/197348

² insyn.stockholm.se/ks/document/2011-10-26/Dagordning/6/06%20d11-0457Bilaga.pdf

³ *PSI-direktivet, Europaparlamentets och Europeiska unionens råds direktiv 2003/98/EG, Lagen (SFS 2010:566) om vidareutnyttjande av handlingar från den offentliga förvaltningen och Proposition 2009/10:175 Offentlig förvaltning för demokrati, delaktighet och tillväxt*

2 Indien

2.1 Sammanfattning

Tjänster inom stora data och användandet av öppna data för olika typer av IT-tjänster ses som tillväxtområden i Indien och initiativ för att främja utvecklingen inom dessa båda områden har tagits såväl av landets regering som av dess näringsliv och enskilda lärosäten. Exempel på detta är etableringen av data.gov.in, en webplattform för att offentliggöra myndighetsdata, samt specifika åtgärder från IT-industrisammanslutningen NASSCOM:s sida i syfte att öka tillämpningen av öppna data liksom att säkerställa kompetensförsörjningen inom stora data. Kompetensutvecklingen sker till stor del tack vare de åtgärder som de stora indiska IT-företagen vidtar, exempelvis Tata Consultancy Services, Infosys och Wipro. Som exempel kan nämnas dessa företags strategiska förvärv av utländska företag med efterfrågad kompetens, samt strategin att ingå i forskningssamarbeten med olika forskningsorganisationer eller universitet, inhemska såväl som utländska. I dagsläget kan statens roll beskrivas som begränsad till att bedriva forskning och utveckling inom några få statliga forskningsorganisationer. Utmaningar i fråga om att säkerställa kompetensförsörjningen inom stora data handlar framför allt om att stärka kompetensen inom landets universitetsväsende och öka samarbetet mellan industrin och den akademiska världen. I dagsläget är detta samarbete mycket begränsat. En annan viktig fråga när stora data diskuteras handlar om att införa lagstiftning som säkerställer att den personliga integriteten respekteras. En process med syftet att införa en sådan lag pågår just nu i Indien.

2.2 Den indiska IT-industrin – en översikt

Den indiska IT-industrin är internationellt sett mycket konkurrenskraftig och omsätter i dagsläget uppskattningsvis cirka 650 miljarder kronor (100 miljarder USD). Branschens ambition är dock att öka omsättningen till hela 1950 miljarder kronor (300 miljarder USD) till år 2020.⁴ Det gångna räkenskapsåret beräknades exporten av IT-tjänster generera en omsättning på 492,7 miljarder kronor (75,8 miljarder USD) medan den inhemska marknaden hade en omsättning på 135,9 miljarder kronor (20,9 miljarder USD)⁵ För att uppnå de uppsatta omsättningsmålen måste IT-industrin inrikta sig mer på den inhemska marknaden liksom främja framväxten av små och medelstora företag. De små och medelstora företagens betydelse för den indiska IT-industrin ökar redan idag. Under 2011–12 stod dessa företag för nio procent av intäkterna, vilket kan jämföras med endast två procent år 2001–02.⁶ Tjänster inom stora data och användandet av öppna data för olika typer av IT-tjänster ses som tillväxtområden och den indiska regeringen har, liksom industrisammanslutningen NASSCOM⁷, tagit flera initiativ för att främja utvecklingen inom båda dessa områden vilket i korthet redogörs för nedan.

⁴ Mint, 4 mars 2013, "Nasscom to expand its focus to products, Internet, domestic IT", tillgänglig på <http://www.livemint.com/Industry/7gINodqepK30gLmgPz7iZO/Nasscom-to-expand-focus-to-products-Internet-and-domestic-I.html>

⁵ The Hindu Business Line, 12 februari 2013, "Indian IT exports to grow 12-14% in FY'14: Nasscom", tillgänglig på: <http://www.thehindubusinessline.com/industry-and-economy/info-tech/indian-it-exports-to-grow-1214-in-fy14-nasscom/article4407167.ece>

⁶ Mint, 13 februari 2013, "Tech start-ups move away from conventional IT services", tillgänglig på <http://www.livemint.com/Industry/foDlqyiFFfpSEZQSaW8NN/Tech-startups-move-away-fromconventional-IT-services.html>

⁷ National Association of Software and Services Companies

2.3 Riktade policyer och strategier

2.3.1 Policyer och strategier avseende stora data

Statliga initiativ som kan sägas vara direkt inriktade på att öka kompetensen inom stora data märks tydligast i form av forskningsverksamhet vid olika forskningsinstitut och genom den utbildning som bedrivs vid olika lärosäten.

Exempel på stora data-relaterad forskning är den verksamhet som bedrivs av Council of Scientific and Industrial Research (CSIR) och dess s.k. *Fourth Paradigm Institute* (CSIR-4PI) med huvudsäte i Bangalore (tidigare *CSIR Centre for Mathematical Modelling and Computer Simulation, CSIR-CMMACS*), liksom *Centre for Development of Advanced Computing* (C-DAC).⁸ *Fourth Paradigm Institute* inriktar sig bland annat på dataintensiv forskning i form av exempelvis klimatformskning, datamekanik (optimering och visualisering) och IT-säkerhet (kryptering, kryptoanalys). *Fourth Paradigm Institute* i Bangalore är utrustad med en superdator med beräkningshastighetskapacitet på 360 teraflop⁹ som är sammankopplad med mindre CSIR-4PI-laboratorier i de indiska städerna Chennai, Chandigarh, Hyderabad, Kolkata, Nagpur, New Delhi, och Pune. Dessa mindre laboratorier ska i sin tur vara utrustade med superdatorer med kapacitet i storleksordningen 10 till 50 teraflop. Samtliga CSIR-laboratorier, inte enbart CSIR-4PI-laboratorierna ska dessutom vara anslutna till det s.k. *National Knowledge Network* som regeringen 2010 fattade beslut om och som nu är under kontinuerlig utveckling. Detta innebär att även forskare vid dessa mindre laboratorier får tillgång till superdatorn i Bangalore.¹⁰ C-DAC bedriver forskning inom IT-säkerhet, hälsoinformatik och distribuerade beräkningar, så kallad *grid computing*. 2013 tog C-DAC i bruk en superdator med kapacitet uppgående till 500 teraflop.

Insatser för att öka kompetensen inom affärsområden relaterade till stora data, både tekniskt ingenjörskunnande och affärsutveckling, görs även inom det indiska universitetsväsendet. *Indian Institute of Management* – Bangalore driver till exempel en exekutivutbildning i *Business Analytics and Intelligence* för personer med minst fem års arbetslivserfarenhet inom denna sektor. Vidare kan nämnas att *Indian Institute of Management* – Calcutta i samarbete med *Hughes Education* erbjuder en ettårig distansutbildning i *Business Analytics* medan *Indian Institute of Technology* – Bombay erbjuder en certifikatutbildning i samma ämne tillsammans med *HughesNet Global Education*. *Indian Statistical Institute* i Pune erbjuder en utbildning inriktad på extrahering av data (data mining). I juli 2012 etablerade Yahoo ett s.k. *Grid Computing Lab* vid *Indian Institute of Technology* – Madras i syfte att möjliggöra forskning inriktad på stora data och digital marknadsföring.¹¹ Avslutningsvis kan nämnas att det indiska läkemedelsföretaget Biocon

⁸ CSIR Fourth Paradigm Institute, <http://www.4pi.in/> CSIR är en del av den indiska regeringens Ministry of Science and Technology, medan C-DAC är en enhet inom Department of Electronics and Information Technology (DeitY) vid Ministry of Communications & Information Technology (MCIT).

⁹ Teraflop är ett mått på beräkningshastighetskapacitet på en miljard flyttalsberäkningar per sekund

¹⁰ "Repositioning of CSIR-CMMACS as nucleus of CSIR-Fourth Paradigm", tillgängligt på http://www.csir.res.in/csir/external/heads/aboutcsir/announcements/om_CSIR-CMMACS_210213.pdf samt "Bangalore to get India's fastest supercomputer" tillgängligt på <http://www.cmmacs.ernet.in/index.php/en/component/content/article/15-c-mmacs-in-the-news/113-bangalore-to-get-india's-fastest-supercomputer>. Se även *the Hindu*, "CSIR all set to launch fastest supercomputer", tillgängligt på <http://www.thehindu.com/todays-paper/tp-national/tp-karnataka/csir-all-set-to-launch-fastest-supercomputer-in-bangalore/article4883184.ece>

¹¹ NASSCOM, 2012, "Big Data – the Next Big Thing" sid 47 – 51, tillgänglig på <http://www.nasscom.in/sites/default/files/researchreports/softcopy/Big%20Data%20Report%202012.pdf>

ingått ett samarbete med *Indian School of Business* (ISB) vilket etablerar *ISB Biocon Certificate Programme in Business Analytics* (CBA), i syfte att utbilda statistiker med kompetens att hantera stora mängder data. Programmet har utvecklats av ISB i samarbete med företag som Bank of America, IT-företagen Cognizant Technical Solutions, IBM, Hewlett-Packard, MuSigma samt Tech Mahindra.¹²

Vid sidan av de initiativ som lanseras av regeringen vidtar naturligtvis Indiens IT-industri egna åtgärder för att utveckla kompetens och säkra kompetensförsörjningen framöver inom stora data. Att det för den globala stora data-marknaden råder enorma behov av kvalificerad arbetskraft råder det inga tvivel om. Konsultfirman Gartners bedömer exempelvis att det i hela världen sammantaget kommer att finnas en efterfrågan på kompetens inom stora data motsvarande 4,4 miljoner arbetstillfällen fram till år 2015. Av dessa möjliga arbetstillfällen kommer dock endast en tredjedel kunna tillsättas, i Indiens fall en något mindre andel, på grund av att den efterfrågade kompetensen helt enkelt inte finns.¹³ De stora indiska IT-företagen, som exempelvis Tata Consultancy Services (TCS), Infosys, Wipro och HCL säkerställer att de har den efterfrågade kompetensen inom stora data genom strategiska förvärv av mindre företag och samarbete med akademiska institutioner eller med andra indiska eller utländska företag. Genom sin globala närvaro kan de dessutom säkerställa att de finns fysiskt etablerade inom områden där den efterfrågade kompetensen också finns.

Som exempel kan TCS etablering av det företaget kallar *Silicon Valley Customer Collaboration Center* i Silicon Valley i USA nämnas. Centret ska fungera som företagets internationella bas när det gäller framväxande teknologier som just stora data och främja tillämpningen av denna nya teknik inom hela företaget.¹⁴ Ett annat exempel utgörs av Infosys som i februari i år meddelade att företaget ingått ett avtal med National ICT Australia (NICTA), de har etablerat ett samarbete som är tänkt att omfatta gemensam forskning, praktik för doktorander, yrkesmässiga utbyten, samt kommersialiseringen av immateriella rättigheter under de kommande fem åren. Samarbetet omfattar bland annat tillämpningen av stora data inom områden som logistik och smarta elnät (*smart grids*).¹⁵ Indiens tredje största IT-företag Wipro illustrerar strategin att förvärva utländska företag med en efterfrågad kompetens inom ett visst område. I maj i år investerade Wipro USD 30 miljoner (SEK 195 miljoner) en minoritetsandel i det amerikanska stora data-företaget Opera Solutions.¹⁶ Avslutningsvis kan nämnas HCL:s samarbete med det amerikanska läkemedelsföretaget Eli Lilly för att etablera ett gemensamt innovationslaboratorium i

¹² *Business Standard*, "ISB ties up with Biocon to offer business analytics course", publicerad 25 september 2013, tillgänglig på: http://www.business-standard.com/article/management/isb-ties-up-with-biocon-to-offer-business-analytics-course-113092500968_1.html senast besökt 20 september 2013

¹³ *Mint*, "Software firms face shortage of skilled engineers", publicerad 14 februari 2013, tillgänglig på <http://www.livemint.com/Specials/ZZnwdYctp9COE7MOIR50CL/Software-firms-face-shortage-of-skilledengineers.html> senast besökt 12 september 2013

¹⁴ *TCS pressmeddelande*, 31 januari 2012, "TCS Inaugurates Silicon Valley Customer Collaboration Center in Santa Clara, California", tillgängligt på http://www.tcs.com/news_events/press_releases/Pages/TCS_Silicon_Valley_Customer_Collaboration_Center_Santa_Clara_California.aspx

¹⁵ *NICTA pressmeddelande*, 13 februari 2013, "NICTA and Infosys ink collaborative research agreement", tillgängligt på http://www.nicta.com.au/_data/assets/pdf_file/0005/37805/Infosys_and_NICTA_joint_innovation.pdf

¹⁶ *Techcircle.in*, 8 maj 2013, "Wipro to invest \$30M in US-based Big Data firm Opera Solutions", tillgängligt på <http://techcircle.vccircle.com/2013/05/08/wipro-to-invest-30m-in-us-based-big-data-firm-opera-solutions/>

Singapore i syfte att främja nya teknologier och öka Eli Lillys internationella konkurrenskraft. Ett av samarbetsområdena avsåg just bearbetning och analys av data.¹⁷

2.3.2 Planerade policyer och strategier avseende stora data

Som nämntes inledningsvis strävar industrisammanslutningen NASSCOM även efter att främja framväxten av nya IT-företag. Som ett led i det arbetet offentliggjorde NASSCOM i mars i år ett särskilt program för att under 10 år stödja 10 000 uppstartsföretag inom den indiska IT-sektorn, såväl ekonomiskt som med rådgivning.¹⁸ De företag som väljs ut att delta i denna satsning får ett s.k. uppstartspaket till ett värde av USD 25 000. Till sin hjälp för att verkställa projektet har NASSCOM företag som Microsoft, Google och Amazon, samtliga tre partnerföretag i satsningen, samt inkubatorer och acceleratorer. Ambitionen är att skapa ett bestående ekosystem som underlättar för uppstartsföretag att etablera en bestående verksamhet. Den första ansökningsomgången för att delta i programmet avslutades den sista maj i år och enligt NASSCOM inkom det 4000 ansökningar. Av dessa är NASSCOM i färd med att välja ut 20–25 företag för stödinsatser. NASSCOM har inte i detalj redogjort för vilka företag som ansökt om att vara med i projektet, däremot har det meddelats att 11 procent av de inkomna ansökningarna avsåg företag verksamma inom stora data.¹⁹

NASSCOM har även tillsatt en särskild arbetsgrupp med uppdraget att utveckla tillväxtstrategier inom affärsområdet stora data, liksom att etablera Indien som global knutpunkt för stora data-tjänster. Gruppen ska även utveckla en strategi för att säkerställa den indiska IT-industrins kompetensförsörjning framöver. Arbetsgruppen består av representanter från det indiska näringslivet och universitetsväsendet.²⁰ NASSCOM har dock inte meddelat när några eventuella slutsatser från studien kan komma att presenteras offentligt. Att frågan om stora data fått ökad betydelse inom NASSCOM illustreras även av att organisation i juni i år för första gången arrangerade en stor konferens om stora data.²¹

När frågan om stora data diskuteras mot bakgrund av händelseutvecklingen i Indien är det viktigt att samtidigt ha etableringen av det indiska personnummersystemet Aadhaar (även kallat Unique ID (UID)) i åtanke. Aadhaar, som det är frivilligt att delta i, kan visserligen inte beskrivas som en policyåtgärd i syfte att främja stora data-tjänster, men som en direkt effekt av lanseringen av Aadhaar etableras nu en marknad för företag att hantera och analysera stora mängder information relaterad till tjänster kopplade till Aadhaar.

Den federala regeringen har beslutat att inrätta en särskild fond avsedd för uppstartsföretag som tillverkar mobilapplikationer som kan användas för att utnyttja olika webbaserade tjänster kopplade till Aadhaar. År 2014 beräknas 700 miljoner indier ha ett Aadhaar-nummer och regeringens beslut att inrätta en särskild fond för att stödja applikationsutvecklare vittnar om den enorma kommersiella potential som finns beträffande utveckling

¹⁷ HCL pressmeddelande, 14 juli 2011, "Lilly and HCL announce unique collaboration to bring new ideas to reality", tillgängligt på <http://www.hcltech.com/press-releases/healthcare/eli-lilly-and-hcl-announce-unique-collaboration-bring-new-ideas-reality>

¹⁸ För ytterligare detaljer se <http://www.google.com/entrepreneurs/initiatives/10000-startups.html> och <http://www.10000startups.com/web/guest/program-details>

¹⁹ Times of India, 22 juni 2013, "NASSCOM receives 4000 applications for 10,000 start-ups programme", tillgänglig på http://articles.timesofindia.indiatimes.com/2013-06-22/pune/40133512_1_startups-programme-start-ups-programme-angel-funding

²⁰ NASSCOM offentliggör skapandet av särskild arbetsgrupp för strategiutveckling och kompetensutveckling inom Big Data, tillgänglig på http://www.nasscom.in/sites/default/files/Article_News/PR_NASSCOM%20Big%20data%20%20analytics.pdf

²¹ "Big Data and Analytics Summit 2013", konferens arrangerad av NASSCOM i Hyderabad, 27 juni 2013

av applikationer utformade för detta system. Fondens totala storlek samt storleken på den ekonomiska stöd som är tänkt att utgå till ett enskilt företag är ännu inte fastställda. Regeringens målsättning sägs dock vara att kunna erbjuda särskilda applikationer med start i mars 2014.²² Tänkbara applikationer är sådana som kan användas för att verifiera ett Aadhaar-nummer, applikationer som kan upprätta ett födelsebevis eller dödsattest, samt applikationer som kan larma polis i en nödsituation. Fonden kommer att administreras av landets *Department of Information Technology* under minst tre år.²³

2.3.3 Policyer och strategier avseende öppna data

Arbetet med att offentliggöra myndighetsdata elektroniskt påbörjades 2012 då Indien inom ramen för sin Strategiska Dialog med USA, i samarbete med USA, etablerade webportalen *Open Government Platform* (www.opengovplatform.org) som idag länkar till Indiens nationella webportal för myndighetsdata, <http://data.gov.in>. Olika typer av data publiceras regelbundet. I dagsläget deltar 51 olika ministerier och departement och över 4000 informationsposter, eller dataset, har offentliggjorts.²⁴ Principen om att tillhandahålla offentlig myndighetsinformation fattades dock redan 2005 genom *Right to Information Act*.

Publiceringen på webportalen regleras av Indiens *National Data Sharing and Accessibility Policy (NDSAP)*, som antogs 2012.²⁵ Policyn slår fast att data ska klassificeras som delbar eller icke-delbar (konfidentiell) data. Delbar data klassificeras i sin tur som antingen fritt tillgänglig data, data tillgänglig för registrerade mottagare eller som data med begränsad tillgänglighet, vilket kräver regeringsbeslut innan data kan offentliggöras. Med NDSAP blir det obligatoriskt för ministerier och departement att offentliggöra icke-konfidentiell information på data.gov.in samt identifiera informationsposter som inte ska offentliggöras. Policyn föreskriver att data bearbetas internt för att säkerställa att kvalitetsstandarder uppnås innan publicering sker, dvs. riktigheten i informationen ska vara kontrollerad, liksom att det inte finns några juridiska hinder för publicering, samt att den personliga integriteten för en enskild individ inte äventyras och inte heller den nationella säkerheten.²⁶

Trots att policyn anger att varje enskilt departement själv får bestämma kriterierna för att skapa informationsposter av så hög kvalitet som möjligt gäller generellt att data ska göras tillgänglig enligt principerna att informationen ska vara så komplett som möjligt, vara av primärdatakaraktär, göras tillgänglig så fort som möjligt utan onödig fördröjning, vara maskinellt läsbar (handskrivna dokument eller pdf-dokument som inte möjliggör bearbetning av data är inte maskinellt läsbara), icke-diskriminerande, baserad på öppna standarder (*Commonly Owned Standards* – vilket möjliggör åtkomst till publicerad data utan krav på mjukvara som kräver dyra licenser), vara fritt tillgänglig utan begränsningar vad gäller

²² *Techcircle.in*, "Govt to fund mobile startups developing e-governance solutions", tillgänglig på: <http://techcircle.vccircle.com/2013/06/05/govt-to-fund-mobile-startups-developing-e-governance-solutions/>, publicerad 5 juni 2013, senast besökt 11 september 2013

²³ *Economic Times* "Boost for mobile tech ventures churning out e-governance solutions", tillgänglig på: http://articles.economicstimes.indiatimes.com/2013-06-05/news/39764396_1_government-websites-e-governance-government-contracts, publicerad 5 juni 2013, senast besökt 11 september 2013

²⁴ För en förteckning över de deltagande ministerierna och departementen samt de informationsposter som offentliggjorts se http://data.gov.in/agency-publications/agency-wise/catalog_type_raw_data, senast besökt 12 september 2013

²⁵ *National Data Sharing and Accessibility Policy – 2012*, tillgänglig på <http://ogpl.gov.in/NDSAP/NDSAP-30Jan2012.pdf>

²⁶ *Implementation Guidelines for NDSAP (April 2013)*, sid 7, tillgänglig på http://data.gov.in/sites/default/files/NDSAP_Implementation_Guidelines-2.1.pdf

möjlighet att sprida informationen vidare (fri/öppen licens), finnas tillgänglig över tid med förändringar i ursprungsdata tydligt markerad.²⁷

2.3.4 Planerade policyer och strategier avseende öppna data

Utöver de initiativ som tagits av den indiska regeringen och initiativ på delstatlig nivå har även den indiska IT-industrisammanslutningen NASSCOM engagerat sig i frågan om den kommersiella potentialen av att utnyttja fritt tillgänglig myndighetsdata (öppna data). Att frågan om öppna data fått ökad uppmärksamhet i Indien illustreras av att NASSCOM i samarbete med det statliga *National Informatics Centre (NIC)* i augusti i år arrangerade en konferens i syfte att främja utvecklandet av applikationer baserade på öppna data.²⁸ Via de sociala medierna Twitter, Facebook och Storify (<http://storify.com/DataPortalIndia>) publiceras även information från s.k. ”brain storming-möten”, av vilka flertalet ägt rum i år i huvudstaden New Delhi med omnejd och i Bangalore, den indiska IT-industrins högsäte. En stor del av diskussionerna berör givetvis de kommersiella möjligheterna för öppna data. Via konferenser riktade till allmänheten och via sociala medier delges även information om pågående tävlingar gällande applikationsutveckling baserad på öppna data. Som exempel kan nämnas ”#OpenDataApps Challenge”, en tävling organiserad av NASSCOM i samarbete med NIC.²⁹ Tävlingen riktar sig till indiska medborgare och av tävlingsreglerna framgår att applikationerna ska baseras på den data som görs tillgänglig via den nationella portalen data.gov.in. De applikationsutvecklare som deltar i tävlingen behåller rättigheterna till sina produkter och får utan begränsningar marknadsföra dessa tjänster för kommersiella ändamål. Applikationerna ska utvecklas enligt någon av de tre licensstyperna *Creative Common*, *MIT License* eller *GPL/LGPL*.³⁰

²⁷ *Implementation Guidelines for NDSAP (April 2013)*, sid 7, tillgänglig på http://data.gov.in/sites/default/files/NDSAP_Implementation_Guidelines-2.1.pdf

²⁸ Konferens på temat ”Open Data Apps for Innovation in Governance”, anordnad av *National Informatics Centre* i samarbete med NASSCOM, 8 augusti i New Delhi. För ytterligare information se <http://www.nasscom.in/nic-and-nasscom-organized-conference-%E2%80%9COpen-data-apps-innovation-governance%E2%80%9D-encourage-use-open-gover>

²⁹ För ytterligare information se <http://data.gov.in/appschallenge>

³⁰ <http://creativecommons.org/licenses/>, <http://opensource.org/licenses/MIT> samt <http://www.gnu.org/licenses/quick-guide-gplv3.html>

Innovationer och värdeskapande – två företagsexempel

Det indiska företaget **Mapunity** i Bangalore kan nämnas för att illustrera tjänsteutbud baserat på användandet av öppna data och tillämpningen av stora Data. Företaget är inriktat på att utveckla teknik för att lösa samhällsproblem och sociala utmaningar. Mapunity ledde utvecklingen av det s.k. Bangalore Transport Information System (BTIS) och samlar in data från flera olika källor – övervakningskameror uppsatta för att bevaka enskilda trafiksträckor, data från telefonoperatörer och från stadsbussar – i syfte att erbjuda allmänheten och trafikansvariga inom statsförvaltningen lägesaktuell trafikinformation. Företaget har efter framgången i Bangalore även utvecklat liknande plattformar i ett dussintal andra indiska städer. I en artikel i tidsskriften *The Economist* noteras att applikationer som erbjuder trafikinformation baserad på öppna data är den enda applikationstjänst som faktiskt nått en större konsumentkrets. Vanligare är istället att många applikationer baserade på öppna data erbjuder en ny tjänst men sällan får något större genomslag.*

Det New Delhibaserade företaget **Medysky** illustrerar potentialen för Big Data-tjänster inom sjukvårdssektorn. Företaget har utvecklat en webbaserad plattform som sammanför läkare, patient, sjukhus, testlaboratorium och apotek i syfte att möjliggöra en mer effektiv patientvård. De patientjournaler som företaget hanterar med information om sjukdoms- och behandlingshistorik är åtkomliga för behöriga läkare när som helst, var som helst. I detta sammanhang är det också relevant att uppmärksamma ovan nämnda *Centre for Development of Advanced Computing* som också nyligen utvecklat ett verktyg för journalhantering med hjälp av molnbaserad teknologi. Verktyget, kallat *Mercury Nimbus Suite*** möjliggör för användare att ladda upp (respektive ladda ner) datafiler innehållande bilder, ljud, video och dokument avseende dessa journaler. Genom denna typ av satsning på IT-infrastruktur kan regeringen sägas främja utvecklingen av Big Data-tjänster inom sjukvårdssektorn.

2.4 Framtida utmaningar

I den indiska debatten om hur kompetensförsörjningen inom stora data bäst kan säkras har frånvaron av ett omfattande samarbete mellan industri och akademiska institutioner lyfts fram som ett stort problem. Det har påtalats från akademiskt håll att det i Indien inte på långa vägar finns samma omfattande samarbete som det mellan industri och universitet som i exempelvis USA och Japan. Indiens *Ministry of Science and Technology* (genom de tre enheterna *Department of Science and Technology*, *Department of Biotechnology* och *Department of Industrial & Scientific Research*) har visserligen tagit initiativ för att främja denna typ av samarbeten, främst i form av att finansiera forskningsprojekt som involverar industrin och akademiska institutioner, men trots det har utvecklingen i denna riktning inte tagit fart. Stora industriföretag samarbetar visserligen med indiska elituniversitet som de s.k. *Indian Institutes of Technology* (IIT) och *Indian Institute of Science* (IISc) i Bangalore, men dessa lärosäten utgör bara en liten minoritet inom det indiska universitetsväsendet. Problemet är att den indiska industrin i stor omfattning misstror universitetens förmåga att leverera vid gemensamma projekt. Tekniska universitet lider av en omfattande brist på kvalificerad personal med rätt kompetens. Av det skälet saknas både intresse och förmåga från universitetens sida att ingå samarbeten med industrin.³¹ Den bristande kvaliteten inom landets högre utbildning är en fråga som uppmärksammats mycket i Indien det senaste året, och något som det kommer att ta lång tid att åtgärda.

När det gäller frågan om hur samarbeten mellan industri och högre lärosäten skulle kunna se ut diskuteras i dagsläget flera alternativ. Exempelvis skulle experter från industrin i

³¹ Korrespondens (130927) med M.H. Bala Subrahmanya, ordförande och professor vid *Department of Management Studies* vid *Indian Institute of Science, Bangalore*

* *The Economist*, "The multiplexed metropolis", 7 september 2013, tillgänglig på

<http://www.economist.com/news/briefing/21585002-enthusiasts-think-data-services-can-change-cities-century-much-electricity-senast-besökt-13-september-2013>

** För ytterligare detaljer se http://www.cdac.in/index.aspx?id=pk_pr_prs_rl215

samarbete med akademiska institutioner kunna svara för undervisningen i enskilda ämnen där de har expertkunskaper. Ett institutionaliserat samarbete där industrin bidrar till att utbilda fakultetsmedlemmar beskrivs som mycket brådskande, och en av de viktigaste åtgärderna, för att säkerställa att rätt kompetens finns inom fakulteter med ansvar för att svara för utbildning i stora data-teknologier vid indiska lärosäten.³² Ett tredje exempel på möjliga samarbeten mellan industri och den akademiska världen utgörs av studenters möjlighet till praktik. Vid *Indian Institute of Science* i Bangalore ingår praktik en termin som ett obligatoriskt inslag i utbildning på avancerad nivå (över kandidatnivå). *Indian Institute of Science* i Bangalore har även etablerat ett så kallat *External Registration Programme* som ger industriföreträdare och medlemmar från olika offentliga forskningsorganisationer möjlighet att doktorera vid lärosätet. Vikten av praktik och behovet av att indiska företag måste bli mer villiga att ta emot praktikanter och inte oro sig för att dessa ska ta med sig affärshemligheter, har också nämnts i den indiska debatten.³³

Andra utmaningar som har nämnts av företrädare för den indiska IT-industrin är att Indien inte svarar för några nya innovationer vad gäller nya teknikområden. Innovationerna inom stora data-industrin sker istället nästan uteslutande i USA vilket gör det svårt för Indien att hålla jämna steg. Om Indien verkligen vill bli en ledande aktör måste den innovativa förmågan förbättras.³⁴ I sammanhanget är det värt att nämna att behovet av att införa ett ökat fokus på kreativitet inom utbildningsväsendet är något som också uppmärksammats i Indiens nya femårsplan.

Behovet av att utveckla lagstiftningen för att skydda den personliga integriteten, är ytterligare en utmaning och en synpunkt som framfördes från flera håll i samband med den stora konferens som IT-industrisammanslutningen NASSCOM anordnade i juni i år.³⁵ När det gäller integritetsfrågor kan Indiens *Information Technology Rules, 2011*³⁶, nämnas. Som en del av denna lagstiftning finns vad som kallas *Security Practices Rules*, en kontroversiell regel, som ger indiska myndigheter rätt att begära ut personlig information från företag. Samtidigt ska det betonas att Indiens regering är i färd med att utarbeta ny lagstiftning avsedd av skydda den personliga integriteten.

I oktober 2012 presenterade en expertgrupp, tillsatt av Indiens Planeringskommission, sitt slutbetänkande inför det fortsatta arbetet med att utarbeta landets s.k. *Privacy Bill*³⁷, ett arbete som leds av *Department of Personnel and Training (DOPT)* vid *Ministry of Personnel, Public Grievances and Pensions*. Expertgruppen föreslog i sin slutrapport att

³² Synpunkt framförd av M.H. Bala Subrahmanya, ordförande och professor vid Department of Management Studies vid Indian Institute of Science, Bangalore och Prakash Durga Devarakonda, Enterprise Business Solutions, Social Media & Business Analytics Leader vid Mahindra Satyam & Tech Mahindra, i samband med NASSCOMs "Big Data and Analytics Summit 2013", Hyderabad, 27 juni 2013

³³ Synpunkt framförd av M.H. Bala Subrahmanya, ordförande och professor vid Department of Management Studies vid Indian Institute of Science, Bangalore och Aditya Khandekar, General Manager, Analytics Advisory Services, Fiserv India, i samband med NASSCOMs "Big Data and Analytics Summit 2013", Hyderabad, 27 juni 2013

³⁴ Synpunkt framförd av Arun Kharbanda, Capability Practice Leader, Research and Analytics Services, WNS, i samband med NASSCOMs "Big Data and Analytics Summit 2013", Hyderabad, 27 juni 2013

³⁵ Synpunkt framförd av Aveek Mukherjee, Managing Director, Wells Fargo och Kiran N. Cavale, Senior Vice Present and India Operations Lead-Business Information Management, Capgemini, i samband med NASSCOMs "Big Data and Analytics Summit 2013", Hyderabad, 27 juni 2013

³⁶ PRS Legislative Research, *The Information Technology Rules 2011*, tillgänglig på: <http://www.prsindia.org/billtrack/the-information-technology-rules-2011-1908/>

³⁷ Press Information Bureau, Government of India, 16 oktober 2012, "Group of Experts on Privacy Submit Report", tillgänglig på: <http://pib.nic.in/newsite/erelease.aspx?relid=88503> För hela rapporten, se http://planningcommission.nic.in/reports/genrep/rep_privacy.pdf

Indiens lagstiftning ska vara teknikneutral och överensstämma med internationella standarder. Rent konkret föreslås inrättandet av en så kallad integritetskommissionär (*Privacy Commissioner*) på federal och regional nivå samt ett system för gemensam reglering, med vilket avses ett system där organisationer (industrisammanslutningar exempelvis) ges möjlighet att själva utarbeta standarder för att värna den personliga integriteten, vilka måste godkännas av ovan nämnda integritetskommissionär. Expertgruppen rekommenderade även begränsningar vad gäller företags rätt att samla in personlig information och hur denna information sedan används. Ett system där individer ges möjlighet att godkänna vilken information som samlas in föreslås också.

Innan expertgruppens slutbetänkande offentliggjordes i oktober förra året hade DOPT spridit två utkast till ett lagförslag. DOPT har däremot inte presenterat något nytt utkast efter att gruppens rekommendationer blivit kända, vilket innebär att myndighetens arbete fortfarande pågår. Avslutningsvis kan nämnas att frågan på vilket sätt individers personliga integritet skyddas kommit att aktualiseras mycket som ett resultat av regeringens beslut att inrätta olika typer organ för informationsinhämtning i stor skala och analys av densamma. Som exempel kan nämnas *National Intelligence Grid (NATGRID)*³⁸ som lanserades i maj i år av *Ministry of Home Affairs* i samarbete med *National Institute of Smart Government*, samt inrättandet av vad som i Indien kallas *Centralised Monitoring System (CMS)*³⁹ som bland annat syftar till att bevaka landets mobil- och internettrafik och genomföra data extrahering (data mining). I dokument avseende etableringen av CMS nämns dessutom uttryckligen att en av programmets komponenter ska bestå av forskning och utveckling för att hela tiden utveckla CMS förmåga att verka.

³⁸ <http://mha.nic.in/pdfs/NATGRID-050613.pdf>

³⁹ <http://pib.nic.in/newsite/erelease.aspx?relid=54679>

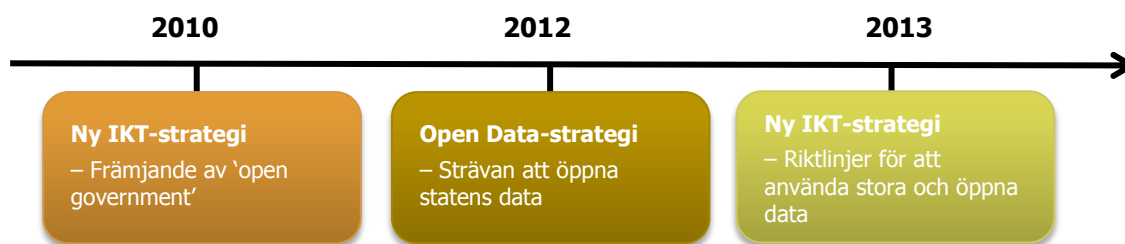
3 Japan

3.1 Sammanfattning

Japan släpar efter när det gäller stora data och öppna data jämfört med USA och Europa. Japanska statstjänstemän tenderar att tveka om att släppa information som samlas in av statsapparaten, delvis på grund av tradition och organisationskultur. Lagrummet för skydd av personuppgifter är föråldrat och gör strategier för stora data och öppna data svåra för såväl förvaltningen som näringslivet. Förutsättningarna avskräcker japanska företag från att lansera tjänster som hanterar stora och öppna data med inslag av personlig information, och de går därför kanske miste om de mest värdefulla möjligheterna. För närvarande är öppna data en gräsrotsrörelse i Japan, där företag under en lång tid har pressat regeringen att göra sina stora datamängder publikt tillgängliga. Det finns dock indikationer på att regeringen har börjat lyssna under senare år. Den inflytelserika industrilobbyn *Japan Business Federation* (Keidanren) har sedan 2012 specificerat förslag på hur staten bättre ska öppna upp data och allt fler ministrar har gett sitt stöd till förslagen.⁴⁰

3.2 Riktad policy och strategier

År 2010 presenterades den förutvarande IT-strategin (se Figur 1 för översikt), vilken inkluderade främjande av 'Open government' som en av tre pelare. Strategin var det första policyinitiativet inriktat mot öppna data i Japan – även om begreppet inte används explicit. Under 2012 lanserades en riktad strategi för öppna data, med syfte att uppmuntra den offentliga sektorn att öppna upp datauppgifter till gagn för företagens affärsskapande och innovation. I juli 2013 kommunicerades den nu liggande IT-strategin, som ännu mer uttryckligt inkluderar främjande av stora och öppna data som ett verktyg för att skapa ekonomisk tillväxt. Den nya strategin är mer detaljerad än tidigare strategier, och innehåller riktlinjer för utnyttjande av stora och öppna data.



Figur 1 Nationella IKT-strategier med bäring på stora och öppna data.

Källa: Tillväxtanalys

I Information- och Kommunikationsteknik-(IKT-)strategin från 2010 föreslås staten bli mer öppna med tydlig inspiration från utländska policytrender. Vid tidpunkten för utarbetandet av strategin besökte japanska tjänstemän länder som ansågs ligga i framkanten inom öppna data-policy, såsom Storbritannien och USA, för att se hur de arbetade med att tillgängliggöra sina datamängder. En arbetsgrupp inom *IT Strategy Headquarters* fortsatte de följande två åren att i detalj utvärdera och bli inspirerade av andra länders strategier.

⁴⁰Exempel återfinns här: www.keidanren.or.jp/en/policy/2012/079.html and www.keidanren.or.jp/en/policy/2013/019.html

Resultaten av arbetet publicerades i en riktad öppna data-strategi 2012 med rekommendationer för staten att göra data tillgänglig och offentlig.

Sedan policyskiftet år 2010 har en del data successivt börjat läggas upp på öppna offentliga webbplatser. Primära exempel rör data om väder, jordbävningar, naturresurser, kollektivtrafik, energiförbrukning, handel och affärer. Ingen av dessa datamängder härrör från eller har ens anonymiserade kopplingar till personlig information. De ansvariga departementen har bedömt att det är för riskabelt att öppna upp data med personlig information, även efter total anonymisering, det vill säga att man tar bort uppgifter som kan koppla data till individer eller grupper av individer. Det mest aktiva departementet för både öppnande av egna datamängder och främjande av en öppna data-policy är det japanska näringsdepartementet *Ministry of Economy Trade and Industry* (METI). METI har nyligen sjösatt sin portal för öppna data, *Open Data METI*, där departementet gradvis kommer offentliggöra datamängder de samlar in som del av sitt uppdrag och som är fritt för alla att ladda ner.⁴¹

En ny övergripande IT-strategi släpptes i juli 2013, med tydliga riktlinjer för hur man ska hantera och arbeta med öppna data. Strategin är i sin tur del av en av de nya pelarna i den totala ekonomiska tillväxtstrategin utarbetad av Japans nya regering som tillträdde vid årsskiftet. IKT beskrivs liksom i tidigare tillväxtstrategier som ett verktyg för ökad hållbar tillväxt inom alla sektorer. En nyhet är dock ett stort fokus på stora data och öppna data för att öka industrins innovationskraft och skapandet av nya arbetstillfällen.

Genom att dra nytta av stora datamängder, och genom att utnyttja insamlad data i nya kombinationer, finns förhoppningen att fler innovationer och företag ska skapas. Förutom framväxandet av nya och innovativa industrier och tjänsteföretag är målet att befintliga företag inom i princip varje sektor ska kunna få draghjälp av att statens dataarkiv öppnas. Ett mål för framtiden är att data med ursprung i (anonymiserade) personuppgifter ska inkluderas. Regeringen hoppas också att stora data och öppna data kommer hjälpa sektorer där IKT i allmänhet och dataanalys i synnerhet inte används i så stor utsträckning, till exempel inom jordbrukssektorn. I den nya IKT-strategin anges en rad utvärderingsvariabler för stora data och öppna data som är tänkta att användas för uppföljning senare. Konkreta mål saknas dock i dagsläget, men förväntas i mer konkreta handlingsplaner och riktlinjer baserade på IKT-strategin. De resultatindikatorer som diskuteras inbegriper hur mycket datamängder som varje departement öppnat upp och hur lättillgängliga de har blivit för utomstående företag och individer, antalet utvecklade applikationer som använder denna öppna data, hur många lagar som reviderats för att främja öppna data, och total omsättning för nya tjänster som företag tar fram baserat på öppna data.

Cabinet Secretariat har satt samman en arbetsgrupp av experter som är arkitekterna bakom riktlinjerna för stora data och öppna data. De kommer att fokusera på åtgärder i statsförvaltningen för att göra det enklare att öppna upp data, samt ge förslag på hur data kan komma näringslivet till godo. Riktlinjerna ska användas för att informera och instruera alla delar – departement, myndigheter, offentliga organisationer – i Japan. Det är främst fem dataområden som ska öppnas upp: 1) budgetarbete och upphandling, 2) förebyggande av katastrofer, 3) white papers, 4) geospatial information, och 5) turism och människors förflytningsmönster. Frågor som man uppfattar som problem kopplade till öppnandet av data, såsom upphovsrätts- och ansvarsfrågor, diskuteras nu aktivt i arbetsgruppen. Man har också diskuterat definitionen och tolkningen av begreppet öppna data för att kunna ge rekommendationer på eventuella avgränsningar av vad som kan öppnas upp utifrån det

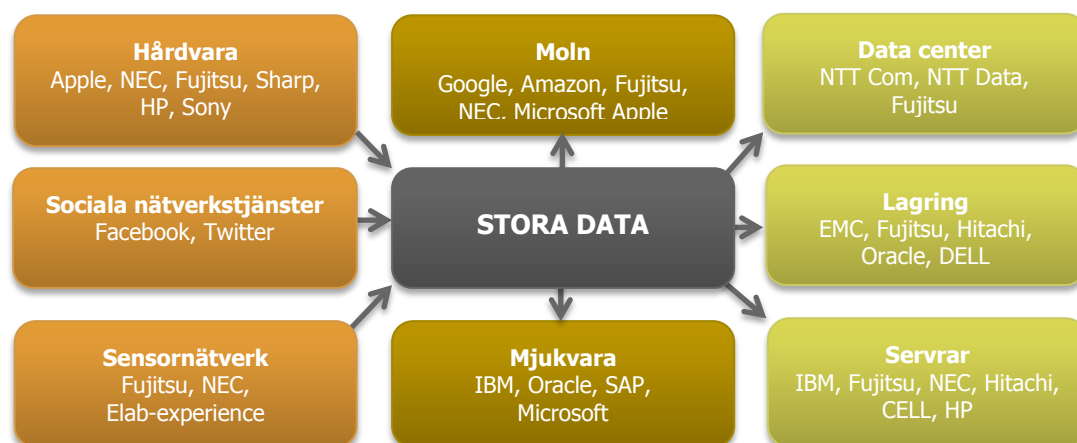
⁴¹ Se här för en betaversion av portalen: datameti.go.jp/data/

befintliga lagrummet. Huvudprincipen är att all data *inte* kan öppnas upp, snarare än det omvända, på grund av kränkningar av privata angelägenheter. Grundinställningen är också att all data som klassas som öppna även ska vara avgiftsfri, men redan tidigt i gruppens diskussioner verkar det finnas acceptans för undantag. Även om arbetet med datainsamling redan är finansierad via skattemedel kan i vissa särskilda fall öppnandet skapa ett stort administrativt merarbete vilket därmed skulle motivera särskilda brukaravgifter. I dagsläget är brukaravgifter de facto en lagstiftad del av vissa myndigheters finansiering. Exempel på sådan befintlig finansiering genom brukaravgifter återfinns till exempel för väderdata, kartdata, och företagsregisterdata. I stort ser arbetsgruppen dessa frågor som detaljer som ska lösas efter hand, och att man nu ska ”vidta åtgärder där åtgärder snabbt kan vidtas”.⁴²

3.3 Resultat inom innovation och värdeskapande

Enligt experter vi intervjuat är det i dagsläget svårt att hitta konkreta exempel på kommersiella framgångar i Japan som härrör från öppnandet av statens data. *Open Knowledge Foundation Japan* har listat 400 företag över hela världen med en affärsidé eller erbjudande baserat på öppna data. Endast sju av dessa är från Japan.⁴³ De japanska exemplen uppvisar inte några ekonomiska positiva resultat i sig utan är mer av typen interna affärsverktyg för att stödja försäljning eller andra applikationer.⁴⁴

Företagen har dock dragit nytta av möjligheterna inom stora data under åtskilliga år – se Figur 2 för en kartläggning av de större japanska aktörerna. Förutom större IT-företag spänner exemplen från mobiloperatörer, matbutiker, biltillverkare och tågoperatörer. Investeringar inom stora data förväntas i Japan växa med i genomsnitt 40 procent per år fram till 2016, då dessa beräknas uppgå till 5,2 miljarder kronor (77 miljarder JPY) per år.⁴⁵



Figur 2 Större företagsaktörer som använder stora data i Japan

⁴² Intervju med Mr. Hiroichi Kawashima, Chairman for the IT strategy working group of METI, June 18, 2013

⁴³ Data från Tomihiko Azuma, Open Knowledge Foundation Japan

<https://docs.google.com/spreadsheets/ccc?key=0AvXTTqI7i6p5dGx1MVQ4eG91VGZfZENmR0dcWizVKE#gid=0> March 07, 2013

⁴⁴ Intervju med Hiroichi Kawashima, PhD, Special Advisor to the government of Saga Prefecture (also a member of Open Knowledge Foundation), 18 juni, 2013

⁴⁵ The Japan Times <http://www.japantimes.co.jp/news/2013/06/18/reference/Big-data-a-digital-sea-of-personal-info-ripe-for-the-taking/> June 25, 2013

Följande tre fallstudier från tillverkningsindustri och hälso- och sjukvård, samt ett exempel på social innovation, är betydande för innovation och värdeskapande som inte skulle ha varit möjligt utan utvecklingen inom stora och öppna data.

3.3.1 Gruvindustri

Wenco International Mining Systems Ltd, ett dotterbolag till *Hitachi Construction Machinery*, levererar främst GPS-baserade IT-system för 'open pit' gruvdrift. *Wenco* använder även sensordata och analyser för förebyggande underhåll av gruvmaskiner. Maskinhaverier är kostsamma, och underhåll i tid kan betyda skillnaden mellan ett kortare driftstopp till större avbrott med hundratusentals dollar i förlorad produktion och ersättningskostnader för totalhavererad utrustning. För att undvika driftstopp arbetar *Wenco* tillsammans med *Hitachi* för att upptäcka fel i maskiner innan de uppstår genom att analysera stora mängder data i realtid.

Sensorer skickar i realtid maskindata och statistik till molnet, som sedan analyseras och resulterar i snabb reaktion på potentiella problem. Exempelvis kan en trög ventil och sänkt avgastemperatur vara tidiga signaler på ett kommande haveri. Varje enskild givare kan på kort tid samla in stora mängder ackumulerad data. Med hjälp av egenutvecklade algoritmer analyseras dessa stora datamängder i realtid för att hitta jämförelser och mönster som kan användas för att prognostisera maskinhaverier innan de är ett faktum. Systemet kan även omedelbart sända automatiserade instruktioner och förslag på åtgärder till maskinoperatörerna.

Optimerad drift och förebyggande underhåll är bara toppen av isberget när det gäller de potentiella tillämpningar av maskindata inom gruvindustrin. Gruvor är ofta belägna på geografiskt avlägsna platser. Det kan vara svårt att rekrytera tillräckligt med utbildade operatörer och gruvverksamheten är till sin natur riskabel för gruvarbetarna. Kostnaden för att bygga en nödvändig omgärdande infrastruktur för de mänskliga operatörerna är hög. Med hjälp av sensorbaserad övervakning och framsteg inom analys av stora data bygger *Wenco* grunden för mer automatisering av gruvdrift i framtiden.⁴⁶

3.3.2 Hälso- och sjukvård

Även om regeringen har rekommenderat vårdpersonal att i största möjliga utsträckning ska ersätta originalläkemedel med generiska läkemedel för att minska sjukvårdsutgifterna, är utvecklingen trög. Ett aktuellt projekt pågår där öppna data är en nyckelkomponent för att kommunicera fördelarna med generiska läkemedel till allmänheten och sätta ytterligare press på förskrivande läkare.

Kure City, Hiroshima prefektur, har en befolkning på cirka 244 000 personer. Nästan en tredjedel av befolkningen är 65 år eller äldre, vilket gör den till en av de städer som har den äldsta befolkningen i Japan. Med åldrande befolkning ökar sjukvårdskostnader per capita varje år vilket pressar stadens ekonomi. En övergång till billigare generiska läkemedel skulle lindra den ekonomiska belastningen avsevärt.

Under juli 2008 startade Kure City en informationstjänst baserad på data från hälsoförsäkringssystemet. 3000 medborgare som fått läkemedel utskrivna där pris-skillnaden mellan originalläkemedel och generiska läkemedel var som störst sållades ut baserat på individuella medicinska kostnadsuppgifter. I Japan betalar medborgare eller

⁴⁶ Data Centre Advisors <http://blogs.hds.com/hdsblog/2013/02/hitachi-machine-data-in-action-open-pit-data-mining.html> June 19, 2013

deras arbetsgivare 30 procent av vård- och läkemedelskostnaden. Staden skickade sedan information till patienterna om att det fanns en möjlighet att byta läkemedel, vilket skulle minska både deras privata och de kommunala sjukvårdskostnaderna. Informationsutskick i omgångar av 3 000 patienter fortsatte genom ett kontrakterat företag. Under perioden mellan juli 2008 och mars 2012 skickades information ut till över 18 000 patienter, och servicen har fortsatt sedan dess. Många kommuner kommer på besök för att lära mer.^{47 48} Resultaten från en uppföljning visar att cirka 80 procent av de patienter som erhållit informationen har växlat till generiska läkemedel, vilket kan översättas med årliga kostnadsbesparingar på 8,1 miljoner kronor (124 miljoner JPY) baserat på 2011 års förbrukning.⁴⁹

3.3.3 Social innovation

Open Knowledge Foundation Japan (OKFJ) är en ideell organisation som främjar öppna data i Japan. OKFJ är den japanska grenen av den globala organisationen *Open Knowledge Foundation (OKF)*, som är verksam i 50 länder. De har bidragit betydligt till utvecklingen mot mer öppna data i Japan genom att tillhandahålla kunskap och expertis. Stiftelsen har tre huvudsakliga mål; att främja öppna data genom att upplysa och utbilda allmänheten, att diskutera vilken data som bör vara öppna och hur detta sker, och verka för nationella strategier för öppna data. OKFJ:s medlemmar kommer främst från akademi och offentliga sektor, däribland representanter från *Ministry of Economy, Trade and Industry (METI)*, *Ministry of Internal Affairs and Communications (MIC)*, och regeringskansliet – the *Cabinet Office*. OKFJ organiserar projekt och evenemang, såsom ”Hackathons” dit de bjuder IT-ingenjörer och IT-ansvariga för att utveckla idéer och program från öppna data under intensiva 24-timmarspass.

Ett resultat från ett dylikt evenemang är den webbplats och tjänst som getts namnet ”Vad används mina pengar till?” som nu är i drift i Chiba City strax utanför Tokyo. Webbplatsen visar på ett pedagogiskt sätt hur den lokala regeringen i Chiba spenderar indrivna skattepengar. Tjänsten använder numera öppna detaljerad data från kommunala budgetar. Webbplatsen har blivit en stor succé och har till dags datum anammats av elva andra kommuner i Japan. Tjänstens syfte är att skapa intresse för det kommunala styret, skapa diskussion över vad kommunen ska satsa på, och generellt öka kommunikationen mellan lokala politiker och tjänstemän i förvaltningen och kommunernas medborgare genom ökad transparens.⁵⁰

3.4 Utmaningar för vidare utveckling

3.4.1 Lagstiftning

Tre huvudsakliga osäkerhetsfaktorer förhindrar utvecklingen för stora data och öppna data i Japan; föråldrad upphovsrättslag, ansvarsfrågor för publicerade data, och hantering av personuppgifter. För att helt lösa dessa frågor kommer lagändringar bli nödvändiga, men det finns inga direkta indikationer på att dessa nödvändiga lagförändringar kommer ske i närtid. Det främsta skälet som anges är trögheten i den japanska lagstiftningsprocessen. Förespråkare för öppna data i den nuvarande regeringen ser nu som sin huvuduppgift att börja öppna den data som inte påverkas av hindrande lagstiftning och att undersöka vilka

⁴⁷ http://www.city.kure.lg.jp/kureinfo/shisei090427_0500.html (in Japanese)

⁴⁸ http://www.mhlw.go.jp/bunya/iryoku/kouhatu-iyaku/dl/03_10.pdf (in Japanese)

⁴⁹ Intervju med Mr. Hiroichi Kawashima, chairman for the IT strategy working group of METI, 18 juni, 2013

⁵⁰ Intervju med Tomihiko Azuma, Chief Fellow, Institute for International Socio-Economic Studies (också medlem av Open Knowledge Foundation), 27 mars, 2013

temporära lösningar som finns för att kringgå ålderstigen lagstiftning innan den hunnit uppdateras. Detta sker samtidigt som man informerar allmänheten och förvaltningen generellt om vad öppna data innebär och verkar för lagändringar på lång sikt.

Ett exempel på hur man försöker hitta vägar framåt är introduktionen av användarlicenser. En lösning i Japan har varit att man gett stor vikt åt att göra klart vem som är ägare och ansvarig för en viss datamängd. Samtidigt är den japanska upphovsrättslagen (jämfört med vissa andra länder som Australien) tillämpbar på datamängder vilket lett till att den ansvarige tvekat att öppna upp data. För att kringgå lösningen och klargöra vem data ursprungligen tillhör ska ett licensförfarande introduceras. Licenser gör det möjligt att fritt släppa data till den privata och offentliga sektorn utan att återropa befintligt lagrum. En licens betyder inte att skaparen/den ansvarige för data hävdar upphovsrätt, utan i stället att denne formellt avkräver sig anspråk på datan i fråga, i vissa fall i kombination med begränsningar i användningen av data, såsom förbud av kommersiellt bruk. Licensieringen har även tydliga rutiner för hur källa ska bifogas till data även när den används i andra eller tredje led. Japan har här i stort följt en licensstruktur från *Creative Commons* för skapandet av licenser. *Creative Commons* är en ideell internationell organisation som möjliggör utbyte och användning av kreativitet och kunskap genom att gratis erbjuda juridiska verktyg. Vad som återstår att diskutera är hur till exempel myndigheter ska arbeta operativt med olika typer av licenser.

Som synes påverkar ett licensförfarande även den andra osäkerhetsfaktorn positivt, nämligen frågan om vem som bär ansvar för data efter att de släppts öppna. Det finns en klar oro bland japanska förvaltningstjänstemän över att data kan komma att missbrukas, eller att data innehåller felaktigheter. I de nya riktlinjerna som är på väg att tas fram av arbetsgruppen ska kvaliteten på data helt separeras från frågan om upphovsrätt. De myndigheter som öppnar data ska inte vara ansvariga för hur dessa data sedan används. Det återstår att se om dessa riktlinjer kommer bidra till förändring i organisationskultur i den japanska förvaltningen som behövs för att påskynda utvecklingen av öppna data.

Gällande den tredje faktorn, hantering av personuppgifter, måste de japanska lagarna ses som mycket restriktiva för vilka uppgifter som får lagras om en individ i syfte att skydda den personliga integriteten. I princip kan till exempel inga data på kunder registreras och lagras utan skriftligt samtycke. Regeringens råd, *Council for Regulatory Reform*, startade under juni 2013 en utredning med syfte att ta fram och föreslå nya riktlinjer för hur data ska kunna anonymiseras helt för att sedan öppnas upp, utan att bryta mot befintlig personuppgiftslag. Rådet ska delge regeringen resultatet av denna utredning före årsskiftet.

3.4.2 Främjande

Barriärerna och skillnaderna i organisationskulturerna mellan olika departement och myndigheter kan i Japan beskrivas som höga, vilket bidrar till stora skillnader mellan departementens intresse att främja och implementera nya riktlinjer för öppna data. *Ministry of Economy Trade and Industry* (METI), *Ministry of Internal Affairs and Communications* (MIC), och *Ministry of Education, Culture, Sports, Science and Technology* (MEXT) är de departement som arbetar aktivt för att förbättra möjligheterna för och även själva utnyttja stora data och öppna data. Exempelvis har *Ministry of Healthcare* och *Ministry of Transport* visat sig tveksamma till stora och öppna data och snarare motverkar än främjar utvecklingen.

Precis som när det gäller främjande av andra IT-relaterade frågor, såsom e-förvaltning, ges de i statsadministrationen inbäddade *Government CIO* (*Chief Information Officers*) en

central roll för att främja öppna data. Varje japanskt departement och myndighet har en särskild *Government CIO* bland sin personal. Denne ansvarar för att främja och ge råd om införande av nya IT-lösningar inom organisationen. Sedan oktober 2012 finns också en central *Government CIO* som övervakar den totala informationsstrategin för hela förvaltningen. *Government CIO* hade tills en lagändring från juni 2013 ingen explicit lagstadgad auktoritet utan kunde endast agera som rådgivande. Förstärkningen av rollen för *Government CIO* sägs syfta till att försöka bryta de beskrivna barriärerna mellan departementen och att bättre samordna IT-strategierna inom förvaltningen.

Baserat på uttryckliga åsikter från representanter från METI har inga allvarliga ansträngningar gjorts för att göra offentliga departement, organ och företag mer medvetna om möjligheterna med stora och öppna data eller skapa uppmuntran om ökad användning. Trots de nya strategierna är inställningen att den offentliga sektorn inte har ett ansvar att påvisa möjligheterna med öppna data för näringslivet – möjligtvis med undantag från viss uppmuntran av småföretags experimenterande i begränsad skala.⁵¹

3.4.3 Standardisering

Vikten av standardisering inom stora och öppna data erkänns av den japanska regeringen, men ännu har inte några beslut tagits vad gäller specificerade standardformat för publicering av data från förvaltningen. En arbetsgrupp inom METI diskuterar för närvarande olika standarder, först som gäller METI och som sedan förhoppningsvis kan anammas av andra departement och myndigheter. Arbete beskrivs som att vara i en uppstartsfas. Ett exempel på *de facto* standardisering är arbetet med licensiering som tidigare nämnts. *Creative Commons* standarder används av METI i deras portal *Open Data METI*.

De vanligaste dataformat som används vid publiceringen av data från statsförvaltningen är csv (*Comma-Separated Values*), xls (Excel) och pdf. API (*Application Programming Interface*) ses internationellt som det mest fördelaktiga för att publicerade offentliga uppgifter, men inga exempel har hittills setts i Japan. Det finns heller inget mål att formellt anamma en internationell standard av typen API, då det för tillfället finns fler trängande interna frågor för Japan att lösa innan de börjar tänka internationellt.⁵²

3.4.4 Forskning och utveckling

Forskning och utveckling inom stora data i Japan sker företrädesvis vid en handfull större företag och inte vid universiteten. Det finns endast ett fåtal större akademiska forskningsprojekt som har fokuserat på utmaningar och möjligheter med stora data. År 2007 finansierade METI projektet *Information Grand Voyage Project* med en miljard kronor (15 miljarder JPY) under tre år där begreppet ”informationsexplosion” användes i stället för stora data men med liknande innebörd. Projektet fokuserade på information på internet, men även stora datamängder i verkliga världen, till exempel, fysisk distribution, transport, hälsovård, digitalt innehåll, och tillverkning. Efter regeringsskiftet år 2009 har projektet har inte förlängts ytterligare, och det finns få synliga resultat från projektet.⁵³ Ett annat angränsande och pågående projekt leds sedan 2009 av Dr Kitsuregawa vid University of Tokyo, som också var en av de ledande medlemmarna i METI *Information Grand Voyage*

⁵¹ Intervju med Mr. Hiroichi Kawashima, ordförande IT strategy working group of METI, 18 juni, 2013

⁵² Intervju med Mr. Tomihiko Azuma, software engineer at Institute for International Socio-Economic Studies, 27 mars, 2013

⁵³ www.meti.go.jp/policy/it_policy/daikoukai/igvp/index_en/project-outline/000about-the-information-grand/index.html

projektet. Trots att även detta projekt varit nära att läggas ner men slutligen fortsatt efter budgetnedskärningar, finns stora förhoppningar på ett genombrott inom utveckling av databas som ska klara av betydligt större datamängder än idag. Dr Kitsuregawas mål är att förbättra hastighetsprestanda i databashantering med en faktor på 1000 gånger.⁵⁴

Som kanske det främsta exemplet på stora data FoU inom den privata sektorn satsar Rakuten – den japanska versionen av Amazon – stora resurser på att bättre kunna hantera och analysera de datamängder som strömmar igenom deras servrar. Rakuten sysselsätter cirka 9 000 personer och av dessa är 2 000 dataingenjörer och tekniker. De utvecklar i allmänhet programvara själva. Liksom för Amazon.com bygger cirka 30 procent av omsättningen på rekommendationer från befintliga användare. Det krävs av Rakuten att i realtid kunna matcha ihop utbud och efterfrågan för varor och tjänster med algoritmer anpassade för stora mängder icke-strukturerade data där naturligt språk och slang ökar komplexiteten.⁵⁵

3.4.5 Brist på experter

Japanska företag har redan signalerat om en brist på kompetens inom stora data och bitt regeringen att ta detta på allvar. Åtgärder från staten för att förhindra denna brist har dock väntat på sig. Det är i stället företag som Fujitsu, NEC och Hitachi, som har börjat erbjuda egen kursverksamhet relaterad till stora data öppna för allmänheten att delta i, samt utökat stöd till universiteten för att skapa nya kurser. Hitachi inrättade 1 juni 2013 ett internt globalt centrum för *Innovative Analytics* för att stärka det globala samarbetet mellan sina analytiker, systemingenjörer och säljare i Japan, USA, Europa och på andra håll i Asien. Hitachi har satt som mål att öka antalet specialister inom stora data från dagens 300 till 500 innan år 2015.

Det finns idag inga högskolekurser eller program som fokuserar på stora och öppna data. Vid University of Aizu, ett universitet specialiserat datavetenskap, rapporteras vissa föreläsningar med externa föreläsare från till exempel NEC och Fujitsu som introducerar de specifika problemen och möjligheterna med stora data.⁵⁶ En handfull forskningsprojekt på universitet med teman som rör stora data har påvisats. Ett exempel är University of Tokyo, som har slagit sig ihop med 13 privata företag för att analysera stora datamängder i syfte att skapa nya produkter och tjänster.⁵⁷

⁵⁴ <http://first-pg.jp/english/en-about-us/kitsuregawa-masaru.html>

⁵⁵ Presentation av Mr. Masaya Mori, Executive Officer, Rakuten vid Open Source Software seminar 26 mars, , 2013

⁵⁶ University of Aiz Hemsida, <http://www.u-aizu.ac.jp/e-index.html>. Project blog <http://blogs.itmedia.co.jp/business20/2012/10/post-04fe.html> (japanska)

⁵⁷ Nikkei.com, <http://e.nikkei.com/e/ac/TNKS/Nni20130508D0805A11.htm?NS-query=%20Firms%20To%20Mine%20Big%20Data>, May 9, 2013

4 Storbritannien

4.1 Sammanfattning

I Storbritannien har regeringen tagit en hög profil kring frågor som rör e-förvaltning och öppna data. I slutet av 2012 antogs digitala strategier, såväl för det brittiska regeringskansliet i dess helhet som för enskilda departement. I en lång rad initiativ har regeringen motiverat att satsningar på en bättre och smartare användning av IT i förvaltningen förväntas ge lägre kostnader, bättre tjänster för medborgarna och stimulera innovation och tillväxt för brittiska företag i informationsekonomin.

Regeringen beskriver Storbritannien som världsledande i att öppna tillgången till offentliga data. En samlad portal (data.gov.uk) gör offentliga dataset lätt tillgängliga, liksom externt utvecklade tillämpningar som använder sig av dessa dataset. *Open Government Partnership (OGP)* är en multilateral organisation som syftar till att få stater att förbinda sig att genom ny teknik främja transparens och förbättra förvaltningen. Under 2013 har Storbritannien tillsammans med Brasilien delat på ordförandeskapet för organisationen, och dess årliga toppmöte hölls i London. Under 2013 har Storbritannien också som ordförande i G8 särskilt lyft fram förslag om att öka tillgängligheten för offentliga data och att de ska publiceras i öppna, standardiserade och användbara format.

4.2 Selektiv politik och strategier

I juni 2012 publicerade den brittiska regeringen sin vitbok om öppna data med titeln *Unleashing the Potential*.⁵⁸ I vitboken konstaterar regeringen att ”data är 2000-talets nya råvara. Dess värde ligger i att möjliggöra ansvarsutkrävande av regeringen, driva på valmöjligheter och förbättringar av offentliga tjänster samt att uppmuntra innovation och företagande som driver på social och ekonomisk tillväxt.”

Strategin som presenteras i vitboken bygger på tre breda målsättningar:

1. Förbättra tillgängligheten för offentliga data

Den brittiska *Freedom of Information Act* gör det obligatoriskt för myndigheter att offentliggöra data, och alla brittiska departement har redan publicerat ”öppna datastrategier” som innehåller utfästelser att offentliggöra ytterligare data. Vitboken fastställer behovet av att stärka medborgarnas rätt till data och att fastställa standarder för att göra data användbara.

2. Förbättra tilliten till offentliga data

Denna målsättning handlar primärt om integritetsfrågor och betonar vikten av att dataset förblir anonymiserade och att personuppgifter förblir personliga.

3. Använd offentliga data smartare

Genom att förändra kulturen inom den offentliga sektorn till att dela data mellan myndigheter och offentliga aktörer kan de datamängder som de offentliga aktörerna innehar användas för mer skräddarsydda och effektiva offentliga tjänster.

⁵⁸ <https://www.gov.uk/government/publications/open-data-white-paper-unleashing-the-potential>

Ett av de konkreta resultaten av vitboken var att tillsätta en oberoende översyn av offentliga data i Storbritannien. Utredningen presenterade sin rapport i maj 2013.⁵⁹ Den 14 juni publicerade regeringen sitt svar på utredningen.⁶⁰ I sitt svar utvecklar regeringen ytterligare vad som ska vara en nationell datastrategi.

En utgångspunkt för offentliga data ska vara att data alltid ska göras tillgängliga där det representerar ett mervärde för skattebetalarna (se Figur 3 för en översikt över datakategorier och dess tillgänglighet), om det inte föreligger legala eller säkerhetsmässiga skäl som talar emot det. I de relativt få fall där öppna publikation av data inte är möjligt kan det, enligt rapporten, finnas goda skäl för att göra det möjligt att dela data i alla fall, under strikta säkerhetskontroller.

Ett exempel på hur data kan delas under begränsade former rör skoldata med resultat på individnivå. Sådana data delas för forskningsändamål med särskilt ackrediterade universitet och forskningsinstitutioner i särskilt etablerade datalab- eller sandboxmiljöer. I vissa fall görs bedömningen att data inte kan delas utanför förvaltningen, men att datadelning är möjlig under strikta säkerhetsbegränsningar mellan departement och myndigheter.



Figur 3 Översikt kategorier av data och dess tillgänglighet

När det gäller fullständighet och kvalitet i offentliga data säger rapporten att man ”inte ska låta det bästa bli det godas fiende”, utan att man ska definiera vissa dataset som en *high quality core* där man koncentrerar ansträngningarna på datasetens kvalitet. För övriga data ska grundprincipen vara att först publicera och sedan gradvis förbättra.

I rapporten utlovas också en förändring av den brittiska *Freedom of Information Act* som ger medborgarna en ”rätt till data”. Därmed ska det bli obligatoriskt för myndigheter att göra sådana data som ska offentliggöras enligt lagen tillgängliga i maskinläsbart format. I samband med att EU:s PSI-direktiv implementeras i brittisk lagstiftning under 2014-15 kommer detta också att innebära att den största delen av offentliga data görs tillgängliga för andrahandsanvändning. Utgångspunkten är att detta ska ske till marginalkostnadspriser, vilket för data i maskinläsbara format torde vara gratis.

⁵⁹ <https://www.gov.uk/government/publications/shakespeare-review-of-public-sector-information>
Tillväxtanalys rapporterade om utredningen i kvartalsrapporten för andra kvartalet 2013, se: <http://www.tillvaxtanalys.se/download/18.7bce3c6e13f3cea9a771d5/1371713635532/Q2+IKT.pdf>

⁶⁰ <https://www.gov.uk/government/publications/government-response-to-shakespeare-review>

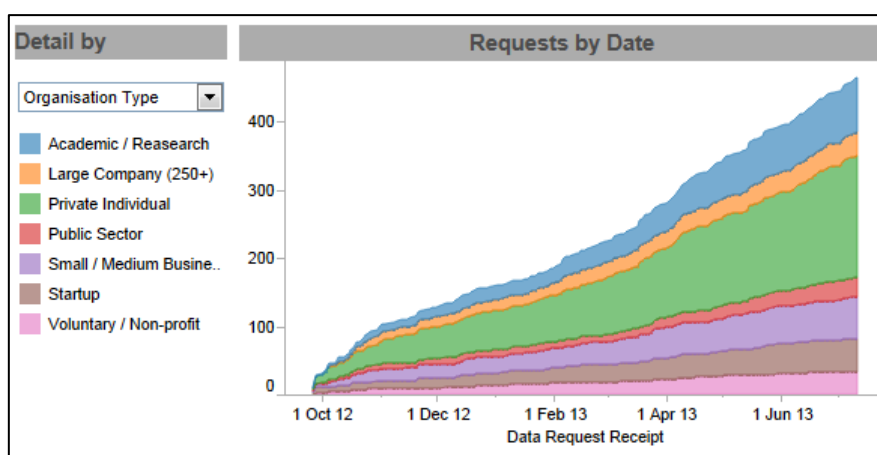
En viktig del av den nationella datastrategin är en (samtidigt publicerad) strategi för informationsekonomin, som är resultatet av ett samarbete mellan regeringen, näringslivet och den akademiska sektorn i ett råd – *the Information Economy Council*.⁶¹

I denna strategi framhålls det stora behovet av kompetens inom IT-området, både hos arbetskraften och bland befolkningen i stort. I strategin konstateras vidare att 95 procent av de cirka 120 000 företagen i den brittiska IT-sektorn har färre än tio anställda. En central del av strategin är därför att inrikta sig på små och medelstora företag för att göra Storbritannien till en attraktiv plats att starta teknikföretag och få dem att växa. Informationsekonomirådet ska bland annat arbeta tillsammans med det brittiska patentverket för att ge riktat stöd till små och medelstora företag i sektorn för att få maximal utväxling av sina immateriella rättigheter.

4.2.1 Lagstiftning och översyn

De olika strategier som regeringen har antagit följs åt av ett antal åtgärder, lagändringar och inrättandet av olika grupper och organ som ska understödja utvecklingen av öppna data i Storbritannien. Exempel är följande:

- *Open Data Institute* inrättades i december 2012 för att främja utvecklingen av en ”öppen data-kultur”. Institutet ska samla experter och fungera som en inkubator för nya företag som bygger på öppna data och sprida kunskap. Institutet finansieras med 10 miljoner pund över en femårsperiod.
- Förändringar i *Protection of Freedoms Act* innebär att sådan information som innefattas i *Freedom of Information Act* ska göras tillgänglig i maskinläsbart format.
- En *Open Data User Group* har inrättats för att driva på processen med att göra offentliga data tillgängliga. Genom portalen *data.gov.uk* kan användare begära att offentliga data ska göras tillgängliga och följa hur deras begäran behandlas.
- Alla departement har publicerat öppna data-strategier. Dessa följs kontinuerligt upp av *Cabinet Office* och utvecklingen offentliggörs i *Ministerial Statements*.



Figur 4 Begäran om tillgängliggörande av dataset.

⁶¹ <https://www.gov.uk/government/publications/information-economy-strategy>

4.3 Exempel på användning inom innovation och värdeskapande

På den officiella websidan för brittiska offentliga data, data.gov.uk, finns idag över 10 000 olika dataset tillgängliga för användare att antingen direkt ta del av uppgifterna eller för att använda dataseten för att utveckla nya tjänster och tillämpningar.

Data.gov.uk fungerar också som en portal för tillämpningar och appar som baseras på de data som publiceras där. En lång rad olika appar med exempelvis väderdata, lokaliserad statistik över trafikolyckor eller lokalisering av skolor baserat på deras resultat i skolmyndigheten *Ofsted*'s utvärderingar finns att ladda ned gratis eller köpa. Applikationerna bygger alltså på data som görs tillgängliga gratis, men där utvecklarna kan ta betalt för sina appar.

Exempel: Public Health England

Public Health England är en myndighet under *Department of Health* som samlar och analyserar information om hälso- och sjukvårdsfrågor i England. Genom myndighetens webbplats kan vårdsektorn, lokalpolitiker eller allmänheten få tillgång till geografiskt nedbrutna data om exempelvis sjukvård och folkhälsofaktorer.

Genom detaljerade jämförelser med andra områden eller med genomsnittsvärden för England kan man bättre prioritera exempelvis inriktning eller reformer i vården eller insatser i folkhälsoarbetet.



Figur 5 Exempel på jämförelsedata från Public Health England.

Exempel: BP

Ett exempel på ett brittiskt företag som mer systematiskt börjat använda stora data i sin produktion är oljebolaget BP. BP använder maskingenererade data från utrustningen på sina borrhplattformar för att förbättra underhåll och service. Genom att löpande analysera data som till exempel temperatur, vibrationer, tryck i oljeflöden kan man tidigt identifiera störningar i driften och förebygga driftavbrott. I en stor industri där själva utvinningen utgör en flaskhals i produktionen innebär ett förebyggt driftavbrott mycket stora ekonomiska besparingar.

BP som driver oljeriggarna i Nordsjön, Mexikanska golfen, Kaspiska havet och utanför Angola har centraliserat övervakningen av driftdata från utrustningen till sitt *Advanced Collaborative Environment Centre* i Aberdeen, Skottland. Olje- och gasriggarna ute till havs är krävande och svårtillgängliga arbetsmiljöer, vilket gör det förmånligt på många sätt att kunna sköta övervakningen på distans. Fjärrövervakningen innebär också att man kan samla expertis på ett ställe, som därigenom blir tillgänglig för alla BP:s produktionsanläggningar.

I och med att "lättillgänglig olja" blir mer och mer av en bristvara etableras utvinningsanläggningar på alltmer krävande platser – till exempel på djup havsbottnar. Utvecklingen ställer allt högre krav på utrustning och arbetsmetoder, och här ser BP avancerad IT-användning, däribland stora data-tillämpningar, som en förutsättning för fortsatt utveckling.

4.4 Utmaningar för fortsatt utveckling

Den brittiska regeringen motiverar behovet av en *Information Economy Strategy* med att utvecklingen på området skär tvärs över flera samhällssektorer. Informationsekonomin är mer av ett ekosystem än en separat sektor. Även om man kan förvänta sig att det mesta av värdeskapandet ska ske inom näringslivet så innebär inte det att staten kan luta sig tillbaka och invänta en positiv utveckling. Staten och näringslivet har ett ömsesidigt beroende. Utan långsiktig planering för att motverka kompetensbrist kommer företagen att ha svårt att rekrytera rätt arbetskraft. Utan rätt infrastruktur, såväl fysisk som virtuell, kommer företagen att få svårt att utvecklas. Utan god IT-säkerhet kommer företag och konsumenter inte att ha tilliten att använda ny teknik och nya tjänster. Om staten inte adresserar marknadsmisslyckanden kommer utvecklingen av informationsekonomin att hämmas.

För att stärka kompetensförsörjningen ska regeringen tillsammans med näringsliv och utbildningssektorn ta fram ett riktat *computing curriculum* till skolornas läroplaner, som ska börja tillämpas i september 2014. Informationsekonomin tar också upp de möjligheter man ser med storskaliga elektroniska utbildningar (*Massive Open Online Courses – MOOC*). Näringslivet och olika utbildningsorganisationer ska enligt strategin tillsammans ta fram ett program med MOOC-kurser för datavetenskaplig utbildning. Företagen i IT-sektorn har bildat ett gemensamt sektorsråd, *e-skills UK*, vars uppgift är att verka för långsiktig kompetensförsörjning inom IT-området. *E-skills UK* arbetar också tillsammans med utbildningssektorn för att ta fram en tillämpad utbildning – *ICT Higher Apprenticeship* som ska kunna ge en akademisk examen.

5 Sydkorea

5.1 Sammanfattning

Sydkorea har en avancerad IT-infrastruktur och ett stort antal IT-användare, som i sig genererar enorma mängder data. Implementering av stora data och öppna data-strategier är i en uppstartsfas, även om utvecklingen går snabbt framåt sedan förra året. Det finns uppskattningsvis omkring tio företag i Sydkorea som på allvar implementerar en stora data-strategi. Både regering och företag blir dock allt mer medvetna om att stora data kan bli en viktig komponent för den nationella konkurrenskraften, särskilt inom IT-industrin. Den sydkoreanska regeringen som tillträdde i årsskiftet har satt upp stora data som prioriterad strategi med stora förhoppningar om att det ska stimulera utvecklingen av nya produkter och tjänster och i slutändan bidra till att skapa arbetstillfällen. Två skäl som talar för stora data som tillväxtmotor är att det krävs ett relativt litet startkapital för att realisera tjänster inom stora data, samt att området framförallt skulle gynna den yngre generationen med kreativa idéer som i dagsläget kämpar för att få anställning. Regeringen har vidtagit åtgärder för att främja infrastrukturen för stora data i olika sammanhang, som exempelvis verkat för offentlig datahantering, genom lagstiftning och satsningar på teknik och strategi samt humankapital.

5.2 Inledning

Den nya regering som tog över i Sydkorea i vintras har känt av såväl yttre som inre tryck på att åstadkomma förändringar inom en rad områden. En långsam återhämtning av den globala ekonomin och svag inhemsk efterfrågan har saktat ner Sydkoreas viktigaste ekonomiska tillväxtmotorer inom bland annat tung industri. Regeringen har väljarnas krav på sig att snabbt identifiera lovande nya tillväxtområden genom att ta tillvara såväl befintligt som nytt kunnande. Stora data har under innevarande kvartal lyfts fram som det första ”nationella projekt” som ska säkra Sydkoreas tillväxt.

Den IT-industri som länge skapat stora värden för Sydkorea kommer att fortsätta spela en viktig roll i landets ekonomi. Förutom några lysande undantag anses industrin dock till stora delar ha stagnerat. Man har blivit alltför beroende av tidigare framgångar och sägs sakna starka visioner inför framtiden. Regeringen har nu bett IT-industrin att göra gemensam sak med staten och bygga upp ett industriellt ekosystem för att samla och odla nya resurser inom just stora data. Regeringen hoppas på betydande resultat i form av nya tjänster från befintliga IT-jättar, liksom så småningom från nya framgångsrika företag.

Regeringens strategi får stöd av aktuella uppgifter från *Korea Information, Science and Technology Institute* (KISTI). De publicerade i april sin årliga status- och trendrapport med en prognos där marknaden för stora data i Sydkorea väntas öka kontinuerligt från 0,6 procent år 2013 till 2,6 procent år 2020, i andel av den totala inhemska IT-marknaden. Det motsvarar en marknad på knappt sex miljarder kronor (894 miljoner USD) år 2020. Rapporten betonar särskilt vikten av industripolitiskt stöd och av innovationer inom små och medelstora företag. Den pekar också på behov av nära samarbeten mellan de mindre företagen och de stora IT-jättarna, till exempel i forum som *Big Data Solution Forum* (BIGSF).

Redan i november 2012, strax före regeringsskiftet, lanserades strategidokumentet *Big Data Master Plan to Realize a Smart Nation*, med förslag på åtgärder. Den nya regeringen bekräftar nu att åtgärderna ska genomföras, och pekar särskilt på behovet av att stärka

användningen av stora data i alla delar av förvaltningen. Staten ska vara proaktiv inom policyutformning och stödja industrin i att ta fram nya applikationer och plattformar. Det skulle till exempel kunna handla om kommande större beställningar av stora data-tjänster till sydkoreanska företag från den offentliga sektorn, liksom lagändringar och myndighetsinstruktioner för att främja användning. För att säkra upp dessa *private public partnerships* från statens sida har en stödbudget antagits fram till år 2016 om knappt 3 miljarder kronor (500 miljarder KRW) enbart avsedd att användas inom stora data.

Regeringen planerar att inrätta ett nytt stort datacenter för att hjälpa de inhemska företagen att komma ikapp och bli ledande internationellt inom stora data. Det kommer att bli Sydkoreas första centrum som tillåter vem som helst att förfina och analysera stora datamängder. Det i frågan drivande departementet *Ministry of Science, ICT and Future Planning* (MSIP) kommer att delegera centrumets uppbyggnad till myndigheten *National Information Society Agency* (NIA). Enligt MSIP anses sydkoreanska företag ligga två till fem år efter globala teknikföretag som Google och Amazon, och sakna ordentlig infrastruktur och erfarenhet för att lyckas inom stora data. Ambitionen är att det nya centrumet ska bli en testbädd för företag med aktiviteter inom stora data, samt ett sätt att främja forskning inom området vid universiteten. Centrumet ska utformas för att små och medelstora företag, nystartade företag, universitetsforskare, och även enskilda medborgare ska kunna testa sina idéer eller forskningsfrågor inom stora data via ett webgränssnitt. NIA kommer att ansluta sin online-plattform till testbädden, med diverse nationell statistik och annan information för att främja öppenheten. Centrumet har redan i september haft en första testkörning med externa parter.

Förutom nya affärsmöjligheter ska nämnas att stora data även är av intresse för statsförvaltningens interna arbete, exempelvis för att få bättre beslutsunderlag. Ett aktuellt exempel kommer från *Korea Institute for Health and Social Affairs* (KIHASA), ett statligt institut direkt under *Prime Minister's Office*. KIHASA startade i maj ett projekt för att kunna förutse och tidigt stävja mobbing genom analys av stora mängder data från flera populära tjänster för sociala media. I projektet analyseras historiska data från 214 populära websidor och tjänster sedan tre år tillbaka, med syfte att ta fram algoritmer som kan ge tidiga varningssignaler på mobbing eller självmordstendenser bland användarna.

5.3 Riktad policy och strategier

Sydkoreas regering implementerar särskilda strategier för stora och öppna data av samma skäl som många andra länder: förutom affärsmöjligheter och tillväxt för inhemska företag nämns också bättre evidensbaserade beslutsunderlag för regeringen i hanteringen av samhällsproblem, samt förbättrad kommunikation med medborgarna.⁶²

Ett centralt element i regeringens initiativ kring stora data är *Big Data Forum* som drivs och leds av *National Information Society Agency* (NIA), en myndighet under *Ministry of Security and Public Administration* (MOSPA). Forumet har samarbetspartners från både offentlig och privat sektor exempelvis *Electronics and Telecommunications Research Institute* (ETRI), *Science and Technology Policy Institute* (STEPI)⁶³, IBM Korea, Saltlux, och SAS Korea. Syftet är att verka som central kraft i framtagandet och implementeringen av Sydkoreas strategier inom stora och öppna data.

⁶² Intervju med HyKon Kim, Executive Director, National IT Policy Division, Head, Big Data Strategy Center, and Paul Cha, Principal Manager, National Information Society Agency, 11 juni, 2013.

⁶³ Både ETRI och STEPI är statsfinansierade forskningsinstitut.

President Park har satsat starkt och tagit igen flera års eftersläpning i arbetet för att främja stora data och öppna data. Det menar representanter från *Big Data Strategy Center*, en del av *Big Data Forum*, och rapporterar att arbetet med strategier, lagstiftning och finansiering inom stora och öppna data har accelererat sedan regeringsskiftet i Sydkorea tidigare i år. Presidenten har också pekat på stora och öppna data som det kanske mest lovande exemplet på hur Sydkorea bör använda sin kompetens inom IT för att förändra och korsbefrukta ekonomin. Regeringen räknar med att kunna skapa 43 000 jobb fram till 2016 genom att tillhandahålla öppna data.

Den nya regeringen och NIA har sett ett behov av att skapa en särskild funktion för öppna data. NIA lanserade därför under juli *Open Data Korea Forum*, vars funktion är att skapa en samlad styrning av användningen av öppna data inom näringsliv, regering, forskning och akademi. Forumets högsta prioritet är att ge råd om öppna data-strategier och att främja användningen av öppna data och/eller nya affärsmodeller som använder öppna data och som baseras på samarbete mellan offentlig och privat sektor.⁶⁴

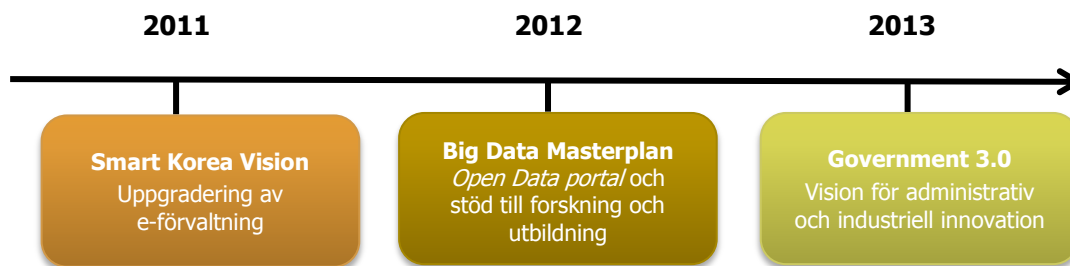
President Park behöver dock inte bygga sina strategier för stora och öppna data från grunden. Flera tidigare strategier har bäddat för att gradvis höja beredskapen för att implementera nya IT-lösningar i den sydkoreanska statsförvaltningen (se Figur 6). Det gäller den nationella *Smart Korea Vision* från 2011, och *Big Data Master Plan*, som släpptes i slutet av 2012. Den senare var utarbetad gemensamt av de dåvarande berörda departementen *Ministry of Education, Science and Technology*, *Ministry of Public Administration and Security*, *Ministry of Knowledge Economy*⁶⁵, samt *Korea Communications Commission* och *National Science and Technology Commission*.

Big Data Master Plan baserades på resultatet av en utredning av öppna data från berörda departement, en undersökning om prioritering av stora data projekt, samt på diskussioner med näringslivet, akademiska forskare och statstjänstemän. Planen innehåller fyra konkreta huvudområden: upprättandet av en webbaserad plattform som möjliggör delning och tillgång till stora data mellan departement och myndigheter i regeringen; framtagandet av en strategi för framtida forskning inom stora data; stöd till universitetens utbildningar inom stora data och finansiering av ”Triple Helix”-forskning; samt insatser för lagstiftning och integritetsskydd. Planens innehåll ingår i hur nuvarande regering prioriterar stora data-projekt för 2013-14, och planerar utöka omfattningen av användande och utveckling av stora data-teknik under 2015-17. Enligt uppgift kommer planen att tilldelas en budget på 3 miljarder SEK (500 miljarder KRW) för år 2016.

Efter *The Big Data Master Plan* släpptes den nya visionen för Sydkoreas hela förvaltning ”*Government 3.0*” i juni i år. Visionen har klar bäring på öppna data, och innehåller riktlinjer för hur offentliga data ska öppnas och delas för att uppnå två viktiga mål; en mer användarvänlig förvaltningstjänst, samt nytt företagande med arbetstillfällen av hög kvalitet.

⁶⁴ www.opendatakoreaforum.kr

⁶⁵ Nuvarande *Ministry of Trade, Industry and Energy*



Figur 6 Nationella IKT-strategier med bäring på stora data och öppna data.

Källa: Tillväxtanalys

Under sommaren 2013 kommunicerade regeringen ett mer konkret initiativ där öppen innovation används som drivkraft för att skapa tillväxt inom området stora data. Under hösten ska *Ministry of Science, ICT and Future Planning* (MSIP) arbeta med NIA i syfte att bygga ett datacenter som ska fungera som en testbädd för företag som använder eller berörs av stora data. Testbädden ska även underlätta för forskare vid universiteten.⁶⁶ Data-centrumet är tänkt att serva små och medelstora företag, uppstarts företag, universitet och medborgare som vill undersöka fenomenet stora data i förhoppning att det kan generera något positivt i deras företagande eller bidra till deras forskning. I NIA:s byggnader kommer centrumets online-plattform att anslutas till regeringens data- och portaltjänst som lagrar nationell statistik och annan offentlig information (data.go.kr).

Förvaltningens portaltjänst data.go.kr ska enligt planerna utökas kraftigt under de kommande åren. Webbplatsen erbjöd data i det öppna API-formatet inom 33 kategorier under 2012 och planerar att öka till 127 kategorier fram till 2015 och 150 dito till 2016. Data kan hittas på områden såsom hälsa, kriminalitet, transporter, väder, ekonomi, historia och resor.⁶⁷

5.3.1 Lagstiftning och översyn

För att de strategier som nämns ovan ska fungera smidigt antogs lagen *Act on Offering and Utilization of Public Data* i slutet av juli och kommer att gälla från den 31 oktober 2013. I slutet av 2012 var 30 000 olika databaser öppna för allmänheten via 2 194 offentliga webbplatser. Det betyder dock inte att dessa databaser nödvändigtvis hade särskild kommersiell potential. Den nya lagen föreskriver att departement, myndigheter och kommuner bör tillhandahålla offentliga uppgifter så länge det inte finns skälig anledning att inte göra det. Lagen garanterar att privatpersoner, företag och forskningsinstitut kan använda offentliga uppgifter i olika syfte, inklusive kommersiella. Regeringen bör underlätta för lätt och effektiv användning av uppgifterna för den privata sektorn.

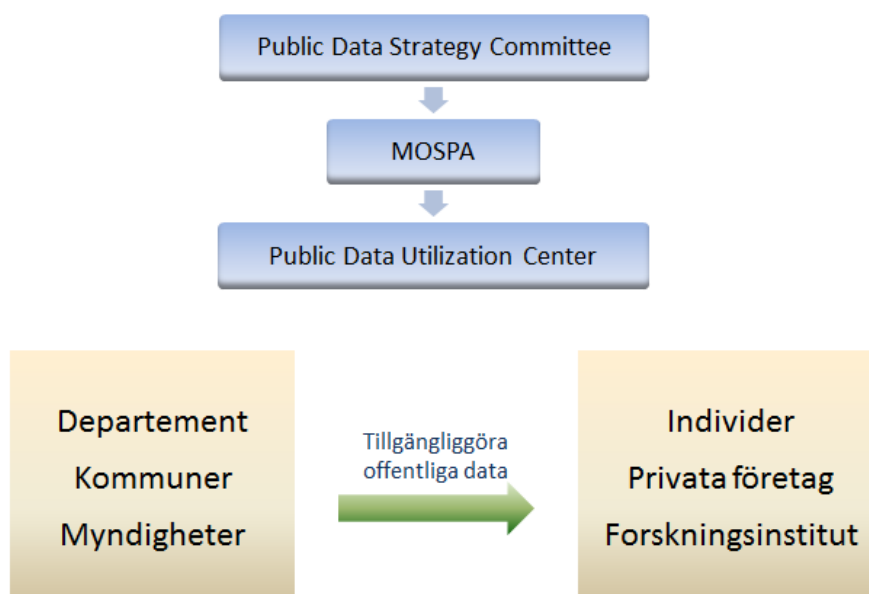
Ansvariga departement för främjande av stora data och öppna data är sedan regeringsskiftet tidigare år 2013 *Ministry of Security and Public Administration* (MOSPA), och det nya "superdepartementet" *Ministry of Science, ICT and Future Planning* (MSIP). Fördelningen mellan dessa är relativt tydlig, där MOSPA enbart fokuserar på att öppna data från

⁶⁶ ZDNet 13 May 2013, www.zdnet.com/s-korea-unveils-Big-data-center-to-help-industry-catch-up-7000015248/

⁶⁷ Big Data Master Plan 2012, www.mospa.go.kr

regeringen, medan MSIP jobbar för att både öppna data och andra former av stora data ska bli en tillväxtmotor för ekonomin.

Den lagstiftande makten anger hur styrningen av stora data-strategin utformas och genomförs. *Public Data Strategy Committee* leds av statsministern och har 33 medlemmar från både departement och privat sektor, inklusive MOSPA, vilka tar yttersta ansvaret för att samordna och bereda policyer för stora data. MOSPA ansvarar för att upprätta ett samordnat system mellan departementen, synkroniserat med lagstiftningen, och driver även *Public Data Utilization Center* som ett sekretariat (se Figur 7).



Figur 7 Översyn, koordinering och stöd för öppnandet av data från förvaltningen.

Källa: NIA

5.4 Resultat inom innovation och värdeskapande

Det finns i dagsläget en rik katalog med startade projekt baserade på öppnandet av offentliga data. Många är drivna av samarbetskonsortier med staten och näringslivet som medlemmar. De flesta projekten är relativt nystartade vilket gör det svårt att bedöma de långsiktiga konsekvenserna på arbetstillfällena och ekonomiskt hållbara affärsidéer.

När ”Big Data Master Plan” släpptes år 2012 innehöll planen tre projekt för år 2013, men innan 2017 ska totalt 16 projekt ha genomförts. Årets tre projekt handlar om att förutspå brott, att tidigt kunna få en uppfattning om naturkatastrofer samt att minska trafikolyckor genom samhällsinsatser. Det nya departementet MSIP har sedan dess adderat ytterligare sex stycken pilotprojekt inom stora data bland annat inom områdena transport, hälsa och medicinering⁶⁸ enligt följande:

- 1 Etablera nya rutter för nattbussar genom att koppla data mellan lokala teleoperatörer (konsortium lett av Korea Telecom)

⁶⁸ Korea IT Times 14 June 2013, www.koreaitimes.com/story/29507/future-ministry-selected-six-Big-data-pilot-projects

- 2 Varningstjänst för sjukdomsepidemier genom att länka uppgifter från den nationella hälsodatabasen med uppgifter från sociala medier (konsortium lett av National Health Insurance Service)
- 3 Tidigt varningssystem för ökad medicinsk säkerhet (konsortium lett av *Security Global Alliance Inc*⁶⁹)
- 4 Preventiv medicinsk diagnostjänst som förutspår ventrikulärrytm (konsortium lett av *Asian Medical Center*)
- 5 Tjänst för analys av köphistorik för att hjälpa småföretagare (konsortium lett av *OPEN mate*)
- 6 Söktjänst för nyheter (konsortium lett av *Advanced Institute of Convergence Technology*)

Den gemensamma nämnaren i dessa projekt är att de ska utveckla offentliga tjänster genom att kombinera stora mängder data från både den offentliga och privata sektorn. MSIP har förhoppningen att projekten kommer att fungera som en katalysator för att hitta andra kombinerade projekt mellan statliga institut och privata företag.

Även inom den privata sektorn är det strategiska användandet av stora data under utveckling. Landets telekomoperatörer och Internetportaler har exempelvis börjat erbjuda stora data-tjänster som är baserade på egna data, insamlad från abonnenter och registrerade medlemmar. Intressant nog har även dessa privata aktörer signalerat en sorts *Open Data Policy*. Sydkoreas största teleoperatör, SK Telecom, har meddelat att de planerar att offentliggöra anonymiserad data om användarmönster hos sina 27 miljoner abonnenter.⁷⁰

Tills dess att informationen släpps fri erbjuder SK Telecom tjänsten *Private Geovision*, som analyserar förutsättningarna för att öppna upp en ny butik. Analysen bygger på information som operatören dagligen samlar in från sina kunder, genom att de ringer samtal från olika fysiska platser. Företaget använder dessutom kreditkortsinformation från sina partnerföretag, vilket möjliggör att en tidsserieanalys av försäljningen i regionen skulle kunna adderas i tjänsten. För en avgift på cirka 15 USD sammanställs en 60-sidig rapport som inkluderar människors rörelsemönster samt rekommendationer för bästa läge för en planerad nyöppnad butik. En rapport kan exempelvis informera en butiksägare om att i ett visst område i Seoul ökar flödet av människor i 20–30-årsåldern runt klockan sex på kvällen, och att de föredrar lågpriskaffe. Rapporten rekommenderar då en butik som matchar dessa förutsättningar.

Ett annat exempel på tjänst som SK Telecom erbjuder är en navigeringsapplikation, *T-map*, som rekommenderar rutter med trafikinformation i realtid. Informationen samlas in från ett köpt trafikdatasystem, anslutet till 50 000 taxibilar och bussar som följs via GPS. Systemet gör det möjligt att få exakt trafikinformation med en felmarginal på max fem minuter, i hela landet.

Väderdata var tidigare begränsat till ett fåtal aktörer, men för fem år sedan öppnades också denna information av de ansvariga myndigheterna. Väderinformationen skapas av *Korea Meteorological Administration* (KMA) och mottas i realtid av *Korea Meteorological Industry Promotion Agency* (KMIPA), som i sin tur skickar vidare informationen genom

⁶⁹ *Security Global Alliance Inc* är en systemintegratör för säkerhetssystem.

⁷⁰ *The Kyunghyang Shinmun*, May 9 2013

http://english.khan.co.kr/khan_art_view.html?code=710100&artid=201305091950107

supportserverar till ett väderföretag. Informationen används exempelvis av försäkringsbolag när de utformar olycksfallsförsäkringar beroende på väderskillnader mellan olika regioner och perioder. Inom distribution används informationen tillsammans med försäljningsinformation från tidigare säsonger för att fatta beslut om skyltningens utformning, om lagerhållning och för att lägga beställningar för kommande säsong.⁷¹

Transportinformationen är ett av de mest populära, men också mest krävande, användningsområdena för stora data. Seoul city, samt 27 andra kommuner, tillhandahåller realtidsinformation för busstrafik, bland annat GPS GIS-information, 3D-information och trafikinformation i realtid. Exempelvis finns det för enbart Seouls lokaltrafik nu drygt 2 500 olika applikationer på marknaden.

Regeringen ser nu en förändring till mer heltäckande tjänster genom kombination av transportinformation med andra data från till exempel resor, logistik, hälsovård och fastigheter. Regeringen har insett att genom att dela och visualisera information om kollektivtrafik kan de reducera samhällets kostnader med 3,7 miljarder kronor årligen (618 miljarder KRW) samt 10 procent av kostnaderna för trafikstockningar, vilket motsvarar 175 miljarder kronor (30 biljoner KRW) per år.

I början av 2012 utvecklade det statligt ägda *Korea Oil Corporation* ett system som kan förutsäga oljepriset. Priset ges av en kombination av data från försäljningspriset på bensin från 12 500 bensinstationer, priset på råolja på den internationella marknaden, växlingskurser och finansiell transaktionsinformation, och kan ses en vecka fram i tiden. Det prognostiserade priset stämmer överens med det verkliga priset till mellan 99.48 och 100 procent, och det blir mer korrekt ju mer data som samlas in.

Det återstår att se om det blir de stora industrikonglomeraten som kommer vara de drivande vid utvecklingen av stora data i Sydkorea, eller om, vilket regeringen hoppas, det blir det små och medelstora företagen som kommer få en energikick av de nya möjligheterna.⁷²

5.4.1 Tillverkningsindustri

Samsung Electronics började tidigt under 2013 att leta efter en standardiserad plattform för samtliga sina stora data-system, och inrättade i juni en organisation i sitt *Media Solution Center* (MSC) för detta ändamål. MSC är den centrala funktionen inom Samsung Electronics som ska övervaka företagets mjukvaru- och innehållsstrategier. Samsung outsourcade tidigare marknadsanalyser av stora data, men planerar från och med hösten att själva analysera kunders och användares historik, användarmönster och preferenser. Analysen ska utgöra ett viktigt inspel till utveckling av nya produkter, tjänster och affärsmodeller. Först ut är datamängder från smarta telefoner, men målet är att även kunna extrahera stora datamängder från smarta TV-apparater och så småningom all annan elektronik företaget säljer.

POSCO, världens fjärde största ståltillverkare med huvudkontor i Pohang, förbereder sig nu för en *Proof of Concept* (PoC) inför införandet av ett större stora data-system. POSCO:s mål är att genom att förutsäga det internationella priset på nickel, bättre kunna styra sina upphandlingskostnader. Man ska också använda stora data för att analysera produktfel och

⁷¹ The Korea Meteorological Industry Promotion Agency, www.kmipa.or.kr

⁷² *Financial News*, June 2 2013,

http://www.fnnews.com/view?ra=Sent0901m_View&corp=fnnews&arcid=201306030100004380000277&cDateYear=2013&cDateMonth=06&cDateDay=02

därmed förbättra kvaliteten på POSCO:s breda sortiment av stålprodukter.⁷³ Planen som guidar projektet heter *Big Data Analysis Infrastructure Implementation Solution PoC Plan*. IT-leverantörer som Oracle och EMC, som redan gjort sig ett namn inom stora data, har meddelat att de avser offerera till detta projekt. Systemet ska baseras på ”text mining”-teknologi, som söker i stora mängder data och plockar ut den information som är intressant.

5.4.2 Hälsa- och sjukvård

Sydkorea introducerade år 1977 ett nationellt sjukförsäkringssystem. Vid starten omfattade det endast cirka nio procent av den totala befolkningen, men ambitionen var att så snabbt som möjligt täcka hela befolkningen. Tolv år senare var ambitionen så gott som uppfylld efter en snabb utbyggnad. Återstående 3–4 procent omfattas nu istället av *Medical Aid Program*, ett alternativ välfärdsprogram för de fattigare delarna av befolkningen. Sedan år 2000 finns det även registrerad digitaliserad hälsoinformation på hela befolkningen. De stora datamängderna från försäkring och hälsovård har identifierats som högtintressant för stora data-tillämpningar av såväl regering som företag.⁷⁴

The *National Health Insurance Service* (NHIS) har börjat diskutera när, snarare än om, man ska öppna dessa data efter anonymisering. NHIS har samlat in information från försäkringar och journaler som omfattar över 800 miljarder olika fall av medicinska undersökningar. Tidigare i år etablerade NHIS en testdatabas, som inkluderade information om behandlingshistorik, undersökningsresultat, boende, medicinska försäkringspremier från ett flertal sjukvårdsinstitutioner mellan åren 2002–2010. Databasen representerar ett urval av 1 miljon människor från totalt 50 miljoner, som representerar olika åldrar, kön, region och inkomstnivå. NHIS plan är att först öppna upp databasen för forskningssyfte och för analys av förvaltningen för att skapa bättre policy. Därefter vill de få till stånd en offentlig diskussion för att bestämma om hur, när och till vem informationen ska göras tillgänglig.⁷⁵

5.4.3 Social innovation

I maj år 2013 startade *Korea Institute for Health and Social Affairs* (KIHASA), tillhörande premiärministerns kansli, ett projekt som behandlar självmords- och mobbningsfrågor med hjälp av stora data. KIHASA hade tidigare utfört forskning som jämför den årliga självmordsfrekvensen med i vilken utsträckning termer associerade med självmord söks på internet. Slutsatsen från forskningen var att det fanns en stark korrelation, samt förslag på att analys av stora datamängder från sociala medier skulle kunna förutsäga och förebygga självmord. Baserat på resultaten utvecklade NIA tillsammans med ett privat företag under tre månader en uppsättning algoritmer som i realtid plockar ut och analyserar självmords- och mobbningsrelaterad information från 214 sociala medietjänster. Arbetet med att utveckla tjänsten var utmanande genom att den språkliga och kulturella kontext som tonåringar befann sig i var tvungen att ”dechiffreras”, vilket involverade flera lingvistik i utvecklingsgruppen. Baserat på projektet kommer KIHASA att presentera förslag på lösningar som kan förhindra självmord och mobbning, samt att upprätta ett övervaknings-system som kan söka efter skadliga hemsidor och fraser.⁷⁶

⁷³ *Korea IT Times*, Tuesday, July 16 2013, <http://www.koreaitimes.com/story/30470/posco-follows-samsung-electronics-and-hyundai-motor-company-use-Big-data>

⁷⁴ *Republic of Korea Health System Review* (2009) http://www.euro.who.int/_data/assets/pdf_file/0019/101476/E93762.pdf

⁷⁵ *Rapportian*, February 13 2013,

⁷⁶ *Digital Times* May 7, www.dt.co.kr/contents.html?article_no=2013050702010960746002

5.5 Utmaningar för vidare utveckling

5.5.1 Främjande

Den sydkoreanska regeringen gör klart att det är högt prioriterat att underlätta främjandet och användningen av stora och öppna data av flera skäl. Det största steget i den riktningen togs genom att lagstiftning öppnade för tillgänglighet av offentliga uppgifter mellan myndigheterna. Utifrån dessa förutsättningar förbereds nu en främjarkampanj på flera fronter för att uppmuntra praktiska tillämpningar bland departement och myndigheter, samt för att stödja andra mål såsom exempelvis utökad teknisk infrastruktur, FoU-investeringar och utbildning av forskare inom stora data.

Kampanjen kommer främst att bedrivas av det nya *Center for Open Data Utilization Support* på NIA och i samarbete med MOSPA och MSIP, vilket förväntas vara operativt den första januari 2014. Centrumet kommer att stödja förvaltningen och industrin för att göra det möjligt för alla intresserade parter att börja analysera stora datamängder. Det nya centrumet förväntas bli en testbädd för företag som är relaterade till stora data och det väntas även vara ett alternativ för att stödja forskare vid universiteten.⁷⁷ Centrumet kommer att kunna gå in med experthjälp, i synnerhet till myndigheter med relativt lägre IT-kompetens, för att få till stånd mer öppnande av data från staten.

Centrumet slutför under hösten ett planeringsdokument för perioden 2013 till 2017 med projekt inom elva områden, där stora och öppna data har en god chans att skapa värde genom konvergens mellan olika områden (exempelvis väder, transport, patent, offentliga finanser, sysselsättning, välfärd, säkerhet). En första version av centrumets webbplattform ska vara färdig för en testkörning i september. Plattformen ska även anslutas till regeringens data- och portaltjänst som lagrar nationell statistik och annan offentlig information. Man fokuserar på en offentlig kampanj mot allmänheten för att öka förståelsen, den ”politiska trovärdigheten”, för stora data och öppna data genom bland annat användarforum och idétävlingar.

Center for Open Data Utilization Support har redan cirka 100 anställda och planerar att öka sin storlek när efterfrågan på support från departement, myndigheter och industri växer.⁷⁸

5.5.2 Standardisering

Standardisering är en viktig del av *The Big Data Master Plan* från 2012, som inte bara pekar på dataformat, utan också på stora data-konceptet i sig; terminologi och konceptuella modeller. Standardisering av det senare ses som det kanske svåraste. Standardisering av handhavande och processer för öppna data på olika statliga organisationer kommer att kräva stort arbete. Som en uppföljning till förra årets strategi har NIA under juni flaggat för ett projekt som ska innehålla en standardiseringsplan – *Big Data Sharing & Utilization System Project*. Detta projekt syftar till att färdigställa ett ramverk för delning och användning av stora data för både offentlig och privat sektor med start 2013, och med vidare fokus på standardisering under 2014.

För närvarande saknas riktlinjer eller rekommendationer för i vilka former förvaltningen ska öppna upp data, även om en hel del *de facto* standardisering håller på att formas.

⁷⁷ ZD Net, May 13, <http://www.zdnet.com/s-korea-unveils-Big-data-center-to-help-industry-catch-up-7000015248>

⁷⁸ Intervju med Sukhyun Song, Director, Information Resource Planning Department, NIA, 19 augusti, 2013

Myndigheter använder data i olika format så som öppna API, FTP, eller Webformat, vilket sänker användarvänligheten för de som är intresserade av öppna data. NIA planerar att undersöka den nuvarande situationen för att dela och utnyttja teknik, policyer och standardisering mellan olika delar av förvaltningen. Därpå vill NIA etablera ett stora data-system som inkluderar ett datautbyte av protokoll, serviceformat, specifika ramverk och så vidare. De planerar även för en publicering av riktlinjer och för en stegvis ansökningsstrategi, tillsammans med standardiseringsstrategin för inhemska och internationella ändamål.⁷⁹

Det internationella samspelet vad gäller standardisering är även en integrerad del av diskussionerna bland policyarkitekterna. En viktig aktör i Sydkorea för IT-standarder är *Telecommunications Technology Association (TTA)*. TTA arrangerade år 2012 ett forum med syfte att säkerställa att Sydkorea proaktivt kan förbereda och svara mot internationella och inhemska trender kring stora data-standarder. Till exempel utvecklar TTA ett stora data GS (*Good Software*) certifikatsystemet, som har i syfte att användas från och med nästa år för att främja inhemska mjukvaruföretag inom stora data. GS-certifikat kräver bedömningar av format, funktioner, och tillämpad teknik för mjukvaruprodukterna som ska användas för hantering av stora datamängder.⁸⁰

Sydkorea har även arbetat fram ett licenssystem baserat på den internationella organisationen *Creative Commons* licenstyper för att underlätta användningen av offentligt innehåll utan intrång i upphovsrätten. Licenssystemet, kallat *The Korea Open Government License (KOGI)*, förenklar processen för ett godkännande och begränsar varken gällande tidsperiod eller användningsändamål.⁸¹

5.5.3 Forskning och utveckling

Experter i Sydkorea från såväl akademiskt som politiskt håll ondgör sig över hur långt efter landet ligger rent tekniskt för att på bästa sätt kunna utveckla framtidens tjänster inom stora data.

Det inflytelserika statliga *Electronics and Telecommunications Research Institute (ETRI)* som varit centrala i mycket av landets framsteg inom IT, gör bedömningen att den inhemska tekniknivån släpar efter två till fem år jämfört med globala stora data-aktörer som Google, Oracle och IBM. Baserat på denna bedömning har berörda departement aktivt börjat finansiera FoU inom stora data för förbättringar av metoder och teknik relaterat till statistik, lingvistik, matematik och så vidare. Dåvarande departementet *Ministry of Knowledge Economy*, numera omorganiserat till nuvarande MOTIE, startade 2012 finansiering av tre stora data-projekt relaterade till programvara med 43 miljoner kronor (7,3 miljarder KRW). ETRI etablerade i början av 2012 *the Big Data Software Institute*, och har samarbetat med IBM inom områden från vattenförvaltning och smarta elnät, för att hitta tillämpningar för stora data.

⁷⁹ *Etnews*, June 16 2013, http://www.etnews.com/news/computing/informatization/2782979_1475.html

⁸⁰ *Telecommunications Technology Association (TTA)*, är en privat IT standardorganisation, som har en ledande roll inom IT-området i Sydkorea. TTA erbjuder övergripande och , särskilda tjänster beträffande IT-standard. Dess huvudsakliga roll innebär IT-standardiseringsplanering och ny IT-standard, etablerande, testning och certifiering av IT-produkter. TTA utför sin verksamhet i nära samarbete med företag, organisationer och grupper av IKT-företag såsom nätverksoperatörer, producenter av tjänster, tillverkare av hårdvara, akademi, FoU- institut osv..

⁸¹ www.kogil.or.kr

Dåvarande *Ministry of Education, Science and Technology* (MEST), nuvarande *Ministry of Education*, finansierade i en utlysning för stora data 14 projekt med 18 miljoner kr (3,1 miljarder KRW), däribland utveckling av bättre mjukvara för att leta igenom stora datamängder. Därtill finansierades med start 2012 ett infrastrukturprojekt inom stora data med 44 miljoner kronor (7,6 miljarder KRW).

The *Big Data Master Plan* konstaterade att den nuvarande tekniken för stora data i Sydkorea är särskilt svag när det gäller insamling av data och användning av distribuerade nätverk, och anser att det är brådskande och nödvändigt att främja FoU-investeringar. Senare under 2013 förväntas en ny uppsättning utlysningar för FoU inom stora data från regeringen.

5.5.4 Brist på experter

En långsiktig och allvarlig svaghet som Sydkorea delar med de flesta länder är brist på relevanta experter inom stora data. Om tillväxtprognoserna för stora data ska realiseras kommer det att behövas betydligt fler unga med examen inom bland annat matematik, statistik och bildbehandling, liksom vidareutbildning av dagens arbetskraft inom IT-industri och statlig förvaltning. I syfte att angripa detta problem har MSIP instruerat flera av sina underliggande myndigheter, främst *National IT Industry Promotion Agency*, *Federation of Korean Information Industries* och *Korea Database Agency*, att införa eller utöka utbildningsprogram inom stora data. Som ett tidigt exempel har *Korea Database Agency* startat ett utbildningsprogram med 200 antagna, som deltar genom både klassrumsundervisning och online-verktyg. Programmet är gratis för deltagare från små och medelstora företag och subventionerat med 80 procent för de större företagens anställda. Ett annat exempel på insatser är ett samarbete vad gäller kurser mellan *National IT Industry Promotion Agency* och Yonsei University.

6 USA

6.1 Sammanfattning

I USA högprioriteras stora och öppna data av Vita huset. Förra året tillkännagav Obama en satsning på 1,3 miljarder kronor (200 miljoner USD) för FoU-investeringar inom stora data. Satsningen syftar till att ta fram verktyg och tekniker för att samla in, lagra, hantera, analysera och dela stora mängder data, utveckla nya tekniker för att producera snabbare forskningsresultat samt öka antalet arbetstillfällen inom området. Hälso- och sjukvård kan dra många fördelar av stora data och bland annat använda data proaktivt för att indikera och/eller identifiera högrisk patienter i ett tidigt skede. Vidare kan stora data bidra till sänkta sjukvårdskostnaderna. Öppna data är en annan viktig prioritering för president Obama. Att tillgängliggöra datamängder lagrade vid offentliga myndigheter för allmänhet och företag anses öka förutsättningarna för innovation och ekonomisk tillväxt. Stora och öppna data är dock politiskt laddat i USA eftersom det har nära kopplingar till försvarsfrågor och nationell säkerhet.

6.2 Inledning

Tack vare Edward Snowden har *cybersecurity* och stora data kommit upp på allmänhetens radar och diskuteras numera av långt fler än de redan insatta. Den tidigare *National Security Agency* (NSA)-anställda Snowdens avslöjanden om PRISM, USA:s hemligstämplade elektroniska övervakningsprogram, har lett till ökad aktualitet kring olika *cybersecurity*-frågor inom USA och i deras förhållande till andra länder, men det har också lett till en bredare diskussion kring stora data. *Cybersecurity* är komplext då det ligger i gränslandet mellan IT, stora och öppna data samt försvarsfrågor och därmed rymmer vissa *policymotsättningar*. Området är politiskt laddat då det har nära kopplingar till försvarsfrågor och nationell säkerhet. Begreppet *cybersecurity* har emellertid vidgats till att numera även omfatta frågor som upphovsrätt, sekretess, driftsäkerhet, internetbedrägerier och dataintrång.

Fenomenet stora och öppna data är dock inte nytt i sig. Redan på 1970-talet gjordes väderdata allmänt tillgängliga av *National Oceanic and Atmospheric Administration* och på 1980-talet öppnade USA upp åtkomst till stora data från *Global Positioning System* (GPS)-satelliter vilka tidigare endast använts för militärt bruk. Så utan att fundera i termer av stora och öppna data har de flesta av oss tacksamt tagit emot både väderrapporter och GPS-data för positionering. Idag torde få ifrågasätta innovationskraften i besluten att offentliggöra väder- och satellitdata som vem som helst i världen kan utnyttja eller använda för att utveckla och bygga upp nya verksamheter kring. Förhoppningen – och ofta övertygelsen – är därför att stora och öppna data kommer skapa både ekonomisk tillväxt samt förenkla och förbättra livet för oss alla.

Stora och öppna data erbjuder svindlande möjligheter vilka kan upplevas både som lovande och skrämmande. Oavsett våra känslor för fenomenet kommer det sannolikt att revolutionera vårt arbetsliv och vår vardag. Med stora och öppna data kommer vi – och andra – kunna samla in och analysera enorma datamängder samtidigt som vi zoomar in på detaljnivåer som aldrig tidigare varit möjligt. Vidare gör stora data oss mindre kvalitetskänsliga – de enorma kvantiteterna av data gör att korrelationer kan dras utan precisa och exakta fakta. Således är det viktigt att notera att korrelationer utifrån stora data

mer tenderar att svara på frågan ”Vad?” än ”Varför?”. Datamängderna i stora data förklarar inte händelser, de identifierar mönster bland dem.

6.3 Riktade policyer och strategier

Efter ett kongressbeslut år 1976 har den amerikanske presidenten det rådgivande kontoret *Office of Science and Technology Policy*, OSTP, till sitt förfogande vilket ger vetenskapliga och tekniska råd i aktuella inrikes- och internationella frågor. OSTP har även mandat och medel att samordna institutionsöverskridande insatser för att utveckla och genomföra policies inom vetenskap och teknik. Ett fokus för OSTP är att koordinera och guida implementeringen av Obamas *Open Data Policy*⁸².

Redan 1991 skapades det federala programmet NITRD, *the Networking and Information Technology Research and Development program*, för att finansiera forskning och utveckling inom IT-området. NITRD är inordnat under OSTP och *National Coordination Office (NCO) for Networking and Information Technology Research and Development*. NITRD består av 17 departement och myndigheter däribland NIST (*National Institute of Standards and Technology*), DOE (*Department of Energy*), NIH (*National Institutes of Health*), NSA, DHS (*Department of Homeland Security*), NASA (*National Aeronautics and Space Administration*) samt NSF (*National Science Foundation*) vilka själva finansierar och delvis utför IT-relaterad forskning. Medlemsorganisationernas forskning och utveckling koordineras inom ett flertal områden, där bland annat cybersäkerhet, nätverksfrågor, framtidens internet, högkapacitetsdatabehandling, människa–maskin-gränssnitt ingår. Vidare hanteras samarbete kring ett antal nya och stora utmaningar i särskilda arbetsgrupper, där bland annat stora data och hälsorelaterad IT ingår.

För att kunna konkurrera globalt har president Obama en uttalad målsättning om att USA bör satsa på ökad utbyggnad och ökat utnyttjande av olika IT-resurser. I Obamas memorandum *Science and Technology Priorities for the FY 2014 Budget*⁸³, framhålls till exempel två fokusområden som myndigheter förväntas prioritera i sina respektive budgetar under 2014; *stora data* samt forskning i enlighet med strategin *Trustworthy Cyberspace*.⁸⁴

6.3.1 Stora data

Efter en rekommendation från PCAST, *Presidents Council of Advisors on Science and Technology*⁸⁵, där det påpekades att USA underinvesterar i stora data, tillkännagav president Obama år 2012 en satsning på 1,3 miljarder kronor (200 miljoner USD) för FoU-investeringar inom stora data, *Big Data Research and Development Initiative*⁸⁶. Satsningen genomförs genom OSTP, som tillsammans med sex olika federala organ och departement ska ta fram verktyg och tekniker för att samla in, lagra, hantera, analysera och dela stora mängder data. Vidare syftar satsningen till att utveckla nya tekniker för att producera snabbare forskningsresultat genom stora data samt öka antalet arbetstillfällen för att utveckla och arbeta inom detta område.

Biomedicinsk forskning genererar och använder sig alltmer av stora data. Dessvärre är förmågan att hantera, integrera och analysera datamängderna – samt lokalisera och an-

⁸² <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>

⁸³ <http://www.nitrd.gov/pubs/2014supplement/FY2014NITRDSupplement.pdf>

⁸⁴ <http://www.whitehouse.gov/sites/default/files/m-12-15.pdf>

⁸⁵ PCAST ett vetenskapligt råd under OSTP. PCAST består av cirka 20 ledande forskare som lämnar råd samt politiska rekommendationer till presidenten inom områdena vetenskap, teknik och innovation.

⁸⁶ http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf

vända data genererat av andra – ofta begränsad vilket tros bero på brist av verktyg, tillgänglighet och utbildning. Mot denna bakgrund lanserade NIH i slutet av 2012 initiativet *Big Data to Knowledge*⁸⁷ (BD2K) vilket just syftar till att stödja biomedicinsk forskning, dess implementering samt utbildning inom stora data. I somras meddelade NIH att de – inom BD2K-initiativet – kommer att finansiera upp till åtta stycken forskardrivna *Big Data Centers of Excellence* med 24 miljoner USD (150 miljoner SEK) årligen under fyra år⁸⁸. Tanken är att de nya excellens-centrumen ska utveckla innovativa metoder, programvara och verktyg för delning, integration, analys och förvaltning av stora data och därigenom öka forskarsamhällets möjligheter att använda sig av stora data. Satsning på stora data-center förväntas att snart följas av andra initiativ inom BD2K.

6.3.2 Öppna data

En annan viktig IT-prioritering för president Obama är att tillgängliggöra datamängder lagrade vid offentliga myndigheter för allmänhet och företag. Vita Huset har därför antagit en *Open Data Policy*⁸⁹ inom ramen för den tidigare existerande *Freedom of Information Act* (FOIA)⁹⁰, vilket ger den legala grunden för allmänheten att begära ut dokument lagrade hos myndigheter. Fokus för amerikansk *Open Data Policy* är att tillhandahålla stora mängder data, potentiellt extraherade från många individers ärenden, för att kunna sammanställa och analysera datamängderna och därigenom exempelvis förstå konsumentbeteende eller energianvändning. Datamängderna ses som en resurs vilka de offentliga myndigheterna förfogar över, och medborgarna beskrivs som konsumenter som kan ta del av och konsumera denna nyttighet. Mer lättillgänglig data anses öka förutsättningar för innovation och ekonomisk tillväxt. Men i en globaliserad värld blir USA:s digitalt publicerade data tillgänglig för hela världen vilket förstås både kan ha sina för- och nackdelar.

Styrdokument gällande öppna data framhåller att fokus på offentliga data bör ligga på informationsinnehåll, snarare än på presentationsformat.⁹¹ IT-lösningar ska därför till exempel inte ha som syfte att generera slutprodukter som specifikt formaterade Word-dokument. Istället är syftet att tillgängliggöra stora mängder enheter av standardiserade 'data'. Den enskilda konsumenten bör kunna extrahera data från databaserna utifrån sina individuella urvalskriterier och därefter sammanställa datan i det format som konsumenten finner mest lämpligt för sina analytiska syften. Datan behöver därför bland annat vara maskinläsbar och associerad med metadata vilket gör den sökbar. Att upprätta sådana stora databaser där olika slags data kan korreleras med varandra förutsätter att myndigheter har gemensamma eller kompatibla IT-plattformar för lagring av sådan data. I denna kontext anses det viktigt att datan är lättillgänglig från den sökande konsumentens perspektiv samt att integritetshänsyn bevaras. Vidare ska webplattformar upprättas genom vilka konsumenter lätt kan söka data. Respektive myndighet uppmanas att snarast identifiera två prioriterade typer av data som ska göras tillgängliga genom sådana plattformar. Vidare framhålls att datan ska vara lättillgänglig från olika former av IT-enheter.

Målsättningen i Obamas öppna data policy skiljer sig i grunden för den svenska offentlighetsprincipen, som i första hand syftar till att kunna följa hur enskilda ärenden handlagts för att avgöra om individernas rättigheter och skyldigheter beaktats på ett adekvat sätt. Obamas öppna data policy vill snarare tillgängliggöra och sammanställa stora

⁸⁷ <http://bd2k.nih.gov/#sthash.eRWgru54.dpbs>

⁸⁸ <http://www.nih.gov/news/health/jul2013/nih-22.htm>

⁸⁹ <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>

⁹⁰ <http://www.justice.gov/ag/foia-memo-march2009.pdf>

⁹¹ *Digital Government, Building a 21st Century Platform for Better Serve the American People.*

mängder data, och analysera dessa för att generera underlag för uppfinningar och innovationer. Fokus för öppna data policy är inte primärt att följa enskilda ärenden, utan snarare att tillgängliggöra stora mängder data, som potentiellt kan extraheras från olika individuella ärenden. Likaså är inte 'handlingen' i fokus i amerikansk öppna data policy, vilket är centralt i svensk offentlighetsprincip. USA:s öppna data policy vänder sig mot framhävandet av specifika helhetsdokument och fokuserar på de enskilda data som kan rymmas inom ramen för en given handling eller dokument.

6.4 Resultat i form av innovationer och värdeskapande

Hälso- och sjukvård kan dra många fördelar av stora data och bland annat använda data proaktivt för att ge indikationer och/eller identifiera högrisk patienter i ett tidigt skede. Sjukhus, medicinska forskare och hälsorelaterade myndigheter har under lång tid journalfört patient- och hälsodata så det finns bra möjligheter att få fram stora mängder data från mätserier gällande exempelvis läkemedelsprövningar, genomsnittliga kostnader för olika medicinska åtgärder, demografisk information om patienter, genomsnittliga väntetider för vård, etc. Förhoppningen är att utvecklingen av verktyg inom stora data ska ge svar på frågor såsom vilka läkemedel ger minst ofta biverkningar. Vilka läkare skriver ut mest antibiotika? Vilka behandlingar är mest kostnadseffektiva? I vilka regioner sprids influensa just nu?

En annan förhoppning är att stora data kan bidra till att sänka sjukvårdskostnaderna. USA lider av höga sjukvårdskostnader och en bidragande orsak anses vara att sjukvården är decentraliserad och okoordinerad vilket kan leda till dupliceringar av behandlingar. Det finns många olika vårdregister i USA ägda av branschorganisationer, försäkringsbolag, vårdgivare, forskare eller andra – men få av registren är landsomfattande. USA:s regering har investerat i att införa elektroniska patientjournaler i sjukvården och betalar ut ersättningar till sjukhus och vårdgivare som implementerar, uppgraderar till eller uppvisar ”meningsfull användning” av godkända elektroniska patientjournaler. I framtiden kommer rapportering av uppgifter till kliniska dataregister att vara ett av de kriterier som yrkesutövare och sjukhus måste uppfylla för att få fortsätta delta i incitamentsprogram för elektroniska patientjournaler. Idag används elektroniska patientjournaler av drygt 50 procent av läkarna och 80 procent av sjukhusen, vilket är en markant ökning. Elektronisk förskrivning av receptbelagda läkemedel har ökat från 0,8 procent i december 2006 till 53 procent i januari 2013⁹².

I USA saknas en central myndighet som säkerställer kvaliteten på patientvården, men det viktigaste federala departementet är *Department of Health and Human Services* (HHS). *The Office of the National Coordinator for Health Information Technology* (ONC), som är förlagt inom HHS, är den främsta federala myndigheten med uppgift att samordna riktäckande insatser för att använda avancerad informationsteknik för hälsoinformation och elektroniskt utbyte av hälsoinformation.

Ett konkret exempel på hur stora data kan användas för att bilda företag inom hälsovården är vårdföretaget iTriage⁹³ vilket grundades 2008 av två akutmiljöläkare. Bolaget har utvecklat en hälso-app som är en av de mest nedladdade i sin kategori, både för iPhone och Android, och den används idag – kostnadsfritt – av nästan 10 miljoner människor. Appen identifierar, lokaliserar och bokar tid hos lämplig vårdgivare utifrån uppgivna symtom och

⁹² <http://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-Sheets/2013-Fact-Sheets-Items/2013-04-23.html>

⁹³ <http://about.itriagehealth.com/>

behov. Vidare kan den jämföra läkemedel, spara påminnelser och lagra patientjournaler. iTriage-appen använder sig av stora och öppna data från bland annat HHS medan de strikt medicinska råden kommer från, och utvärderas av, Harvard Medical School. Företaget iTriage har idag 90 anställda och samarbetar med tusentals vårdgivare.

Ett annat exempel på hur stora data levererar innovativa lösningar, men till andra branscher än hälso- och sjukvården, är företaget Opower⁹⁴. Bolaget startades 2007 och använder stora data för att minska energikostnader åt sina idag drygt 15 miljoner kunder världen över, däribland fler än 50 miljoner amerikanska hushåll. Opower använder sig av programvara som behandlar och lagrar stora data samt analyserar tiotals miljoner smarta elmätare med 15 minuters mellanrum. Analyserna ger riktad energiinformation till kunderna, utan att de har några enheter installerade i hemmet, vilket gör att kunderna sparar både energi och pengar. Fler än 2,8 terawattimmar energi och mer än 330 miljoner USD i energikostnader har hitintills sparats genom Opowers stora data-verksamhet. Företaget har idag fler än 400 anställda.

6.5 Framtida utmaningar

I USA högprioriteras stora data och öppna data av Vita huset – liksom IT- och kommunikationsnät och allmänhetens möjlighet att delta i IT-samhället.

Idag har i stort sett hela den amerikanska befolkningen tillgänglighet till någon form av internet-uppkoppling med en hastighet på 3 Mbps eller högre. Denna utbyggnad har i princip skett av enskilda företag på marknadsbasis och utan större insatser från Vita Huset eller andra politiska organ. Drygt 80 procent av hushållen har möjlighet att ansluta sig till bredband med hastigheter på minst 100 Mbps. Privata aktörer har i stort sett täckt in de lönsamma geografiska områdena, medan glesbefolkade områden fortfarande har sämre täckning. Policyfokus har därför delvis skiftat från utbyggnad av markbundna kablar till satsningar på 4G med målet att inkludera glesbygden.

En stor IT-infrastrukturssatsning om sammanlagt fyra miljarder USD hanteras av *National Telecommunications and Information Administration* (NTIA) genom två anslagsprogram: *Broadband Technology Opportunities Program* (BTOP)⁹⁵ och *the State Broadband Initiative* (SBI)⁹⁶. I juni 2012 undertecknade Obama likaså ett verkställande direktiv, *Accelerating Broad-band Infrastructure Deployment* (ABID)⁹⁷. I ett alltmer uppkopplat och utbrett IT-samhälle är det viktigt att öka den generella kunskapen om och intresset för IT. En allmän användning av internet anses vara viktigt ur många avseenden såsom tillgänglighet, närbarhet men även ur ett demokratiperspektiv. För att stärka allmänhetens digitala kompetens har Obama-administrationen, startat ett *Digital Literacy Initiative*⁹⁸ för att sporra utvecklingen av det digitala kunnandet i alla åldersgrupper.

⁹⁴ <http://opower.com/>

⁹⁵ <http://www2.ntia.doc.gov/>

⁹⁶ <http://www2.ntia.doc.gov/SBDD>

⁹⁷ <http://www.whitehouse.gov/the-press-office/2012/06/14/executive-order-accelerating-broadband-infrastructure-deployment>

⁹⁸ <http://www.digitalliteracy.gov/>

7 Sammanfattande resultat

Baserat på landstudierna har följande sammanfattande huvuddrag identifierats:

7.1 Forskningsinsatser

Stora data är fortfarande ett relativt nytt och växande område. Visserligen finns det områden där man sedan länge har hanterat mycket stora datamängder. Det som är nytt är att digitaliseringen av många funktioner ger upphov till mycket stora datamängder inom många fler områden, samt att tekniken för att lagra, hantera och analysera data har blivit så billiga och tillgängliga att fler aktörer kan använda den inom nya områden.

En hel del av utvecklingen av nya tillämpningar är drivna och finansierade av näringslivet, men det finns stor potential för offentligt initierad eller finansierad forskning med syfte att öka kunskapen och tillämpningar från stora och komplexa datamängder. Kunskapsområden omfattar allt från att samla, lagra och bevara stora och komplexa datamängder till att hitta bättre sätt att organisera, analysera och dela data. En integritetsaspekt i detta sammanhang är att data samlas in för ett specifikt ändamål och sedan eventuellt kan användas i ett helt annat sammanhang.

Ett relaterat område är att göra data tillgängligt över geografiska gränser för att möjliggöra eller underlätta forskning. Detta kräver överenskommelser mellan länder och aktörer kring format och standarder. Förutom rent tekniska områden öppnar stora och öppna data även en rad forskningsfrågor inom bland annat etik och förvaltningsfilosofi.

I de studerade länderna konstateras att forskningsinsatser har initierats för att öka kompetensen och utvecklingen inom området. I USA har till exempel resurser allokerats till *Big Data Centers of Excellence*. I Sydkoreas *Big Data Masterplan* ingår en strategi för framtida forskning om stora data som ett av fyra strategiska huvudområden.

7.2 Infrastruktur

De genomförda landrapporterna visar också att insatser görs på infrastruktur av olika slag, såsom superdatorer och datacenter, men även utbyggnad av bredband för att möjliggör överföring av datamängder. Etablering av gemensamma portaler för öppna data är en insats för att främja utvecklingen som återfinns i de flesta studerade länderna.

En viktig komponent är att etablera stödprogram för uppstarts företag inom IT-sektorn och har lanserats i Indien och Sydkorea, där företagen erbjuds tillgång till inkubatorer och acceleratorer. I Sydkorea ser vi även uppbyggnaden av ”virtuella sandlådor” där vem som helst kan få tillgång till grundläggande verktyg samt datamängder för att själva experimentera fram nya tjänster.

I Storbritannien erbjuds små och medelstora företag stöd för att få utväxling på sina immateriella rättigheter. Här har också olika grupper och organ inrättats för att främja utvecklingen inom det immaterialrättsliga området.

Ett område med ökande internationella samarbeten är kring forskningsinfrastruktur. Här finns exempel på OECD-initiativ, till exempel standardisering för att göra det möjligt att skapa gemensamma databaser.

7.3 Kompetens

Vidare konstateras att tillgång till kompetens framträder som en utmaning. Från näringslivsperspektiv handlar det om att företag ska kunna erbjuda och utveckla nya tjänster, vilket kräver ny och rätt kompetens som emellertid ännu inte alltid finns tillgänglig vid universitet och högskolor. För Indien, som har en ambition att produktionen inom IT-industrin ska gå uppåt i värdekedjan, konstateras en kompetensbrist för att kunna attrahera företag inom stora data och ansträngningar görs nu för att utveckla denna. Detta görs genom utveckling av både distans- och vidareutbildningar. Olika utbildningar kommer att tas fram i samverkan mellan näringslivet och utbildningsinstanser. Utveckling av distansutbildningar görs även i Storbritannien.

En annan kompetensdimension är att användare måste ha en hög kompetens för att kunna dra nytta av de möjligheter som nya tjänster baserade på stora datamängder innebär. Men även för att kunna göra egna riskbedömningar vad gäller de nya möjligheterna. Exempel på initiativ kommer från USA, där ett *Digital Literacy Initiative* har startats för att stärka den digitala kompetensen i alla åldersgrupper.

7.4 Insatser för att stödja myndigheter och andra aktörer

Det finns också behov av kompetenslyft bland de statliga organisationer som genom ny lagstiftning och policy ska offentliggöra data på ett strukturerat och lättillgängligt sätt. Därtill krävs också resurser för att löpande kunna tillgängliggöra stora datamängder.

Datakvalitet är också en viktig fråga. För att tillgängligheten på data ska få den effekt som policyn syftar till bör den hålla en genomgående hög kvalitet och vara lättillgänglig. Datakvaliteten kommer inte att säkerställas enbart genom att öppna stora dataarkiv, utan löpande arbete måste göras för att hålla data uppdaterad, korrekt och tillgänglig. Detta kan kräva ny kompetens och/eller resurser beroende på hur väl förberedd till exempel en myndighet är. Från Sydkorea och Japan kommer exempel på att aktörer, både privata och offentliga, erbjuder stöd för att öppna upp data. Bland annat har man inrättat ett supportcenter i Sydkorea som myndigheter kan få stöd av. Dessa stödinsatser grundar sig på att öppna data både kräver kompetens och resurser, och har setts som nödvändiga för att policy inom området ska kunna implementeras inom rimliga tidsramar.

Att införa gemensamma riktlinjer och rekommendationer för hur offentliga organisationer ska tillgängliggöra data har också setts som viktigt. Sydkorea har utsett ett departement och en myndighet som ansvariga för strategier för stora data och öppna data i hela förvaltningen.

7.5 Kostnader för datatillgång

En fråga som kan ha stor betydelse för användningen av öppna data är huruvida data görs tillgängliga mot någon avgift eller om det är gratis för användaren. Vissa bedömare menar att det måste vara helt gratis för att kunna klassificeras som öppna data. I Sverige har e-delegationen till exempel i sin vägledning för vidareutnyttjande av offentlig information rekommenderat att myndigheter som regel ska avstå från att ta ut avgifter för utlämnandet av data om man inte har en särskild rättslig skyldighet att ta ut avgifter.

I de studerade länderna skiljer sig inställningen något i denna fråga. I Japan är brukaravgifter från användare av offentliggjorda data en lagstadgad del av vissa myndigheters finansiering. I kontrast till detta kan man nämna Storbritannien, där huvudregeln är att

utlämnande av data ska ske till marginalkostnadspriser, vilket för elektroniska data i maskinläsbart format i princip är gratis.

7.6 Integritet och säkerhet

Stora data och öppna data medför även en diskussion kring integritet och säkerhet. Japan tillämpar en försiktighetsprincip och den gällande lagstiftningen är inte tillåtande. Men försiktigheten med att öppna upp data hör också samman med inställningen att garantera att den är korrekt. Detta kan jämföras med utvecklingen i Storbritannien där inställningen är att allt ska ut i befintligt skick, dock förutsatt att data är anonymiserad.

För länder som tidigare inte har tillämpat offentlighetsprincipen handlar det också om att ändra både den formella lagstiftningen och förvaltningskulturen när data ska släppas och kommunikationen med medborgaren ska förbättras. Sydkorea har sett befintlig, och eventuellt föråldrad, lagstiftning som ett hinder för att implementera policy inom området – både för att offentliga organisationer ska kunna tillgängliggöra data, och för att näringslivet ska kunna använda data för att skapa nya tjänster. Även om dessa motiv varit framträdande vid översyn och revidering av lagstiftningen, så har detta även öppnat för en debatt om medborgarnas rätt till insyn i statsförvaltningen i allmänhet.

Här kan man ställa sig frågan om internationellt integrerade regelverk och standarder kommer att krävas då allt mer pekar mot att data kommer att cirkulera över landsgränser.

Det är lätt att anta principen att all data av princip ska offentliggöras, om så helt anonymiserad. Det finns starka argument för att skattefinansierad data, även till största möjliga mån, direkt och utan extra kostnad ska komma företag och individer till godo. En rad praktiska frågor uppstår dock direkt, särskilt när det gäller de kanske mest intressanta datamängderna (till exempel sjukvårdsdata) som i de flesta fall innehåller kopplingar till personliga karakteristika. Anonymisering måste göras effektivt, helst så nära realtid som möjligt, och med minsta möjliga risk för dataläckage. Detta kräver hög systemkompetens och löpande översikt, något som ansvariga aktörer tidigare inte behövt investera i eftersom principen har varit att hålla all data inom slutna system.

Genomgående i landfallen finns en förhoppning om att stora data ska kunna hjälpa till att sänka sjukvårdskostnaderna i samhället – samtidigt som patientdata är en av de mest svårhanterliga typerna av data att göra tillgänglig när den personliga integriteten ska säkras.

7.7 Attraktionskraft, synlighet och tillgänglighet

Det kan konstateras att vissa länder, såsom Storbritannien inom särskilt öppna data och Sydkorea speciellt inom stora data, anstränger sig för att synas och positionera sig inom fältet och bygga ett starkt nationellt varumärke. Denna ansats kan delvis förklaras med en önskan om att attrahera lätttrörliga multi-nationella företag.

Storbritannien betonar exempelvis att landet har unika förutsättningar för öppna data; hög IT-kompetens, stor offentlig sektor, etc. Sydkoreas traditionellt starka IT-industri, med stöd från regeringen, ser sig nu efter nya tillväxtmotorer inom mjukvaruutveckling och särskilt inom tjänstesektorn. Stora data ses som ett lovande område där landet inte har råd att ”missa tåget”.

8 Lärdomar för Sverige

Utifrån de olika exemplen på politiska initiativ kring öppna data och stora data vill Tillväxtanalys särskilt uppmärksamma några frågor, där det kan finnas inspiration och lärdomar för Sverige:

- Storbritannien lyfter fram att landet har unika förutsättningar för öppna data; hög IT-kompetens, omfattande offentlig sektor, etc. Detta är förutsättningar som i minst lika hög grad finns även i Sverige. Men saknas ”sense of urgency” i Sverige? I flera av de studerade länderna gör man bedömningen att staten spelar en viktig roll för att utveckla öppna data och stora data, åtminstone för att skapa förutsättningar för ett ekosystem där ny teknik och nya tjänster utvecklas. I Sverige är inställningen möjligen mer att det är näringslivets roll att skapa dessa på egen hand, och att statens roll är väsentligen mindre.
- Aktiviteter och tjänster förekommer i Sverige inom en rad olika sektorer, t ex öppna data inom sjuk- och hälsosektorn, personligt hälsokonto som kan göras tillgängliga för företag etc. Denna utveckling görs dock inte särskilt i explicita termer av ”öppna data”. Möjligen skulle en relativt billig reform/ satsning att mer medvetet samla initiativ i en portal och öka uppmärksamheten för frågan ge resultat.
- Samtliga av de studerade länderna har riktade insatser för att säkra den framtida tillgången på kompetens inom områden som är relevanta för en ökad användning av öppna data och stora data. En framtida kompetensbrist inom områden som statistik, databaser, matematik, bildbehandling identifieras av såväl de studerade länderna som av fristående analytiker som den kanske främsta flaskhalsen för utveckling av sektorn. Här kan Sverige sannolikt göra mer för att stärka såväl forskning som högre utbildning.
- I exempelvis Sydkorea och Storbritannien finns tydliga policyer och incitament för att tillgängliggöra offentliga data. Sydkorea har till och med särskilda stödinsatser för att hjälpa offentliga aktörer att göra sina data tillgängliga. Som en kontrast till detta är vissa offentliga aktörer i Sverige i sin finansiering delvis beroende av intäkter från att sälja tillgången till sina data.
- I flera av de studerade länderna ser man små och medelstora företag som nyckelspelare i att ta tillvara den innovationskraft som finns i ett bättre användande av öppna data och stora data. Här skulle Sverige kunna vinna på att genomföra särskilda riktade insatser för att informera om öppnandet av datamängder till Sveriges små och medelstora företag.
- Ett område som den offentliga sektorn tydligt har egen makt över är att föregå med gott exempel. I detta sammanhang handlar det exempelvis om att använda de nya möjligheterna hanteringen av stora data medför för att utveckla det evidensbaserade beslutsfattandet. Vidare kan den offentliga sektorn bidra med konkreta demonstrationer av ökad effektivitet som en smartare dataanvändning kan bidra till.

Tillväxtanalys, myndigheten för tillväxtpolitiska utvärderingar och analyser, är en gränsöverskridande organisation med 60 anställda. Huvudkontoret ligger i Östersund och vi har verksamhet i Stockholm, Brasilia, New Delhi, Peking, Tokyo och Washington D.C.

Tillväxtanalys ansvarar för tillväxtpolitiska utvärderingar och analyser och därigenom medverkar vi till:

- stärkt svensk konkurrenskraft och skapande av förutsättningar för fler jobb i fler och växande företag
- utvecklingskraft i alla delar av landet med stärkt lokal och regional konkurrenskraft, hållbar tillväxt och hållbar regional utveckling

Utgångspunkten är att forma en politik där tillväxt och hållbar utveckling går hand i hand. Huvuduppdraget preciseras i instruktionen och i regleringsbrevet. Där framgår bland annat att myndigheten ska:

- arbeta med omvärldsbevakning och policyspaning och sprida kunskap om trender och tillväxtpolitik
- genomföra analyser och utvärderingar som bidrar till att riva tillväxthinder
- göra systemutvärderingar som underlättar prioritering och effektivisering av tillväxtpolitikens inriktning och utformning
- svara för produktion, utveckling och spridning av officiell statistik, fakta från databaser och tillgänglighetsanalyser

Om Working paper/PM-serien: Exempel på publikationer i serien är metodresonemang, delrapporter och underlagsrapporter.

Övriga serier:

Rapportserien – Tillväxtanalys huvudsakliga kanal för publikationer.

Statistikserien – löpande statistikproduktion.

Svar Direkt – uppdrag som ska redovisas med kort varsel.