

# डाटा “लोगों का, लोगों द्वारा, लोगों के लिए”

“नदी पार करते समय उसके पत्थरों को स्पर्श करते चलो।”

—डेंग जियाओपेंग

एक अनिश्चित, अस्थिर दुनिया में आगे बढ़ते समय, रीयल टाइम संकेतकों का उपयोग करते हुए अर्थव्यवस्था के गमन-मार्ग की सतत् निगरानी की आवश्यकता होती है। इस प्रक्रिया में डाटा की भूमिका नदी में पड़े उन पत्थरों के समान है जो नदी पार करते व्यक्ति का मार्गदर्शन करते हैं। हाल के वर्षों में आंकड़ों (डाटा) की उपलब्धता में आशातीत वृद्धि के फलस्वरूप आंकड़ों की सीमांत लागत में घातांकी कमी आई है और इनसे समाज के सीमांत लाभों में कई गुना वृद्धि हुई है। इसलिए सामाजिक स्तर पर आंकड़ों का उपयोग अब पहले की तुलना में अधिक हो रहा है। जहां निजी क्षेत्र लाभकारी आंकड़ों से लाभ लेते हुए अच्छा कार्य कर रहा है वहीं दूसरी ओर सरकार को भी कुछ ऐसे सामाजिक क्षेत्रों में हस्तक्षेप करने की जरूरत है जहां आंकड़ों के संदर्भ में निजी क्षेत्र की ओर से किया जाने वाला निवेश अपर्याप्त है। सरकार के पास पहले से ही नागरिकों के बारे में प्रशासनिक, सर्वेक्षण, संस्थागत और संव्यवहार संबंधी आंकड़े उपलब्ध हैं, किंतु ये आंकड़े अनेक निकायों में प्रकीर्ण रूप में पड़े हुए हैं। इन असंयोजित डाटा-समूहों को एक साथ संयोजित कर देने से सरकार को होने वाले अन्य लाभों के साथ-साथ, नागरिकों की जीवन-सुविधा में सुधार करने, वास्तविक प्रमाणाधारित नीति तैयार करने, कल्याण स्कीमों के लक्ष्यनिर्धारण में सुधार करने, अपूर्ण जरूरतों का पता लगाने, विखंडित बाजारों का एकीकरण करने, लोक सेवाओं में जवाबदेही का स्तर बढ़ाने, शासन में नागरिकों की आर्थिक भागीदारी सुनिश्चित करने, इत्यादि, में सहायता मिलेगी। गोपनीय सूचनाओं का संरक्षण करने तथा इनका आदान-प्रदान करने के लिए पहले ही उन्नत प्रौद्योगिकी की उपलब्धता सुनिश्चित करते हुए, सरकारें आंकड़ों की निजता (डाटा प्राइवसी) के कानूनी ढांचे के अंतर्गत लोक कल्याणकारी साधन के रूप में आंकड़ों (डाटा) का सृजन कर सकती हैं। भारत के संविधान की मूल भावना के अनुसार, ये आंकड़े (डाटा) “लोगों के, लोगों द्वारा, लोगों के लिए” होने चाहिए।

## आंकड़ों ( डाटा ) और सामाजिक कल्याण का अर्थशास्त्र

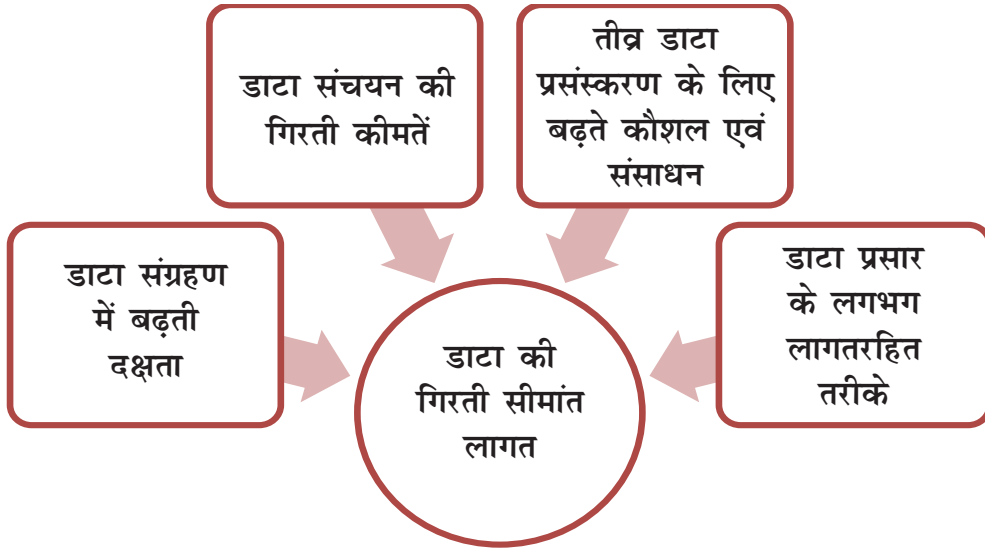
4.1 हाल के वर्षों में, विश्व सूचनाओं के विस्फोट, यानी प्रकाशित आंकड़ों के परिणाम में चरघातांकी वृद्धि का साक्षी रहा है। चूंकि लोग एक दूसरे से बात करने, सूचनाएं प्राप्त करने, वस्तुओं एवं सेवाओं का क्रय करने, बिलों का भुगतान करने, वित्तीय बाजारों में लेन-देन करने, टैक्स फाइल करने, कल्याणकारी सेवाओं का लाभ लेने तथा स्थानीय नेताओं से संपर्क रखने के लिए डिजिटल सेवाओं का भरपूर उपयोग कर रहे हैं अतः एक अभूतपूर्व पैमाने

पर आंकड़ों (डाटा) का सृजन किया जा रहा है। वैश्विक डाटा अवसंरचना आंकड़ों की इस भरमार को संभालने के लिए व्यापक रूप से विश्वसनीय, तीव्र और सुरक्षित सिद्ध हुई है। (मैकिसे, 2011)।

4.2 इस डाटा-विस्फोट के अनुरूप, आंकड़ों की सीमान्त लागत में घातांकी कमी आई है और इसके उपयोग से समाज के सीमान्त लाभों में कई गुना वृद्धि हुई है।

4.3 जब लोग अपनी दिन-प्रतिदिन की गतिविधियों को ऑनलाइन तरीके से करते हैं तो वे इन गतिविधियों के

चित्र 1: डाटा ( आंकड़ों ) की गिरती हुई सीमान्त लागत



डिजिटल फुटप्रिंट भी छोड़ते हैं। इसको दूसरे तरीके से देखें तो लोग अपने बारे में, प्रतिदिन आंकड़ों का सृजन करते हैं और अपनी इच्छानुसार इन आंकड़ों को सरकारी और निजी सर्वरों पर डालते हैं। कुछ दशक पूर्व तक, जिन आंकड़ों को प्राप्त करने में एक श्रमसाध्य सर्वेक्षण करना पड़ता था वे आज शून्य लागत पर संचित हो रहे हैं, यद्यपि वे विविध स्रोतों में प्रकीर्ण रूप में हैं।

4.4 वस्तुतः प्रत्येक व्यक्ति डिजिटल अर्थव्यवस्था में हिस्सा नहीं लेता है। बड़ी संख्या में गरीब व्यक्तियों के पास अभी भी कोई डिजिटल फुटप्रिंट नहीं होता है। उन व्यक्तियों के मध्य जो इनसे संबंधित होते हैं, ऑनलाइन निष्पादित गतिविधियों की रेंज पूर्णतः सीमित होती है। तथापि, डाटा एकत्रित करने की लागत अभी उस लागत से काफी कम है जो कतिपय दशकों पहले हुआ करती थी। यद्यपि प्रत्येक घर का सर्वेक्षण कुछ खास किस्म के डाटा एकत्रित करने का एक मात्र तरीका है, फिर भी हम एक थकाऊ डाटा प्रविष्टि प्रक्रिया द्वारा अपनाए गए अन्यथा श्रमसाध्य कागजात आधारित सर्वेक्षण को रोकते हुए, वास्तविक समय में डाटा ऑनलाइन लॉग करने की सस्ती तकनीकों को प्रस्तुत करते हैं।

4.5 डाटा एकत्रित करने की घटती हुई लागत के साथ, भंडारण लागतों में भी निरंतर कमी आई है। प्रति गीगाबाइट भंडारण लागत वर्ष 1981 में 500,000 अमेरिकी डालर से गिरकर अद्यतन स्थिति के अनुसार 0.05 अमेरिकी डालर से भी कम रह गई है। तथापि,

डाटा अतिरेक और इसे भंडारित करने की असीमित क्षमता तब तक उपयोगी नहीं हो सकती है जब तक कि कोई डाटा की इन भारी भरकम मात्राओं की युक्तिसंगत समय में जानकारी हासिल न कर सके। सौभाग्य से, डाटा प्रक्रमणित करने की मानवीय और तकनीकी पूंजी डाटा-अधिष्ठापन के साथ-साथ विकसित हुई है। डाटा संबंधी विज्ञान एक अलग, निधि संपन्न अध्ययन क्षेत्र के रूप में उभरा है जिससे डाटा का कार्यक्षम प्रयोग करने के लिए निरंतर नवाचारी तरीके प्राप्त हो रहे हैं विश्लेषण विज्ञान में पाठ्यक्रम देशव्यापी बन गए हैं। बढ़ती हुई संख्या में व्यक्तियों को बृहद डाटासेटों को संभालने का कौशल प्राप्त कराया गया है। डाटा प्रक्रमण के लिए अभी भी अपेक्षाकृत अधिक खर्चीला है क्योंकि इसमें मुखर, विषमरूपक और समग्र स्रोतों से असंगत होने की प्रकृति होती है, किन्तु प्रौद्योगिकी से इन समस्याओं के समाधान निरंतर निकल रहे हैं।

4.6 एक बार प्रक्रमणित हो जाने पर, इसकी पूरी जानकारी प्रसारित करने की लागत नगण्य होती है—अर्थात् इसकी कोई लागत नहीं आती है और सूचना को इंटरनेट के जरिए भेज दिया जाता है तथापि, डाटा प्रसारण में एक अन्य लागत शामिल होती है—जोकि डाटा की गोपनीयता और सुरक्षा सुनिश्चित करने की लागत है। डाटा को प्रसारित किए जाने से पहले, इसको वैयक्तिक पहचान सूत्रों से मुक्त और समेकित किया जाना आवश्यक होता है। यह एक प्रत्यक्ष लागत है, इसके अलावा एक अप्रत्यक्ष लागत भी है जो कि डाटा के दुरुपयोग की लागत होती है।

डाटा के अचानक उजागर हो जाने से कानूनी जटिलताएं और पर्याप्त वित्तीय विवक्षाएँ उत्पन्न होती हैं। तथापि, इन जोखिमों का शमन करने के लिए प्रौद्योगिकी (तकनीकी) में बड़े पैमाने पर कदम उठाए गए हैं।

4.7 इसके साथ-साथ, डाटा एकत्रित करने, भंडारण, प्रक्रमण और प्रसारण में समुन्नति होने से डाटा की सीमांत लागत में अप्रत्याशित स्तरों तक कमी आई है। चित्र एक इस घटना को संक्षिप्त रूप में प्रस्तुत करता है। साथ-साथ, डाटा का सीमांत लाभ अब तक का उच्चतम लाभ है। एक जिला शिक्षा अधिकारी को यदि अपने जिले के प्रत्येक स्कूल की, छात्रों और शिक्षकों की उपस्थिति दरों, औसत परीक्षा प्राप्तांकों और स्कूल के शौचालयों की स्थिति के बारे में जानकारी हो, तो वह बेहतर निर्णय ले सकता है। इसी प्रकार, यदि माता-पिता को अपने ग्राम में शिक्षकों की औसत अनुपस्थिति दर की जानकारी हो, तो वे इस बारे में बेहतर निर्णय ले सकते हैं कि उन्हें अपने बच्चों को किस स्कूल में दाखिल करना है और वे अनुपस्थिति दर की आसपास के ग्राम से तुलना कर सकते हैं। ऐसे अनेक दृश्यविधान मौजूद हैं जिनमें डाटा की सीमांत यूनिट को काम में लाने से लोक कल्याण में तीव्र वृद्धि हो सकती है।

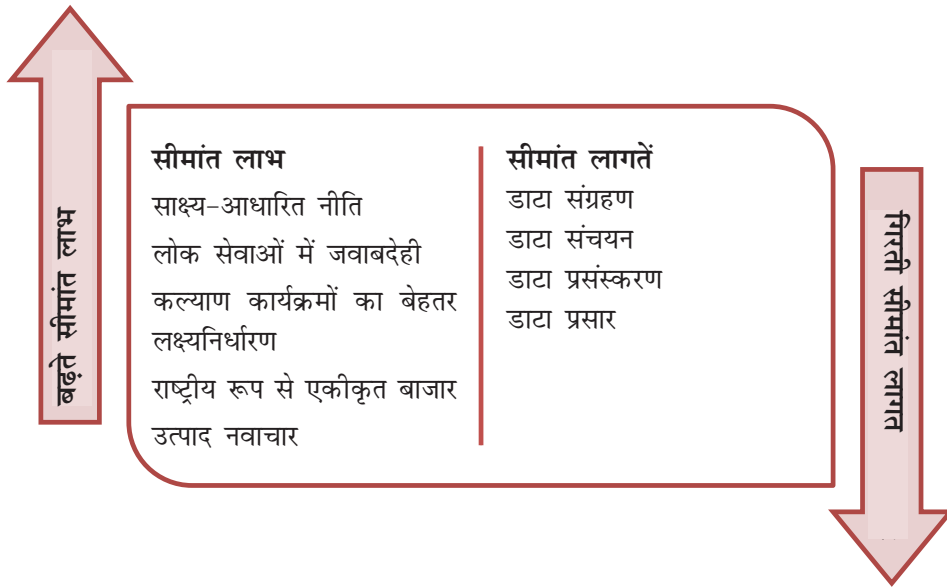
4.8 उस डाटा के बारे में विचार करें जिसे सरकार अपने नागरिकों के बारे में रखती है। वर्तमान समय में, अधिकतर डाटा विभिन्न मंत्रालयों द्वारा रखी गई विभिन्न रजिस्ट्री से प्रसारित किया जाता है। ऐसा इसलिए होता है क्योंकि प्रत्येक बार किसी नागरिक को किसी नई सेवा तक पहुंचना होता है और व्यक्तियों को अपनी पहचान सिद्ध करने के लिए और प्रक्रिया संबंधी अपना दावा सिद्ध करने के लिए सभी दस्तावेज एकत्रित करने के लिए कहा जाता है। उदाहरण के लिए, किसान को देय, निःशुल्क बिजली जैसी सब्सिडी अथवा लाभ लेने के लिए, दस्तावेजों की अपूर्ण सूची में ग्राम प्रशासन अधिकारी द्वारा जारी स्वामित्व प्रमाणपत्र, चित्त आदंगल (भूमि रजिस्ट्र से सार), पट्टा (अधिकारों का अभिलेख) अथवा बिक्री विलेख, पास की किसी सरकारी परियोजना से अनापत्ति प्रमाणपत्र और यह सिद्ध करने के लिए कि वह उस भूमि का कृषक है, अन्य दस्तावेज शामिल हैं। इस अनुप्रयोग में सन्निहित प्रयासों में अधिकांशतः किसी सरकारी विभाग से कोई नई जांच पड़ताल अथवा निरीक्षण किया जाना शामिल नहीं होता है। नागरिकों को अपनी हकदारी के लाभों को प्रकट करने के लिए सरकारी

तंत्र के अंदर दस्तावेजी फाइलों में बंद डाटा प्राप्त करने के लिए असुविधा का सामना करना पड़ता है। सरकार विविध मंत्रालयों में बिखरे हुए डाटा सेट को समेकित करके नागरिकों को बेहतर अनुभूति प्रदान कर सकती है।

4.9 यदि डाटा को मिला दिया जाता है, तो ऐसे डाटा से कल्याणकारी स्कीमों को लक्षित करने की कठिनाई को कम किए जाने की संभावना होती है। उदाहरण के लिए, एक ऐसे काल्पनिक व्यक्ति को लें जो कार रखने में पर्याप्त सक्षम है किन्तु वह बीपीएल कल्याणकारी योजनाओं का लाभ उठाने में समर्थ हो जाता है, भले ही वे अनुचित हों। यदि डाटासेट असंयोजित होते हैं तो वाहन रजिस्ट्री से सार्वजनिक वितरण रजिस्ट्री के बारे में पता नहीं चलता है। परिणामस्वरूप, सार्वजनिक वितरण प्रणाली से इस व्यक्ति को गलत रूप में सब्सिडी दिया जाना जारी रखा जाता है। तथापि यदि दो डाटासेटों को एकीकृत किया जाता है, तो इस प्रकार के समावेशी उपाय से गलतियों को न्यूनतम किया जा सकता है, और बहुमूल्य सरकारी संसाधनों की बचत की जा सकती है। इसी प्रकार से बहिष्कार त्रुटियों का परिशोधन किया जा सकता है।

4.10 वस्तुतः डाटा की गिरती हुई लागत से ऐसे नए लाभों की उत्पत्ति हुई है जो कुछ वर्ष पहले उपलब्ध नहीं थे। उदाहरण के लिए ऐप को ही लें जो किसानों को संपूर्ण देश में उत्पादों की कीमतों की जानकारी देता है। एक ऐसा समय था जब, इस ऐप के होते हुए भी यह अनुपयोगी रही होगी और यह जानना कि अन्य जिले में कीमतें उच्चतर हैं, किसान के लिए महत्वपूर्ण बात नहीं रही होगी क्योंकि उसके पास परिवहन, भंडारण और वितरण में अत्यधिक कीमत (दर्जनों बिचौलियों की लागत सहित) चुकाए बिना इन तक पहुंचने के लिए कोई रास्ता नहीं रहा होगा। लेकिन आजकल, ई-नाम इलेक्ट्रॉनिक राष्ट्रीय कृषि बाजार जैसे प्लेटफार्मों की सहायता से, हमारे पास ऐसी तकनीकें हैं जिनकी सहायता से किसान अपनी बिक्री ऑनलाइन कर सकते हैं और लगातार डिलीवरी करके उच्चतर कीमतें प्राप्त कर सकते हैं। इस प्रकार से, डाटा में बाजारों को राष्ट्रीय स्तर पर एकीकृत करने, बिचौलियों की आवश्यकता को कम करने, अंतिम उपभोक्ताओं के लिए कीमतों को घटाने और किसानों के लिए कीमतों को बढ़ाने की सामर्थ्य होती है। चित्र 2 में इन लागतों और लाभों का संक्षिप्त विवरण दिया गया है।

चित्र 2: बढ़ते हुए सीमांत लाभ और डाटा की घटती हुई सीमांत लागत



### डाटा को सार्वजनिक संपत्ति क्यों माना जाना चाहिए?

4.11 सीमांत लाभ में वृद्धि के साथ किए गए डाटा की सीमांत लागत में गिरावट का अर्थ है कि डाटा की वह अनुकूलतम मात्रा जिसका उपयोग समाज को करना चाहिए, पहले से काफी अधिक है। यदि ऐसा है, तो आर्थिक सिद्धांत से अनुमान लगाया जाता है कि अर्थव्यवस्था में अब तक, डाटा को काम में लाने और प्रयोग करने के लिए किए गए प्रयासों में वृद्धि देखी जानी चाहिए। वास्तव में ऐसा हुआ है, किन्तु यह आंशिक रूप से ही संभव हो सका है।

4.12 डाटा संबद्ध प्रयासों में निजी क्षेत्र का निवेश पहले सदैव के निवेश से उच्चतर है। फोर्ब्स सर्वे 2017 में यह पाया गया है कि 53 प्रतिशत कंपनियों ने निर्णय लेने में बड़ी मात्रा में डाटा का सक्रिय रूप से प्रयोग किया है (ड्रैसनर, 2017)। सभी कंपनियों में रूझान स्वास्थ्य देखभाल और वित्तीय सेवाओं की तरह ही विषम है और भौगोलिक स्थिति व कंपनी आकारों के अनुसार है। वस्तुतः पिछले दो दशकों में विश्व फेसबुक, ऐमोजोन, इंस्टाग्राम आदि जैसी कंपनियों के प्रादुर्भाव का साक्षी रहा है जो केवल व्यक्तियों के डाटा से राजस्व अर्जित करती हैं।

4.13 किन्तु, ऐसे अनेक क्षेत्र हैं जहां डाटा को व्यापक रूप में काम में नहीं लाया जाता है। उदाहरण के लिए कृषि बाजार पर विचार करें। यदि कीमत सूचना पाने वाले

किसी किसान का सीमांत लाभ उस सूचना की सीमांत लागत से उच्चतर है, तो वह उस सूचना के लिए भुगतान करेगा। परिणामस्वरूप, निजी क्षेत्र व सूचना एकत्रित करके और उसकी उसे बिक्री करके जिसे वह चाहता है, उसकी आवश्यकता की व्यवस्था करेगा। ऐसा करने से अंततः एक मूल्य के साथ राष्ट्रीय रूप से एकीकृत कृषि बाजार अस्तित्व में आएगा। हालांकि भारत में अभी तक राष्ट्रीय स्तर पर एकीकृत कृषि बाजार नहीं है। ऐसा बाजार भारत में हो जाना चाहिए यदि आज डाटा का सीमांत हितलाभ वास्तव में लागत से अधिक है। जो प्रवाह कारपोरेट क्षेत्र में चल रहा है वह कृषि क्षेत्र में क्यों नहीं हो पा रहा है।

4.14 अभी तक आंकड़ों के जिस अर्थशास्त्र पर विचार किया गया है, वह यह है कि वे आंकड़े समाज से संबंधित होते हैं, न कि किसी व्यक्ति विशेष से, वह भले ही किसान या कोई फर्म हो। किसी निजी कंपनी के पास उपलब्ध डाटा जो कि संसाधित करके कृषकों को उपलब्ध कराया जाता है वह उस डाटा से भिन्न होता है, जोकि “सोशल प्लेनरों” के पास उपलब्ध रहता है। निजी फर्म के लिए, राष्ट्रीय एकीकृत कृषि बाजार, और समाज कल्याण के परिणाम की संभावना, एक सकारात्मक बाह्यता होती है। व्यवहारिक तौर पर यह कोई ऐसा लाभ नहीं है जो कि फर्म को प्राप्त होता है क्योंकि फर्म अर्थव्यवस्था में विभिन्न एजेंटों जो कि उक्त से लाभान्वित होंगे, कोई प्रभार नहीं ले पाती। इसलिए, निजी फर्म को प्राप्त होने वाला सीमांत लाभ

समाज के सीमान्त लाभ से कम होता है। क्योंकि फर्म समाज कल्याण की प्रसुविधा को आंतरिक नहीं करती है, इसलिए फर्म डाटा की इष्टतम मात्रा सामाजिक इष्टतम से कम हो जाएगी।

4.15 दूसरे, डाटा कई रूपों में आता है जिसमें से प्रत्येक रूप एक अलग प्रसुविधा प्रदान करता है। किसी व्यक्ति विशेष का डाटा लिंक उसकी अत्यंत निजी जानकारी—जैसे कि उनके बायोमीट्रिक विवरण, से लेकर अत्यंत सार्वजनिक स्थिति—जैसे उनका नाम तक हो सकता है। इसमें मानव क्रियाकलापों से जनित, विश्लेषणों से व्युत्पन्न डाटा शामिल होते हैं जिनमें साधारणतः एक एल्गोरिथम भी शामिल होता है। उदाहरण के लिए क्या किसी व्यक्ति ने करों का भुगतान किया है अथवा नहीं, यह घटक जनरेट किया गया डाटा प्वाइंट होता है। किंतु, कर अभिलेखों और अन्य डाटा बिंदुओं का उपयोग करते हुए, क्रेडिट ब्यूरो किसी व्यक्ति को क्रेडिट स्कोर दे सकता है, जो व्युत्पन्न डाटा बिंदु होता है। डाटा जो किसी विशिष्ट व्यक्ति से संयोजित नहीं किए जाते हैं किंतु कि किसी विशेष व्यक्ति से असम्बद्ध डाटा में भी एक व्यक्ति के स्तर का सूक्ष्मतरंग विवरण भी उपलब्ध रहता है उसे अनाम डाटा कहते हैं। चिकित्सा अनुसंधान जैसे कुछ क्षेत्रों में अनाम डाटा महत्वपूर्ण होते हैं। जो डाटा किसी व्यक्ति से लिंक नहीं होते हैं या न ही सूक्ष्मता के वैयक्तिक स्तर पर संयोजित होते हैं ऐसे डाटा को सरकारी डाटा कहते हैं जैसे जनगणना।

4.16 जबकि प्राइवेट सेक्टर ने उसे डाटा को काम में लाने का प्रभावशाली कार्य किया है जिसे प्राइवेट हितलाभ में परिवर्तित किया जा सकता है—अन्य क्षेत्रों में सरकारी हस्तक्षेप अपेक्षित होता है जहां पर डाटा में प्राइवेट निवेश अपर्याप्त रहता है। अर्थव्यवस्था के सामाजिक क्षेत्र यथा शिक्षा और स्वास्थ्य देखभाल ने डाटा के संदोहन में वाणिज्यिक क्षेत्रों को पीछे छोड़ दिया है। क्योंकि प्राइवेट क्षेत्र इन क्षेत्रों में डाटा के सामाजिक लाभों को आंतरिक नहीं कर सकता है, इसलिए इन क्षेत्रों में डाटा के लिए बाजार अभी तक विफल रहा है।

4.17 यह सुनिश्चित करने के लिए कि सामाजिक रूप से इष्टतम मात्रा में डाटा तैयार किया जाता है और उपयोग में लाया जाता है, सरकार डाटा की प्रकृति और संवेदनशीलता पर निर्भर करते हुए, या तो स्वयं डाटा उपलब्ध कराती है या प्राइवेट क्षेत्र द्वारा अनुपयुक्त प्रोत्साहन संरचना को ठीक करने के लिए, पहल करती है। वास्तव में, कृषि क्षेत्र में, सरकार ने ई-नाम इलेक्ट्रॉनिक नेशनल एग्रीकल्चर मार्केट का सटीक निर्माण किया है,

क्योंकि इस बात की संभावना नहीं है कि प्राइवेट क्षेत्र इस प्रकार के समाधान तैयार करेगा।

## लोकोपयोगी तौर पर डाटा सृजन करते हुए डाटा गोपनीयता सुनिश्चित करना

4.18 डाटा को लोकापयोगी बनाने के प्रयास में, डाटा के प्रकार के आधार पर, निजता के निहितार्थ और उपयोग किए जा रहे डाटा की अंतर्निहित निष्पक्षता पर विचार करना बहुत महत्वपूर्ण होता है। कहने की आवश्यकता नहीं है, कि अंतरंग डाटा की निजता सुनिश्चित करने के लिए अपेक्षित प्रक्रिया उपनाम या सरकारी डाटा के लिए अपेक्षित प्रक्रिया से भिन्न होती है। डाटा के इन विविध प्रकारों के निपटान में सबसे प्रमुख अंतर डाटा सिद्धांत की जानकारी और सहमति देना है। इस अध्याय में जिस डाटा के बारे में बात की गई, उस डाटा के बारे में भले ही स्पष्टतः वर्णन न किया गया हो, तो भी इसके बारे में ऐसा अनुमान किया जाता है कि डाटा संसाधन, स्वीकृत गोपनीयता मानदंडों और आगामी गोपनीयता विधि, जिसे वर्तमान में संसद पटल पर पेश किया गया है, के अनुपालन में किया जाएगा।

## डाटा सृजन प्रक्रिया में पैमाने और वितान की मितव्ययताएं

4.19 इस स्पष्ट आवश्यकता के अतिरिक्त कि डाटा शुद्ध होने चाहिए, किसी सरकार नियंत्रित डाटा क्रांति की आवश्यकता उन तीन मुख्य विशेषताओं से प्रेरित होती है जो प्राप्त होने वाले सहक्रियात्मक लाभों के लिए किसी डाटा में आवश्यक रूप से होनी चाहिए। डाटा सृजन प्रक्रिया विशेष रूप से पैमाने और वितान की महत्वपूर्ण मितव्ययताएं प्रदर्शित करती है।

4.20 सर्वप्रथम जब डाटा की बात आती है तो डाटा में इसके भागों को मिला देने से यह काफी बड़ा हो जाता है अर्थात् जब अन्य डाटा के साथ मिला दिया जाता है तो यह अत्यधिक उपयोगी हो जाता है। उदाहरण के लिए विभिन्न सरकारी एजेंसियों भिन्न-भिन्न डाटा सेटों का विलय जैसे कि वित्तीय सेवाएं विभाग, वित्त मंत्रालय के जन-धन खाते से लिए गए लेन-देन डाटा का उपयोग ग्रामीण विकास मंत्रालय के अधीन मनरेगा के लिए किया गया। चूंकि ग्रामीण संकट के वास्तविक (रियल टाइम) जिस पर इस सतीक्षा के इस खंड के अध्याय 10 में चर्चा की गई है। संकेतक रूप में मनरेगा का नाम लिया जा सकता है, अतः जन-धन खातों के लेन-देन आंकड़ों का उपयोग करते हुए प्राप्त क्रेडिट का उपयोग संकट



का सामना कर रहे जिलों/पंचायतों में क्रेडिट प्रदान करने के लिए किया जा सकता है। असमान डाटासेट का ऐसा संयोजन कल्याणकारी नीतियों को डिजाइन और कार्यान्वित करने के लिए आवश्यक क्षमता प्राप्त करने में अत्यंत उपयोगी हो सकता है।

4.21 द्वितीय, आंकड़े पर्याप्त संख्या में व्यक्तियों/फर्मों के होने चाहिए ताकि उपयोगी नीतिगत अंतर्दृष्टि विकसित करने के लिए व्यक्तियों/फर्मों की तुलनाओं एवं सहसंबंधों का आंकलन किया जा सके। अतः पूरे देश के विभिन्न उत्पाद बाजारों से किसी व्यवसाय की कीमतों से संबंधित आंकड़ों को एकत्र करने के लिए बहुत बड़ी संख्या में उत्पादकों और खरीदारों को रियल टाइम में एक मंच पर अपने लेन-देन को लॉग इन करने की आवश्यकता होती है। नियमित रूप से लेन-देन की जानकारी की रिपोर्ट करने के लिए कई एजेंटों को प्रेरित करना एक ऐसा कार्य है, जिसमें महत्वपूर्ण प्रारंभिक निवेश की आवश्यकता होती है, जो निजी क्षेत्र के लिए निषेधात्मक साबित हो सकता है।

4.22 अंत में आंकड़ों के लिए एक पर्याप्त समय श्रृंखला होनी चाहिए ताकि नीति निर्माण के लिए गतिशील प्रभावों का अध्ययन करते हुए नियोजित किया जा सके। नीतियों की प्रभाविता के आंकलन हेतु ऐसे पूर्व-उत्तरवर्ती मूल्यांकनों के लिए वे आंकड़े महत्वपूर्ण होते हैं, जो एक काफी लंबी अवधियों तक विस्तृत होते हैं।

4.23 आंकड़े जिसमें तीनों विशेषताएं होती हैं, तीन अलग-अलग और असंगत डाटा सेटों की तुलना में अधिक मूल्यवान हैं, जिनमें अलग-अलग विशेषता होता है। अतः आंकड़ों के निर्माण की प्रक्रिया से संभावनाओं की महत्वपूर्ण मितव्ययता प्रकट होती है। इसके अलावा, ऐसे डाटा के निर्माण के लिए उच्चतम स्तर के प्रयास की जरूरत होती है जिसमें सभी तीन विशेषताएं मौजूद हों, अर्थात् डाटा के निर्माण की प्रक्रिया से महत्वपूर्ण पैमाने की मितव्ययताओं की गहराईयों का पता चलता है। साथ ही उक्त डाटा के निर्माण में संलग्न (अग्रिम) निर्धारित लागत का भी पता लग जाता है जो कि अधिक महत्वपूर्ण हैं। निजी क्षेत्र के पास जोखिम उठाने की क्षमता नहीं है या निजी क्षेत्र के पास इतनी पूंजी नहीं है कि वे तीनों विशेषताओं अर्थात् असंगत डाटासेटों का संयोजन करना, व्यक्तिगत/फर्मों के पास पर्याप्त संख्या में आंकड़ों का होना, और काफी लंबी अवधि के आंकड़ों को स्टोर करने-की क्षमता से युक्त डाटा का निर्माण करने

के लिए अपेक्षित पूंजी निवेश कर सकें। यदि किसी भी तरह से निजी क्षेत्र उक्त समृद्ध डाटा का संयोजन एक साथ कर देते हैं तो इससे भविष्य में एकाधिकार की समस्या पैदा हो सकती है जिससे एक तरफ तो नागरिकों के कल्याणकारी कार्य प्रभावित होंगे तो दूसरी तरफ इस कदम से—“डाटा नागरिकों के लिए” के सिद्धांत का उल्लंघन होगा।

4.24 सबसे महत्वपूर्ण बात यह है कि डाटा सार्वजनिक वस्तुओं की कुछ विशेषताओं को वहन करता है। यह गैर-ऋणात्मक है अर्थात् किसी व्यक्ति द्वारा डाटा का उपभोग करने पर अन्य लोगों के लिए उपलब्ध डाटा की मात्रा कम नहीं होती है। सिद्धांत रूप में डाटा को बाह्य रखा जा सकता है अर्थात् डाटा एक्सेस करने से लोगों को बाहर करना संभव है क्योंकि डाटाबेस कम्पनियां पे वॉल के जरिए ऐसा करती हैं। हालांकि कुछ ऐसे भी डाटा हैं—विशेषकर सामाजिक हितों के मुद्दों पर सरकारों द्वारा एकत्र किए गए डाटा—जिन्हें सामाजिक कल्याण के हित में लोकतांत्रिक बनाया जाना चाहिए। इस तरह के डाटा को सार्वजनिक साधन के रूप में इस्तेमाल करने पर बढ़ावा दिया जाना चाहिए। चूंकि निजी क्षेत्र किसी अक्षय, अबहिष्करणीय कल्याणकारी डाटा को इष्टतम परिमाण में उपलब्ध कराने में सफल नहीं हो पाएगा, अतः इस संबंध सरकारी हस्तक्षेप की आवश्यकता है।

4.25 डाटा का सृजन लोगों के द्वारा लोगों के संबंध में और लोगों के लिए होना चाहिए। लोककल्याणकारी साधन के रूप में डाटा को लोकतांत्रिक स्वरूप प्रदान किया जा सकता है और सर्वोत्तम संभव उपयोग हेतु उपलब्ध कराया जा सकता है। बॉक्स 1 में सरकार द्वारा की गई ओपन गवर्नमेंट डाटा पहल का उल्लेख किया गया है जो लोक कल्याणार्थ डाटा की मूल भावना का एक उदाहरण है। चूंकि यह एक उत्कृष्ट प्रारंभ है, अतः लोक कल्याण के साधन के रूप में डाटा के उपयोग से जो व्यापक लाभ प्राप्त किए जा सकते हैं वे इस बात में निहित हैं कि सरकार इस दिशा में दोहरी ताकत से प्रयास करें।

## सिस्टम का निर्माण

4.26 परिकल्पित डाटा सिस्टम में मुख्य रूप से उस डाटा को शामिल किया जाता है, जिसे लोग पूर्ण सूचना और सहमति के आधार पर सरकारी निकायों के साथ साझा करते हैं अथवा उस डाटा को शामिल किया जाता

### बॉक्स 1: ओपन गवर्नमेंट डाटा और नागरिकों का जुड़ाव

केंद्र सरकार के ओपन गवर्नमेंट डाटा प्लेटफार्म के जरिए नागरिक एक ही स्थान पर मशीन रीडेबल रूप में कई तरह के सरकारी डाटा प्राप्त कर सकते हैं। पोर्टल केन्द्रीय मंत्रालयों और विभागों को सार्वजनिक उपयोग के लिए उनके द्वारा एकत्र किए गए डाटासेट, दस्तावेज सेवाओं, उपकरणों और अनुप्रयोगों को प्रकाशित करने की अनुमति देता है। वे डाटासेट्स जिनमें गोपनीय जानकारी होती है, को छोड़कर अन्य समस्त डाटासेट आम लोगों को उपलब्ध कराये जाते हैं जिसमें कल्याणकारी योजनाओं के डाटा से लेकर व्यापक आर्थिक संकेतक तक शामिल होते हैं। प्लेटफार्म में नागरिकों के जुड़ाव से संबंधित टूल्स जैसे फीड बैक फार्म, डाटा विजुअलाइजेशन, एप्लिकेशन प्रोग्रामिंग इंटरफेस (एपीआई) आदि सम्मिलित हैं।

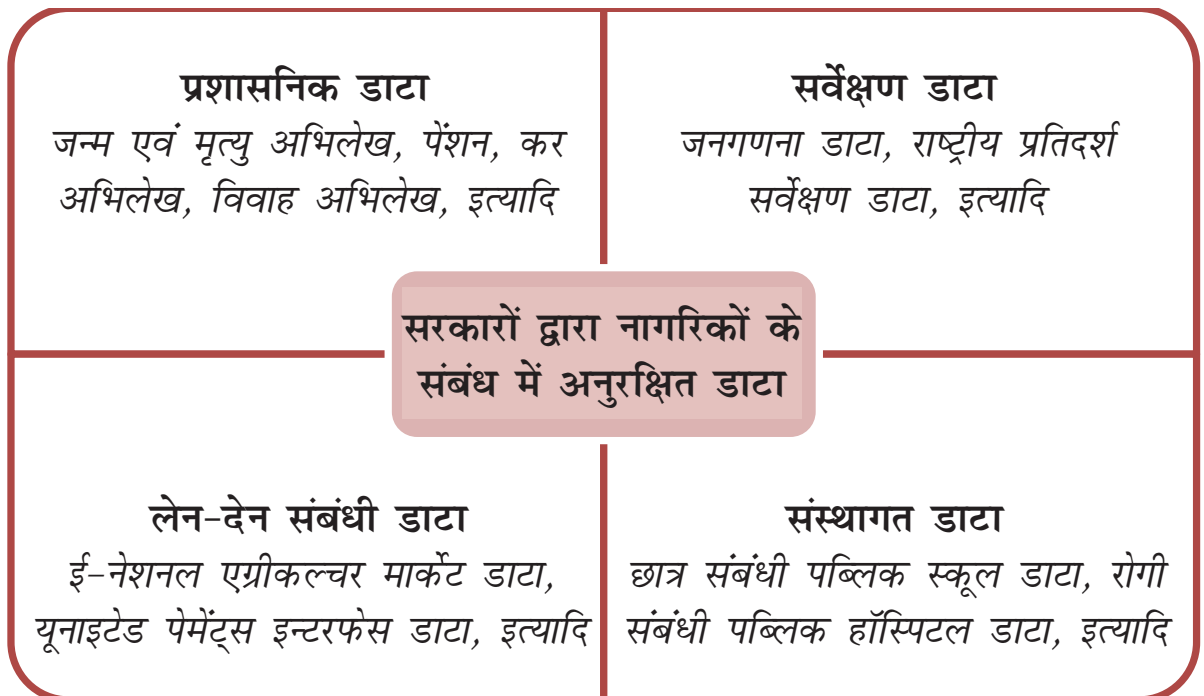
ओपन डाटा न केवल सरकारी अधिकारियों को बेहतर निर्णय लेने में मदद करता है, बल्कि लोगों की समस्याओं को हल करने में भी मदद करता है। जनता को ओपन डाटा उपलब्ध कराने से लोगों में डाटा का विश्लेषण करने और डाटा से दूरदर्शिता की सीख प्राप्त करने वालों की संख्या में अत्यधिक बढ़ोतरी हुई है। फलस्वरूप डाटा की उपयोगिता स्वतः ही बढ़ जाती है।

मानव संसाधन विकास मंत्रालय ने हाल ही में लोगों की समस्याओं को सार्थक तरीके से सुलझाने के लिए स्मार्ट इंडिया हैकथॉन की शुरुआत की है—जो कि एक खुला नवाचार मॉडल है। इस प्रौद्योगिकी के जरिए भारत की ज्वलंत समस्याओं को सुलझाया जा सकता है। स्मार्ट इंडिया हैकथॉन उत्पाद विकास प्रतियोगिता है जिसमें प्रतिभागियों को प्राब्लम स्टेटमेंट और संगत डाटा मिलता है, जिसके उपयोग से वे एक प्रोटोटाइप सॉफ्टवेयर या हार्डवेयर विकसित करते हैं। ये प्रतियोगिताएं क्राउड-सोर्स सोल्यूशन (भीड़ तन्त्र से समाधान जुटाने) के तौर पर उपयोग में लायी जाती हैं ताकि गवर्नेंस को बेहतर बनाया जा सके और कल्याणकारी स्कीमों की प्रभावोत्पादकता को बढ़ाया जा सके। निःसंदेह विश्वसनीय डाटा के बिना इसमें से कुछ भी संभव नहीं है।

है जिसे राज्य द्वारा कर संग्रह, या कल्याणकारी वितरण जैसे स्पष्ट प्रयोजन के लिए संग्रह करने की विधिक मंजूरी प्रदान की गई हो। भारत सरकार लोगों के बारे में डाटा के चार अलग-अलग सेट रखती है—प्रशासनिक,

सर्वेक्षण, संस्थागत और लेनदेन डाटा (सारांश के लिए चित्र-3 और विवरण के लिए बॉक्स-2 देखें)। यद्यपि बाद के दो डाटाबेस एक आरंभिक अवस्था में हैं, और पहले के दो डाटाबेस व्यापक और सुदृढ़ हैं।

### चित्र 3. भारत सरकार द्वारा संगृहीत किया जाने वाला डाटा



## बॉक्स 2: नागरिकों से संबंधित समृद्ध आंकड़े जिनका सरकार अपने नागरिकों के कल्याण के लिए उपयोग कर सकती है

सरकार प्रशासनिक आंकड़ों (डाटा) का रखरखाव मुख्यतः सांख्यिकीय प्रयोजनों के लिए करती है। प्रशासनिक डाटा समूहों के अंतर्गत जन्म-मृत्यु संबंधी अभिलेख, अपराध-रिपोर्ट, भूमि एवं संपत्ति पंजीकरण, वाहन पंजीकरण, राष्ट्रीय सीमाओं पर लोगों की आवाजाही, कर-अभिलेखों, आदि से जुड़े आंकड़ों को सम्मिलित किया जाता है। सरकार कल्याणकारी स्कीमों के मूल्यांकन के लिए भी आंकड़े एकत्र करती है; उदाहरण के लिए, पेयजल एवं स्वच्छता मंत्रालय स्वच्छ भारत मिशन की प्रभावोत्पादकता का आंकलन करने के लिए शौचालयों के उपयोग से संबंधित आंकड़े एकत्र करता है।

दूसरी ओर, सर्वेक्षण संबंधी आंकड़े, वे आंकड़े होते हैं जिन्हें व्यवस्थित, आवधिक सर्वेक्षणों के माध्यम से विशेष तौर पर सांख्यिकीय प्रयोजनों के लिए एकत्र किया जाता है। उदाहरण के लिए, राष्ट्रीय प्रतिदर्श सर्वेक्षण कार्यालय रोजगार, शिक्षा, पोषण, साक्षरता, आदि पर देशभर में बड़े पैमाने पर प्रतिदर्श सर्वेक्षण करता है। चूंकि इन आंकड़ों को सांख्यिकीय विश्लेषण के लिए एकत्र किया जाता है अतः सहभागियों की पहचान असंगत होती है और इन्हें सूचित नहीं किया जाता, तथापि इन पहचानों को सुरक्षित तरीके से गोपनीय रखा जाना आवश्यक होता है।

संस्थागत आंकड़े लोगों के बारे में ऐसे आंकड़े होते हैं जो सार्वजनिक संस्थाओं द्वारा रखे जाते हैं। उदाहरण के लिए सरकार द्वारा संचालित जिला अस्पताल अपने मरीजों के चिकित्सा संबंधी समस्त अभिलेखों का रखरखाव करता है। सरकार द्वारा संचालित विद्यालय अपने छात्रों के संबंध में उनकी वैयक्तिक सूचनाओं का रखरखाव करता है। राज्य द्वारा संचालित विश्वविद्यालय छात्रों की शैक्षिक उपलब्धियों तथा उनको प्रदान की गई डिग्रियों से संबंधित अभिलेखों का रखरखाव करते हैं। ऐसे समस्त आंकड़े स्थानीय रूप से मुख्यतः दस्तावेजी रूप में, रखे जाते हैं, जहां ऐसे आंकड़ों को डिजिटलीकृत करने तथा क्षेत्रीय या राष्ट्रीय स्तर पर एकीकृत करने की गुंजाइश कम होती है।

लेन-देन संबंधी आंकड़े किसी व्यक्ति के लेनदेनों से संबंधित आंकड़े होते हैं, जैसेकि, एकीकृत भुगतान इंटरफेस (यूपीआई) पर या भीम आधार पे पर किए जाने भुगतान आंकड़ों की यह श्रेणी उदीयमान स्थिति में है, परंतु जैसे-जैसे लेन-देनों के कैशलैस भुगतान करने वाले लोगों की संख्या में वृद्धि होती जाएगी वैसे-वैसे ही इसका विकास होते जाने की संभावना है।

4.27 भारत में डाटा संग्रहण की व्यवस्था अत्यधिक विकेंद्रित स्थिति में है। सामाजिक कल्याण के प्रत्येक संकेतक के लिए आंकड़े एकत्र करने की जिम्मेदारी संबंधित संघीय मंत्रालय और इसके राज्य समकक्षों की होती है। फलतः एक मंत्रालय द्वारा एकत्र किए गए आंकड़ों को अन्य मंत्रालय द्वारा एकत्र किए गए आंकड़ों से पृथक रखा जाता है। उदाहरण के लिए, किसी व्यक्ति के वाहन के रजिस्ट्रेशन संबंधी आंकड़ों को एक मंत्रालय द्वारा रखा जाता है, जबकि उसी व्यक्ति के सांपत्तिक स्वामित्व संबंधी अभिलेखों का रखरखाव किसी अन्य मंत्रालय द्वारा किया जाता है। ये डाटा-समूह राज्य द्वारा संचालित विश्वविद्यालय द्वारा उसी व्यक्ति की शैक्षिक उपलब्धियों संबंधी अभिलेखों से भी तथा दशवार्षिक जनगणना सर्वेक्षणों में एकत्र की गई अन्य

जनांकिकीय सूचनाओं से भी परस्पर पृथक हो जाते हैं। चूंकि ये डाटा-सेट असंयोजित अवस्था में हैं (नीचे चित्र 4 देखें), अतः प्रत्येक मंत्रालय के पास व्यक्तियों/फर्मों के आंकड़ों का एक छोटा-सा पृथक अंश ही है। तथापि यदि इन छोटे-छोटे पृथक अंशों को हम एक साथ मिला सकें तो इस सूक्ति को सिद्ध होते पाएंगे कि कोई भी समग्र अपने समस्त अंशों के योग से वृहत्तर हो जाता है। वैश्विक सर्वोत्तम प्रक्रियाओं से अधिगम के रूप में, बॉक्स 3 और 4 ट्रांसपोर्ट फॉर लंदन डाटा इनिशिएटिव तथा संयुक्त राज्य सरकार के डाटा इनिशिएटिव संबंधी अध्ययनों पर प्रकाश डाला गया है। बॉक्स 5 में, तेलगांवा सरकार के क्लोजर होम: समग्र वेदिका पहल से इसी प्रकार के अधिगम की स्थिति को दर्शाया गया है।

## बॉक्स 3: ट्रांसपोर्ट फॉर लंदन डाटा को विमुक्त करने का लाभ

ट्रांसपोर्ट फॉर लंदन (टीएफएल) किसी भी व्यक्ति के उपयोगार्थ, एक मुफ्त प्रारूप में काफी मात्रा में आंकड़े निःशुल्क उपलब्ध कराता है—जैसेकि समय-सारणियां, सेवा स्थिति और व्यवधान सूचना, आदि। यात्रा संबंधी इस तरह के मुक्त आंकड़े, प्रौद्योगिकीय अर्थनीतिक विकास और सार्वजनिक परिवहन की उपयोग-वृद्धि में सहायता करते हुए, समय बचाने, अनिश्चितता को कम करने और सूचना-लागतों को नीचे लाने के लिए ट्रेवल एप्स और रीयल-टाइम अलर्ट्स को सपोर्ट कर सकते हैं। अद्यतन स्थिति के अनुसार, विशेष रूप से टीएफएल ओपन डाटा फीड्स के उपयोग से 600 से अधिक एप्स को समर्पित किया जा रहा था, जिनका 42 प्रतिशत से अधिक लंदन वासियों द्वारा उपयोग किया गया।



टीएफएल ने दर्शाया है कि आंकड़े जारी कर देने से सार्वजनिक परिवहन उपयोग करने वाली जनता के समय की बचत का मान 15 मिलियन पाउण्ड से 58 मिलियन पाउण्ड वार्षिक रहा है। डेलॉयट के एक अध्ययन में पाया गया है कि परिवहन सूचना ऐप्स तथा रियल टाइम अलर्ट से मिलने से ही 70-95 मिलियन पाउण्ड के समान समय की बचत हो जारी है। यह अनिश्चितता में कमी का मूल्य मान है, साथ ही टीटीएफएल द्वारा स्वतंत्र आंकड़े जारी करने से लंदन की तकनीकी अर्थव्यवस्था में 14 मिलियन पाउण्ड की अतिरिक्त सकल मूल्य वृद्धि हुई है और 700 नए रोजगार जुड़े हैं। इसने अनेक नए आय और बचत के अवसर सुलभ कराए हैं; टीएफएल में कार्य करने की विधियां संभव हुई हैं; यात्रियों को जानकारी सुलभ होने मात्र से स्थानीय बसों में 20 मिलियन पाउण्ड की अधिक टिकटें बिकी हैं। इस डेटा का प्रयोग अनेक प्रकार की ऐप्स द्वारा हो रहा है जिनमें प्रारंभ स्टार्ट-अप्स से लेकर वैश्विक व्यापार प्रौद्योगिकी प्लेटफार्म तक शामिल है-सभी ने इस डेटा के माध्यम से समय बचाया है और तनाव घटाया है। इस समय टीएफएल मुक्त डाटा के 13700 पंजीकृत प्रयोक्ता हैं।

लंदन इन्फ्रास्ट्रक्चर मैपिंग अनुप्रयोग/एप्लिकेशन एक ऐसा नवीन प्लेटफार्म है जो जनोपयोगी सेवाओं, नगरों (नगर जो संसद में अपना प्रतिनिधि भेजे), लंदन के मेयर तथा टी.एफ.एफ. को अवसरचर्चात्मक निवेश तथा आयोजना से प्रासंगिक जानकारी को साझा करने की अनुमति प्रदान करता है। डेटा की रेंज को एक साथ लाकर यह अनुप्रयोग (एप्लीकेशन), विभिन्न सेक्टरों के बीच बेहतर तालमेल, निर्माण और प्राधिकरण के प्रति सम्मिलित बेहतर दृष्टिकोण तथा भावी मांग एवं क्षमता संबंधी बाधाओं की पहचान को सुगम बनाता है। सूचना को स्थान की दृष्टि से एक ऐसी बिस्पोक मैपिंग एप्लीकेशन के द्वारा देखा जाता है जिसे प्रयोक्ताओं के परामर्श से विकसित किया गया है। पूर्व साक्ष्यों से ज्ञात हुआ है कि यह टूल संपूर्ण लंदन में, सड़क निर्माण कार्य जैसी परियोजनाओं को बेहतर समय पर अनुमति देकर तथा निर्माण कार्य के प्रति संयुक्त दृष्टिकोण के माध्यम से लागत कम करते हुए (जैसे:- उपयोगिता केबल और पाईप की संयुक्त डाकिंग) हाउसिंग/आवासन विकास के द्वारा खोलते हुए तथा दुर्घटनाओं को कम करके, निवेश संरक्षण को बेहतर ढंग से समर्थन प्रदान करता है।

वर्तमान में टी.एफ.एम. एक ऐसे नए टूल को विकसित कर रहा है जो अनुरक्षण कर्मचारियों को ट्रेन/रेलगाड़ी की स्थिति के बारे में लाईव सूचना/साक्षात् जानकारी उपलब्ध कराने के लिए ट्यूब ट्रेन्स/भूमिगत रेल से प्राप्त द्वारा फीड्स का विश्लेषण करेगा। इस टूल के उपयोग द्वारा कर्मचारी डाटा का विश्लेषण कर सकते हैं तथा यह पता लगा सकते हैं कि खराबी कहां पर है या कहां पर आ सकती है तथा सेवाबाधित होने से पहले ही उनका निवारण भी कर सकते हैं। इस टूल में अनुरक्षण योजना को और अधिक कुशल बनाने तथा सेवा संबंधी मंहगे दोषों को घटित होने से पहले ही रोकने की प्रबल संभावनाएं हैं। इन-हाउस क्षमता, तृतीय पक्ष व्यय को कम करते हुए टी.एफ.एल. द्वारा धनराशि की बचत करने में भी सहायता प्रदान करेंगी। वर्तमान में पिछले पांच वर्षों से टी.एफ.एल. द्वारा एक बाह्य अनुरक्षण समर्थन सविदा पर 46 मिलियन पाउण्ड खर्च किए जा रहे हैं।

स्रोत: डेटा एफ.ए. पब्लिक गुड पर नेशनल इन्फ्रास्ट्रक्चर रिपोर्ट <https://www.nic.org.uk/wp-content/uploads/Data-for-the-Public-Good-Nic-Report.pdf> पर उपलब्ध है।

#### बॉक्स 4: अमेरिकी सरकार द्वारा किए गए डाटा प्रयास/पहल:

29 जनवरी, 2009 को अमेरिका के राष्ट्रपति बराक एच. ओबामा ने उदार और पारदर्शी सरकार पर एक ज्ञापन जारी करते हुए अपने प्रशासन से “एक पारदर्शी जन-भागीदारी और सहयोग प्रणाली” स्थापित करने के लिए कहा। उदार-सरकार प्रयास/पहल, संघीय सरकार और सार्वजनिक एजेंसियों को मशीन द्वारा पठनीय प्रारूप में सार्वजनिक उपयोग के लिए उनका डाटा, ऑन लाईन प्रकाशित करने के लिए बाध्य करती है। अपने डाटा को लोकतांत्रिक बनाने के लिए एक बोली में, अमेरिका सरकार ने 138,000 से अधिक डाटा सेट्स को जनता के लिए सुलभ बना दिया। फरवरी -2015 में अमेरिका सरकार ने डी.जे. पाटिल को, अमेरिकी जनता के हितार्थ डाटा की शक्ति को उन्मुक्त करने के लिए विज्ञान एवं प्रौद्योगिकी गति कार्यालय में सर्वप्रथम प्रमुख डाटा वैज्ञानिक के रूप में नियुक्त किया। उन्होंने वर्ष 2017 तक दो वर्षों की सेवाएं प्रदान की।

**Data.gov**, जो कि अमेरिका सरकार का एक श्रेष्ठ प्लेटफार्म है, विभिन्न सार्वजनिक एजेंसियों से डाटा और मेरा-डाटा साझा करता है और इस प्रमुख मंच में निम्नलिखित मूल्य संवर्धित/वर्धित विशिष्टताएं हैं:-

- (i) स्थान, डाटा-सेट प्रकार, टैग, फॉर्मेट, समुदाय इत्यादि द्वारा फ़िल्टर करने की क्षमता।
- (ii) इसमें डाटा-सेट्स एकीकृत करने और विजुअलाइजेशन का एक टूल उपलब्ध है।
- (iii) डाटा-समूहों की गुणवत्ता में सुधार के लिए, विभिन्न मंचों, ब्लॉगों और समुदायों में भागीदारी और प्रतिक्रिया देने के लिए जनता के साथ वचनबद्धता।

- (iv) ऐसे संसाधन जो समस्त संघीय ऐजेंसी, ए.पी.आई. को लिंक उपलब्ध कराते हैं।
- (v) सॉफ्टवेयर डेवलपमेंट किट ओपन-सोर्स संसाधनों तथा कूटांक-संग्राहक (रिपॉजिटरी ऑफ कोड) सहित डैवलैपर हब।
- (vi) शैक्षिक एप्लीकेशन (कॉलेज की खोज, पब्लिक स्कूल ढूँढना, जिले), ऊर्जा एवं पर्यावरण एप्स (वैकल्पिक ईंधन स्थाननिर्धारक) तथा खाद्य एवं पोषण एप्लीकेशन (फूडयूकेर, गुड़गाईड) के माध्यम से आंकड़ा संचालित (डाटा-ड्रिवन) निर्णय लेना।
- (vii) अमेरिका में विभिन्न शहरों से डाटासैट्स प्रकाशित करने के लिए [cities.data.gov](http://cities.data.gov) प्लेटफॉर्म।
- (viii) आपदा-प्रबंधन, स्वास्थ्य-सेवा, कानून-व्यवस्था पुलिसिंग और जलवायु परिवर्तन में सहयोगात्मक प्रयासों के लिए [disaster.data.gov](http://disaster.data.gov) वैबपोर्टल का लोकार्पण किया गया।
- (ix) स्वास्थ्य देखभाल के क्षेत्र में सटीक औषध पहल (पीएमआई) एक ऐसे मुख्य खुला डाटा परियोजना है जिसका उद्देश्य हर व्यक्ति का उपचार करना, आनुवंशिक आधारित रोग-विषयक कैंसर जांच केंद्रों को बढ़ाना और राष्ट्रीय “कैंसर जागरूकता नेटवर्क” की स्थापना करना है ताकि उपचार निर्णयों को मदद मिल सके।

यूनाइटेड स्टेट्स कृषि विभाग (यूएसडीए), ऐसी विभिन्न कृषि संबंधी उपयोगी वस्तुओं, जिनका व्यापक रूप से उपयोग किया जाता है, की कीमतों मूल्यों के संबंध में नियमित पूर्वानुमान जारी करता है इन अनुमानों में भावी कृषि संबंधी बाजारों को प्रभावित करने वाली मुख्य ताकतों एवं अनिश्चितताओं, दीर्घकालिक वैश्विक आर्थिक विकास की संभावनाओं, कृषि संबंधी उत्पादन, खपत एवं व्यापार और यूएस मुख्य कृषि पण्यों एवं भावी कीमत उतार-चढ़ाव की पहचान की जाती है। इन अनुमानों का प्रयोग वैकल्पिक नीति परिदृश्यों के प्रभावों का विश्लेषण करने के लिए भी किया जा सकता है। डाटा अनुमानों की यह सुगम उपलब्धता से वस्तुतः विश्व भर में कृषि संबंधी बाजार संचालित होते हैं।

### बॉक्स 5: सरकारों में शिक्षण में संघवाद: तेलंगाना सरकार की समग्र वेदिका पहल

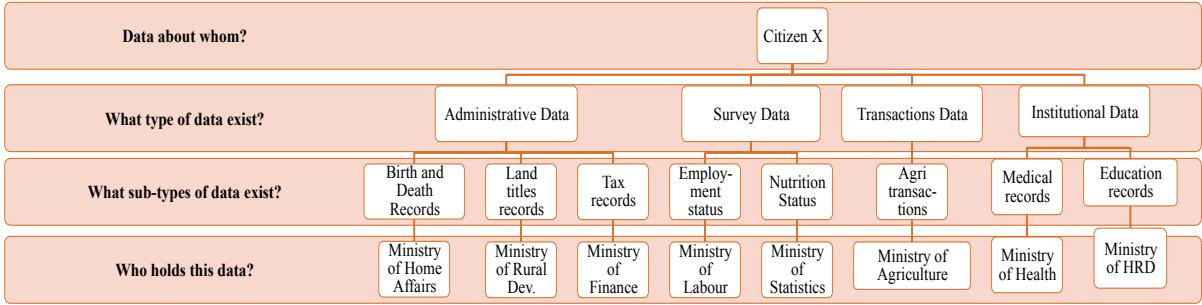
तेलंगाना सरकार की ‘समग्र वेदिका’ पहल से डाटा सेट एकीकरण कार्य के महत्वपूर्ण लाभ की महक आती है। इस पहल का संबंध सामान्य आइडेंटीफायर किसी व्यक्ति का नाम एवं पता का उपयोग करने वाले लगभग 25 मौजूदा सरकार के डाटा सेट से है। प्रत्येक व्यक्ति के बारे में सात श्रेणियों की सूचना को इस समूहन कार्य से जोड़ा गया था जो— अपराध, परिसंपत्तियां, उपयोगिताएं, सब्सिडी, शिक्षा, कर एवं पहचान थे। इसके अलावा प्रत्येक व्यक्ति को पति-पत्नी, भाई-बहन, माता-पिता और अन्य सगे संबंधियों से जोड़ा गया था। इस पहल में सुरक्षा के सभी उपायों पर ध्यान दिया गया है ताकि डाटा से किसी प्रकार की छेड़छाड़ या गोपनीयता के अतिक्रमण को रोका जा सके। डाटा बेस में डाटा जोड़ने और डाटा संपादन का अधिकार मंत्रालय अथवा विभाग वार अलग अलग होता है। निर्दिष्ट विभाग केवल विशेष क्षेत्र के लिए डाटा तैयार कर सकता है— उदाहरण के लिए मोटर वाहन विभाग सभी उपलब्ध डाटा को देख सकने के बावजूद शिक्षा से संबंधित डाटा में फेरबदल नहीं कर सकता।

4.28 हाल ही में डाटा सेटों को जोड़ने—मुख्यतः आधार नंबर को पैन डाटा बेस, बैंक खातों, मोबाइल नंबर आदि जैसे डाटाबेस से जोड़ने के पर कुछ चर्चाएं की गई हैं। यहा स्पष्टीकरण आवश्यक है। जब कोई व्यक्ति आधार संख्या को बैंक खातों की डाटाबेस जैसे मौजूदा डाटाबेस से जोड़ता है, तो एक अतिरिक्त कॉलम होता है जो सारणी में जोड़ा जाता है। इस प्रकार से जोड़ने का यह कार्य ‘एक तरफा’ है। दूसरे रिकार्ड को जोड़ने और बेनामी खातों की छंटनी करने के लिए बैंक अब

प्रमाणित आधार संख्या (अर्थात प्रतिव्यक्ति 64-डिजिट नंबर (जो व्यक्ति के 12-डिजिट आधार संख्या पर आधारित है परन्तु उसके समतुल्य नहीं है) का उपयोग करते हैं परन्तु इसका मतलब यह नहीं है कि यूआईडीएआई अथवा सरकार अब व्यक्ति से संबंधित बैंक खाते अथवा अन्य डाटा की जानकारी प्राप्त कर सकती है।

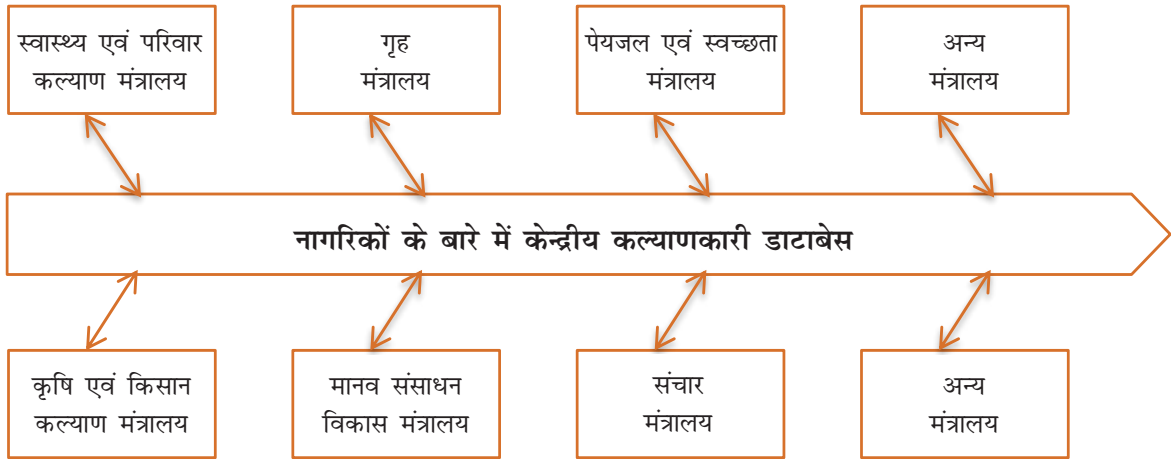
4.29 चित्र 4 में वर्तमान प्रणाली को स्पष्ट किया गया है। एक काल्पनिक नागरिक ‘एक्स’ से संबंधित डाटा

### चित्र 4: किसी काल्पनिक नागरिक 'एक्स' के बारे में आंकड़े ( डाटा )



टिप्पणी : संक्षिप्तीकरण किए गए कुछ मंत्रालयों के नाम, डाटा किसमें और उपकिसमें दृष्टांत उदाहरण मात्र है और सर्वांगपूर्ण नहीं हैं।

### चित्र 5: असंयोजित डाटा सेटों को जोड़ने के लिए उद्यम बस संरचना



(सूचना) को सभी मंत्रालयों में संसूचित कर दिया गया है। वास्तव में इस उदाहरण को बहुत सरल कर दिया गया है। उदाहरणार्थ 'ग' से संबंधित "चिकित्सा रिकॉर्ड" सरकारी जिला अस्पताल में रखे गए हैं, जहां उसका हाल ही में उपचार किया गया है। समस्त संभावनाएं ऐसी हैं कि यह "चिकित्सा रिकॉर्ड" कागजों तक ही सीमित होकर रह जाएंगे और राज्य स्तर तक संकलित होकर विश्लेषण के लिए कभी नहीं पहुंच पाएंगे, इसकी केन्द्र तक तो पहुंचने की कल्पना ही नहीं की जा सकती है। तथापि, इस उदाहरण में चित्रित किया गया है कि प्रत्येक प्रकार के डाटा के लिए केन्द्रीय मंत्रालय अंततोगत्वा जवाबदेह है।

4.30 चित्र 5 में "डाटा एक्सेस फ़िड्यूशरी आर्किटेक्चर (डाटा पहुंच प्रत्ययी संरचना)" को सचित्र स्पष्ट किया गया है। सरकार के प्रत्येक विभाग का यह दायित्व है

कि वह उसके पास उपलब्ध होने वाले डाटा को आगे उपलब्ध करवाए जोकि वह 'डाटा प्रबंधक' के रूप में धारण किए रखता है। इन विभागों को 'निजी डाटा' तथा 'लोक डाटा' द्वारा अपेक्षित मानकों के अनुसार संसाधित करते समय सावधानी रखनी होगी। 'डाटा निवेदक' को तत्पश्चात यह डाटा, "डाटा एक्सेस फ़िड्यूशरी" के माध्यम से उपलब्ध करवा दिया जाएगा। डाटा निवेदक सार्वजनिक अथवा निजी संस्थान दोनों में से कोई भी हो सकता है किंतु डाटा तक तभी पहुंच हो पाएगी जब उसके पास समुचित 'प्रयोक्ता स्वीकृति' होगी। एंड-टू-एंड एनक्रिप्शन होने के कारण स्वयं 'डाटा एक्सेस फ़िड्यूशरी', डाटा को नहीं देख सकेगा। डाटा को "सार्वजनिक उपयोगी वस्तु" बनाने वाली सरकार की इस पहल से ऐसे 'मॉडल' द्वारा "प्रयोक्ता-स्वीकृति" अत्यंत महत्वपूर्ण बन जाएगी।

4.31 एक नागरिक को ही केवल विश्लेषण की संभावित इकाई नहीं समझना चाहिए। कोई एक व्यक्ति स्कूल या गांवों या अस्पतालों के विश्लेषण की मांग कर सकता है। उदाहरणस्वरूप यदि प्रत्येक ग्राम की एक “अनन्य-पहचान (यूनीक आइडेंटिफायर) विद्यमान हो तो गांवों के विभिन्न डाटाबेस को मिलाया जा सकता है। दुर्भाग्यवश अधिकांश मामलों में मंत्रालयों और विभागों के अपने-अपने कूट किसी एक प्रदत्त-भौगोलिक क्षेत्र के लिए होते हैं। उदाहरणार्थ प्रत्येक ग्राम का एक विशेष पिन कोड होता है जोकि डाक विभाग द्वारा निर्धारित किया गया होता है, इसी प्रकार ग्रामीण विकास विभाग द्वारा एक “विलेज कोड (ग्राम कूट)” प्रत्येक ग्राम के लिए किया गया है और, “स्वास्थ्य और परिवार कल्याण मंत्रालय” द्वारा “हेल्थ ब्लॉकों” का निर्धारण किया गया है। “सर्वनिष्ठ पहचान (कॉमन आइडेंटिफायर) की कमी के कारण सूचना का संकलन करने में कठिनाई होती है।

4.32 इस समस्या का हल करने की जो पहल है, वह है, स्थानीय सरकारी निर्देशिका जो पंचायती राज मंत्रालय द्वारा विकसित किया गया एक कंप्यूटर-प्रोग्राम (एप्लिकेशन) है। इस स्थानीय सरकारी निर्देशिका में सभी स्थानीय एकांकों स्थानीय सरकारी निकायों के भू-क्षेत्र के प्लेटफार्म मानचित्र (जैसे कि ग्राम और उनकी संबंधित ग्राम पंचायतें) को सम्मिलित करते हुए वर्ष 2011 की जनगणना के अनुसार लोकेशन कोड निर्धारित किए गए हैं। “स्थानीय सरकारी निर्देशिका” डाटा को एक “सार्वजनिक वस्तु” समझने का सबसे उत्तम उदाहरण है। पंचायती राज मंत्रालय ने जटिल समस्या का निवारण करने की दिशा में महत्वपूर्ण प्रयास किया जो ग्रामीण भारत में सेवाएं देने वाली प्रत्येक सरकारी और निजी संस्थानों दोनों के लिए उभयनिष्ठ थी जैसे औपचारिक पते का नहीं होना, -इसके लिए प्रत्येक स्थान का एक कूट निर्धारित किया गया है। मंत्रालय ने केवल डाटा ही नहीं बनाएं अपितु सार्वजनिक उपयोग के लिए उसको प्रकाशित भी किया गया है। सभी ‘सरकारी डाटा बेसों’ जिनमें ‘स्थानीय-कूटों (लोकेशन कोड)’ की आवश्यकता होती है, उनको यदि इस निर्देशिका के कूटों (कोडों) से मिला दें तब सभी डाटा बेसों को कॉमन लोकेशन फील्ड मिल सकेगी और जिससे “डाटा मर्जिंग (डाटा संविलय)” में सहायता प्राप्त होगी तथा कल्याण कार्यों के लाभ वितरण के दौरान संशुद्धता में होने वाली त्रुटियों को घटाया जा सकेगा।

### सभी प्रकार के दुराव को समाप्त करने के लिए विकसित प्रौद्योगिकी का उपयोग

4.33 एकीकृत प्रणाली की दक्षता तीन महत्वपूर्ण गुणों पर

निर्भर करती है। पहला, जब कोई मंत्रालय समग्र डाटाबेस को देखने में सक्षम होता है तब वह मंत्रालय केवल उसी डाटा फील्ड में परिवर्तन कर सकता है, जिसके लिए वह जिम्मेवार होता है। दूसरा, डाटा को तत्काल रूप से (रीयल-टाइम) अद्यतन किया जाना चाहिए और यह इस रीति से करना चाहिए कि जब एक मंत्रालय डाटाबेस में परिवर्तन कर रहा हो तब दूसरे मंत्रालय की डाटा तक पहुंच में किसी तरह की बाधा नहीं हो। तीसरा और सबसे महत्वपूर्ण डाटा बेस अत्यंत सुरक्षित रहना चाहिए और इसमें किसी भी प्रकार से हेर-फेर की संभावना नहीं होनी चाहिए।

4.34 प्रत्येक नागरिक से संबंधित ऐसी व्यापक, सर्वांगपूर्ण सूचना द्वारा सरकार को सशक्त करना प्रथमदृष्टि में भयभीत करने वाला हो सकता है। तथापि, यह सत्य से बहुत दूर है। पहला, अभी तक सरकारी रिकॉर्डों में विशाल मात्रा में डाटा विद्यमान हो चुका है और इस योजना का प्रयोजन केवल इस उपलब्ध डाटा को अधिक प्रभावी विधि द्वारा उपयोग करना है। इस प्रस्ताव में विलक्षणता यह है कि यहां किसी प्रकार की कोई नई सूचना एकत्र नहीं की जाएगी अपितु जो डाटा सरकार के पास उपलब्ध है उसी को ‘प्रयोक्ता स्वीकृति’, ‘समुचित निजता’ तथा ‘न्याय संगति संबंधी नियंत्रण’ को ध्यान में रखते हुए नागरिकों, सरकार, निजी व सार्वजनिक संस्थानों को उपलब्ध करवाया जा सकेगा।

4.35 दूसरा, जहां लोग डाटा को प्रकट नहीं करना चाहते हैं तो वे “बाहर आने के विकल्प” का चयन सदैव करते हैं। उदाहरणार्थ कोई भी व्यक्ति सर्वेक्षण में भाग नहीं लेने का विकल्प का चयन कर सकता है अथवा सरकार द्वारा प्रोत्साहित की जा रही भुगतान संबंधी सेवाओं का उपयोग नहीं करने के लिए स्वतंत्र है, इसके अपवाद भी हैं। लोग बिना लाइसेंस और बिना रजिस्ट्रेशन प्रमाण पत्र के वाहन नहीं खरीद सकते हैं। किंतु इस डाटा को यदि अनिवार्य नहीं बनाया गया तो संपत्ति अधिकार के कार्यान्वयन, सड़क सुरक्षा और राष्ट्रीय सुरक्षा को खतरा हो सकता है, जिनपर समझौता नहीं किया जा सकता है। किंतु डाटा के शेष वर्ग जैसे कि संस्थागत, सर्वेक्षण तथा लेन-देन संबंधी डाटा के संबंध में लोगों के पास विकल्प रहेगा।

4.36 तीसरे कुछ सार्वजनिक सेवाओं का कोई व्यवहार्य प्राइवेट बाजार का विकल्प न होने के बावजूद ऐसी सेवाओं से प्राप्त प्रत्येक व्यक्ति से संबद्ध डाटा को शेयर करने का विकल्प डाटा सुलभ न्यासी संरचना (डाटा एसेस फाइंडुसियरी आर्किलेक्चर) के अधीन सदैव नागरिक के

पास ही उपलब्ध होगा। इसके अलावा आपके सभी प्रकार के डाटा के लिए अपरिवर्तनीय लॉग सुविधा उपलब्ध होगी जिससे कि नागरिकों को यह जानकारी हो सके कि उनके डाटा को किसने और क्यों देखा है।

4.37 सिद्धांत यह है कि अधिकतर डाटा व्यक्तियों द्वारा और व्यक्तियों के संबंध में तैयार किए जाते हैं और इसलिए उनका प्रयोग व्यक्तियों के लिए ही किया जाना चाहिए। डाटा सेटों में सूचना साझा करने से सामाजिक कल्याण के कार्य करने में सुधार होगा, व्यक्ति बेहतर निर्णय लेने में सशक्त होंगे और एक महत्वपूर्ण लोक हित का लोकतंत्रीकरण होगा।

### भारत के डाटा अवसंरचना का रूपांतरण

4.38 निसंदेह, ऊपर यथा उल्लिखित सिस्टम के लिए एक वृहत डाटा अवसंरचना अपेक्षित होती है। डाटा सेट को संयोजित करने से अपने आप ही इसके लाभ प्राप्त होंगे, और यदि डाटा अनिश्चित गुणवत्ता का है तो इसका लाभ भी सीमित होगा और इसका सुगम प्रक्रमण करना आदि भी सरल नहीं होगा। डाटा को प्रयोग में लाने के संबंध में चार कार्रवाई की जाती है यथा डाटा को एकत्रित करना, डाटा का भंडारण करना, डाटा का प्रक्रमण करना और डाटा का प्रसारण करना। इनमें से

प्रत्येक के संबंध में भारत में सुधार की गुंजाइश होती है (चित्र 6 देखें)।

### डाटा एकीकरण : कागज से डिजिटल के बजाए प्रत्यक्ष डिजिटल

4.39 जब तक डाटा डिजिटल, मशीन पठनीय प्रारूप में नहीं होगा, इसकी उपयोगिता सीमित रहेगी। सभी हितधारकों द्वारा डाटा के कागजी रिकॉर्ड तक नहीं पहुंचा जा सकता है। भले ही योजनाओं के कागजी रिकॉर्ड को स्कैन करके डिजिटल निधान (संग्रह) में स्टोर किया गया हो, लेकिन वे विश्लेषण के लिए सुलभ नहीं होते हैं। डाटा का जो मूल प्रयोजन है, उसे पूरा करने के लिए डाटा को एक डिजिटल, मशीन पठनीय रूप में प्रसारित करना होगा जिसका डाउनलोड और विश्लेषण किया जा सकता है। हालांकि अधिकांश राज्य और राष्ट्रीय स्तर के डाटा पहले से ही इस रूप में हैं, लेकिन डिजिटल रिकॉर्ड अभी तक अधिक व्यापक प्रशासनिक स्तरों पर सर्वव्यापी नहीं है। उदाहरण के लिए एक नागरिक अपने गांव की स्वच्छता स्थिति को डिजिटल डेटाबेस पर नहीं देख सकता है। ज्यादा से ज्यादा, वह अपने जिले को देख सकता है। कई योजनाओं के लिए डाटा केवल राज्य स्तर पर उपलब्ध है।

चित्र 6 डाटा संग्रहण, संचयन, प्रसंस्करण और प्रसार संबंधी रूपांतरण

<b>डाटा संग्रहण</b>	<ul style="list-style-type: none"> <li>● मौजूदा कागज आधारित डाटा को डिजिटल बनाना</li> <li>● स्रोत पर डिजिटल डाटा एकत्रण करना</li> </ul>
<b>डाटा संचयन</b>	<ul style="list-style-type: none"> <li>● चयनित डाटा का रियल-टाइम संचयन करना</li> <li>● एकत्रण और डाटा प्रविष्टि के बीच समय अंतराल को कम करना</li> </ul>
<b>डाटा प्रसंस्करण</b>	<ul style="list-style-type: none"> <li>● डाटा विश्लेषण के लिए सरकारी निकायों की क्षमता बढ़ाना</li> <li>● विश्लेषण और अंतरदृष्टि निर्माण में निजी क्षेत्र को भागीदार बनाना</li> </ul>
<b>डाटा प्रसार</b>	<ul style="list-style-type: none"> <li>● स्कीम डैशबोर्ड निर्मित करना</li> <li>● जनता के लिए जिला स्तरीय डैशबोर्ड खोलना</li> <li>● तृतीय पक्षकार अध्ययनों से प्राप्त डाटा को जनता के लिए खोलना</li> </ul>



4.40 हाल ही में शुरू की गई डिजिटल इंडिया पहल, कागज-आधारित डाटा को डिजिटल रूप में परिवर्तित करने के थकाऊ कार्य का एक सरल समाधान है। इस पहल के माध्यम से स्वैच्छिक नागरिकों को स्कैन किए गए दस्तावेजों का स्निपेट देकर, जिसका टंकण, वे डाटा प्रविष्टि पोर्टल में करते हैं, डाटा प्रविष्टि प्रयास को क्राउड-सोर्स किया जा सकता है। योगदानकर्ता सही प्रविष्टियों के लिए नकद पुरस्कार अर्जित करते हैं (विवरण के लिए बॉक्स 6 देखें)।

4.41 हालांकि, यह योजना मौजूदा कागजी रिकॉर्ड को डिजिटल करने के लिए एक लंबा रास्ता तय करती है, फिर भी डाटा संग्रह की वर्तमान प्रक्रिया को डिजिटल में बदलने के लिए, डाटा को पहले कागज पर संग्रह करने और बाद में एक डिजिटल प्रारूप में परिवर्तित करने की प्रक्रिया के विपरीत एक समानांतर पहल की आवश्यकता है। निःसंदेह, प्रत्येक अस्पताल, ग्राम पंचायत, विद्यालय या ब्लॉक कार्यालय को इलेक्ट्रॉनिक डाटा-एंट्री डिवाइस की आपूर्ति करने की कवायद काफी महत्वाकांक्षी है, लेकिन इस लक्ष्य को प्राप्त किया जा सकता है। स्रोत पर डिजिटल डाटा संग्रह यह सुनिश्चित करता है कि डाटा की प्रविष्टि बिल्कुल उसी रूप में की गई है, जिस रूप में इसे देखा गया, डाटा प्रविष्टि के अतिरिक्त कार्य को कम करता है और डाटा प्रविष्टि त्रुटियों को समाप्त करता है। वास्तव में डिजिटल पारिस्थितिकी (इकोसिस्टम) तभी प्राप्त हो सकती है जब पारिस्थितिक (इकोसिस्टम) को बनाने वाली सूचना की सभी परतें डिजिटल हों।

## डाटा स्टोर करना

4.42 डाटा के रियल टाइम स्टोरेज से सार्वजनिक सेवा

प्रदाय तंत्र को लाभ प्राप्त हो सकता है। सेंट पीटर्सबर्ग, फ्लोरिडा शहर ने हाल ही में सेंट पीट स्टेट नामक एक नागरिक डैशबोर्ड बनाया गया है, जहां नागरिक शहर के विभागों के प्रदर्शन, विकास परियोजनाओं की स्थिति और विभिन्न इंटरैक्टिव उपकरण, जिसमें शहर में सभी पुलिस कॉल के इंटरैक्टिव मानचित्र शामिल है, के विषय में गतिशील रूप से अद्यतन की गई जानकारीयां देख सकते हैं।

4.43 सभी प्रकार के डाटा रियल-टाइम स्टोरेज के लिए अनुकूल नहीं होते। इसके अलावा, एक रियल-टाइम डाटाबेस की वृद्धिशील लागत हमेशा इसके आंशिक लाभ को सही नहीं ठहरा सकती है। हालांकि, जिन क्षेत्रों में रियल-टाइम नहीं है, वहां कम से कम सर्वर पर डाटा लॉग इन की हाई फ्रिक्वेंसी उपलब्ध होनी चाहिए, जिससे पर्याप्त लाभ प्राप्त हो सकता है।

4.44 शिक्षा ऐसा ही एक उदाहरण है। यहां तक कि छः माह की एक लघु अवधि थी एक बच्चे के जीवन में काफी लंबा समय होता है—यह एक शैक्षणिक वर्ष था—आधी अवधि बनती है। मान लें कि एक विद्यालय में लड़कियों के काम में आने लायक शौचालय नहीं है जिसकी वजह से कन्याओं/ लड़कियों को विद्यालय से अनुपस्थित रहना पड़ता है। इस समस्या का निवारण यथासंभव शीघ्र किया जाना चाहिए; अन्यथा ये लड़कियां अपने मूल्यवान विद्यार्थी जीवन को खो देंगी। ब्लॉक या जिला शिक्षा अधिकारी को शौचालयों की स्थिति पर दैनिक या कम से कम साप्ताहिक रिपोर्ट प्राप्त करनी चाहिए ताकि यदि कोई समस्या एक या दो सप्ताह की अवधि के भीतर निपटाई नहीं जाती है तो वे उस स्थिति

## बॉक्स 6: डिजिटल इंडिया कैसे काम करता है

सरकारी विभाग डिजिटल इंडिया प्लेटफॉर्म पर कागजी रिकॉर्ड की स्कैन की गई प्रतियां अपलोड करते हैं। इन स्कैन किए गए दस्तावेजों को सार्थक आंकड़ों के साथ स्निपेट में विभाजित कर दिया जाता है। ये स्निपेट यादृच्छिक रूप से डिजिटल योगदानकर्ता को दिया जाता है। डिजिटल योगदानकर्ता ऐसे नागरिक हैं जो पोर्टल पर अपना समय स्वच्छा से देते हैं। स्निपेट्स प्राप्त करने पर, योगदानकर्ता सूचनाओं को पढ़ता है और इसे डाटा प्रविष्टि पोर्टल पर टाइप करता है। सभी परिवर्तित डाटा को संबंधित स्निपेट के आधार पर सत्यापित किया जाता है। योगदानकर्ता सटिक प्रविष्टियों पर रिवाइड प्वाइंट अर्जित करता है, जिसे या तो नकद के लिए रिडीम किया जा सकता है या डिजिटल इंडिया पहल के लिए दान किया जा सकता है। एक बार किसी विशेष दस्तावेज से संबंधित सभी स्निपेट को डिजिटल डाटा में बदल दिए जाने पर प्लेटफॉर्म दस्तावेज को डिजिटल रूप में पुनः संयोजित करता है और इसे वापस सरकारी विभाग को सौंप देता है। कोई भी भारतीय नागरिक आधार संख्या के साथ एक डिजिटल योगदानकर्ता के रूप में सहभागिता कर सकता है। नागरिकों को रिकॉर्ड पॉइंट्स और ‘डिजिटल योगदानकर्ता’ की मान्यता देने के माध्यम से प्रोत्साहित किया जाता है तथा वे ‘डाटा एंट्री ऑपरेटर’ के रूप में प्रमाण पत्र भी अर्जित कर सकते हैं। स्वाभाविक रूप से, कार्यक्रम में एक मोबाइल ऐप भी है, ताकि जिन नागरिकों के पास लैपटॉप या डेस्कटॉप कंप्यूटर नहीं हैं, वे भी सहभागिता कर सकें।

अपेक्षित कार्रवाई कर सकें। यदि किसी समस्या को ब्लॉक अधिकारी या जिला अधिकारी के संज्ञान में आने में महीनों लग जाते हैं तो विद्यालय की छात्राओं को अधिगम के कई माह गंवाने पड़ेंगे। वास्तविक काल में या कम से कम साप्ताहिक डिजिटल डैशबोर्ड अद्यतनीकरण, इस प्रकार के कठोर परिणामों को टाल देगा। इसी प्रकार कृषि पर भी विचार करें जहां पर कृषकों को वास्तविक समय पर मौसम की स्थितियों, अपेक्षित वर्षा, आदान तथा उत्पाद कीमतों की, दैनिक निर्णय लेने के लिए आवश्यकता पड़ती है। इस प्रकार के डाटा तक वास्तविक समय में पहुंच की सुविधा कृषि उत्पादकता तथा कृषकों की आय में सुधार लाने के लिए आवश्यक है।

4.45 सार्वजनिक सेवा प्रणाली में आईसीटी दृष्टिकोणों के व्यापक रूप से अपनाए जाने से अब वास्तविक काल डाटा संग्रहण तथा भंडारण अब कोई महत्वाकांक्षी तथा दीर्घकालीन स्वप्न नहीं रह गया है वरन् कम से कम चुनिंदा क्षेत्रों तथा संदर्भों में कार्यान्वित/साकार होने योग्य है।

#### डाटा प्रसंस्करण:-

4.46 जब डाटा को डिजिटल रूप में एकत्रित किया जाएगा, वास्तविक समय पर संगृहीत किया जाएगा तथा मौजूदा डाटा के साथ समेकित कर दिया जाएगा तो डाटा की भरमार हो जाएगी। इस प्रचुर डाटा में शासन को बदलने की जबरदस्त क्षमता है परंतु इस क्षमता को हासिल करने के लिए कौशल की आवश्यकता है। उदाहरण के लिए एक जिला शासकीय कार्मिक में उस जिले में डाटा उपयोग करने के विश्लेषणात्मक कौशल होने चाहिए। इस प्रकार के कौशल के अभाव में डाटा अवसंरचना में निवेश का केवल सीमित उपयोग ही किया जा सकता है।

4.47 सभी प्रशासनिक स्तरों पर सरकारों को वास्तविक समय में डाटा उपयोग करने, विश्लेषण करने तथा डाटा को एक अर्थपूर्ण सूचना में परिवर्तित करने के लिए अपनी आंतरिक क्षमताओं के निर्माण में निवेश करना चाहिए। हालांकि प्रत्येक सरकारी विभाग में एक समर्पित एनालिटिक्स या डाटा इन्साइट प्रभाग हो सकता है परंतु सांख्यिकी और कार्यक्रम कार्यान्वयन मंत्रालय तथा इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी मंत्रालय, राष्ट्रीय स्तर पर ऐसे प्रयासों के संचालन के लिए मॉडल विभागों के तौर पर कार्य कर सकते हैं।

4.48 सरकार सभी आवश्यक सुरक्षा संरक्षोपाय करते हुए कुछ प्रकार का डाटा, प्राइवेट लेजर्स के लिए खोले जाने पर भी विचार कर सकती है। खुला सरकारी डाटा पोर्टल पर जनता के साथ साझा किया गया डाटा। पूर्णतः गुप्त होता है तथा कई व्यक्तियों के पास एकत्रित किया जाता है। इस प्रकार का डाटा वहां पर सावधानीपूर्वक सांख्यिकी विश्लेषण का प्रचारण करता है जहां पर व्यक्तिगत स्तर की टिप्पणी/प्रेक्षण की आवश्यकता होती है। हालांकि व्यक्तिगत रूप से पहचान योग्य समस्त जानकारी अस्पष्ट करने के पश्चात् भी यदि डाटा निधि क्षेत्र के साथ साझा किया जाता है तो सरकार डाटा से सर्वाधिक संभव परिज्ञान प्राप्त करने के लिए डाटा एनालिटिक्स वृत्तिकों के कौशल और उत्साह का दोहन कर सकती है।

#### डाटा प्रसार:

4.49 ओपन सरकारी डाटा पोर्टल, जनता तक डाटा प्रसार करने का एक प्रभावी माध्यम है परंतु अधिकांश नागरिकों के पास डेटाबेस को विच्छेदित करने के लिए विश्लेषण आत्मक उपकरणों का उपयोग कर पाने का कौशल या समय नहीं है, सरल दृष्टिगत उपकरण महत्वपूर्ण हैं। इस पोर्टल ने कई दृष्टिगत उपकरणों को पहले से ही सक्रिय कर दिया तथा इनमें निम्नलिखित द्वारा संवर्धन किया जा सकता है।

4.50 सर्वप्रथम, सरकार कणिकामय डाटा वाली ऐसी प्रत्येक बड़ी सरकारी योजना के लिए योजना डैशबोर्ड की शुरुआत कर सकती है जो सभी तरह से ग्रामीण स्तर तक जाती है। डैशबोर्ड संबंधित सरकारी निकाय से डाटा प्राप्त करने के लिए तत्पर प्राप्तकर्ता तथा उसी डाटा को जनता के समक्ष वास्तविक समय में तत्परता से प्रदर्शक/प्रस्तुत करने वाला होना चाहिए। स्वच्छ भारत अभियान इस दिशा में एक अनुकरणीय नियंत्रण पर (डैशबोर्ड) है जिसे अन्य योजनाओं में भी प्रतिकृति के रूप में उपयोग किया जा सकता है ताकि नागरिक कल्याण योजनाओं के प्रत्यक्ष और वित्तीय प्रदर्शन के बारे में जान सकें। वर्तमान में इस प्रकार के डैशबोर्ड केवल कुछ ही योजनाओं के लिए मौजूद हैं। उदाहरणार्थ:-आई.सी.डी.एस. आंगनवाड़ी सेवा योजना जो कि भारत में महिला एवं बाल विकास की सबसे बड़ी योजना है, इसमें भी सार्वजनिक रूप से उपलब्ध डैशबोर्ड नहीं है।

4.51 मौजूदा डैशबोर्ड में, विशेष रूप से प्रशासन के अधिक व्यापक स्तर पर प्रगति दिखाने के लिए, मैट्रिक्स को संवर्धित किया जा सकता है। उदाहरण के लिए,

स्वच्छ भारत डैशबोर्ड में जिला स्तर पर तैनात किए जाने वाले स्वच्छता दूतों (स्वच्छता सहायकों) की संख्या या उस जिले में सूचना, शिक्षा और संचार (आईईसी) अभियानों की संख्या की जानकारी को शामिल किया जा सकता है। सभी संभावित क्षेत्रों में, इन मैट्रिक्सों पर डाटा सरकार के पास उपलब्ध है, लेकिन यह जनता के लिए उपलब्ध नहीं है, क्योंकि शायद यह डाटा डिजिटल रूप में नहीं है।

4.52 दूसरा, कई राज्य सरकारों ने विभिन्न कार्यक्रमों के लिए प्रबंधन सूचना प्रणाली (एमआईएस) के आधार पर जिला-स्तरीय डैशबोर्ड की स्थापना की है। इन राज्यों में आंध्र प्रदेश, पश्चिम बंगाल, मध्य प्रदेश और छत्तीसगढ़ शामिल हैं। आंध्र प्रदेश और पश्चिम बंगाल को छोड़कर अन्य राज्यों में ये डैशबोर्ड नागरिकों के लिए आसानी से उपलब्ध नहीं हैं। इसके लिए एक पासवर्ड की आवश्यकता होती है, जो केवल स्थानीय प्रशासन के पास उपलब्ध रहता है। कुछ डैशबोर्ड अब प्रचालन में नहीं हैं (टूटी लिंक), अन्य पुराने डाटा प्रदर्शित करते हैं। जबकि इस तरह के डैशबोर्ड को शुरू करने का निवेश सराहनीय है, यदि राज्य सरकारें सूचना प्रसारित करने के लिए लगातार इन डैशबोर्ड को बनाए रखती हैं तो लाभ अधिक होगा।

4.53 अंत में, कई केन्द्रीय मंत्रालयों और राज्य विभागों ने योजनाओं की आवश्यकता का मूल्यांकन अथवा प्रभाव मूल्यांकन के लिए डाटा संग्रहण करने का प्रयास किया है। हालांकि, इनमें से अधिकांश अध्ययनों को सार्वजनिक किया गया है, लेकिन अंतर्निहित आंकड़ों को सार्वजनिक नहीं किया गया है। चूंकि इस तरह के अध्ययन सरकारी निकायों के साथ साझेदारी में किए जाते हैं, इसलिए उन्हें जनता को उपलब्ध कराया जाना चाहिए ताकि इन अध्ययनों के निष्कर्षों को मान्य करने के लिए स्वतंत्र विश्लेषण किया जा सके।

## अनुप्रयोग

4.54 एक बार अवसंरचना सही हो जाने के बाद अनगिनत एप्लिकेशन चालू किए जा सकते हैं। एक सुदृढ़ डाटा बैकबोन समाज के प्रत्येक हितधारकों, केन्द्र सरकार से लेकर एक स्थानीय सरकारी निकाय को, नागरिकों से लेकर निजी संस्थानों को सशक्त कर सकता है। कुछ महत्वपूर्ण लाभों का वर्णन इस प्रकार है। (असर्वांगपूर्ण और निर्देशात्मक) संभावित लाभों का वर्णन नीचे किया गया है।

### बॉक्स 7: मनरेगा में नरेगा सॉफ्ट और ई-गवर्नेंस

नरेगा सॉफ्ट मनरेगा योजना का एक व्यापक ई-गवर्नेंस सिस्टम है। यह हितधारकों के एक रेंज के लिए सुलभ है और यह केन्द्र से लेकर पंचायत तक के सभी स्तरों पर समस्त मनरेगा कार्यों के पूर्ण प्रवाह को प्रग्रहित करता है। यह सिस्टम नागरिकों के सूचना के अधिकार को ध्यान में रखते हुए, उपस्थिति नामावली, पंजीकरण आवेदन रजिस्टर, नौकरी कार्ड/रोजगार रजिस्टर/उपस्थिति नामावली निर्गम रजिस्टर और उपस्थिति नामावली पावती रजिस्टर जैसे दस्तावेजों, जो अन्यथा जनता के लिए दुर्गम हैं, उपलब्ध कराता है।

सिस्टम के उपयोग में भाषा कोई बाधा नहीं है; इसका उपयोग कई स्थानीय भाषाओं में आसानी से किया जा सकता है। वास्तव में, यहां तक कि अनपढ़ भी इंटरफेस का उपयोग कर सकता है क्योंकि इसके टच-स्क्रीन कियोस्क मॉडल में ध्वनी और आइकन की सुविधाएं मौजूद हैं। इसे इस तरह से डिजाइन किया गया है कि इसका उपयोग कामगार, जो योजना के लाभार्थी हैं, से लेकर ग्राम पंचायतों तक, जिला कार्यक्रम समन्वयकों से लेकर बैंकों और डाकघरों तक के हितधारकों की एक रेंज द्वारा आसानी से किया जा सकता है। यहां तक कि जो नागरिक योजना के लाभार्थी नहीं हैं, वे भी पोर्टल पर उपलब्ध सूचनाओं को देख सकते हैं।

सॉफ्टवेयर में कई मॉड्यूल होते हैं, जो सभी कार्यकलापों और सभी हितधारकों से व्यापक रूप से संबंधित होते हैं। उदाहरण के लिए, जबकि कामगार प्रबंधन मॉड्यूल सभी कामगार-संबंधित सेवाओं की रीढ़ है, एक निधि प्रबंधन मॉड्यूल केन्द्रीय मंत्रालय से कामगार की जेब तक के धन के संचलन के मार्ग को ट्रैक करता है, एक शिकायत निवारण मॉड्यूल शिकायत दर्ज करने एवं उसके उत्तर को ट्रैक करने में अनपढ़ों के साथ-साथ हितधारकों की भी सहायता करता है और एक बैंक/पोस्ट ऑफिस मॉड्यूल वित्तीय संस्थानों को मजदूरी की जानकारी प्राप्त करने और खातों में जमा धन का विवरण दर्ज करने की अनुमति देता है।

## सरकार स्वयं लाभार्थी के रूप में

4.55 यदि सरकार प्रामाणिक डाटा और दस्तावेजों को तुरंत प्राप्त करने में सक्षम हो जाती है तो वह समावेशी और बहिष्करण दोनों प्रकार की त्रुटियां कम करके कल्याणकारी योजनाओं और सब्सिडी के लक्ष्य में सुधार कर सकती है। विलय किए गए डेटासेट के माध्यम से सार्वजनिक सेवा वितरण में भी सुधार किया जा सकता है। उदाहरण के लिए, जीएसटी रिटर्न के साथ रिटर्न का प्रति सत्यापन (क्रॉस-वेरिफिकेशन) करने से संभावित कर चोरी को उजागर किया जा सकता है।

## लाभार्थी के रूप में निजी क्षेत्र की फर्में

4.56 निजी क्षेत्र को व्यावसायिक उपयोग के लिए डाटाबेस सलेक्ट करने का अधिकार दिया जा सकता है। सार्वजनिक हित में डाटा के भाव को देखते हुए, लाभ के लिए इस डाटा के व्यावसायिक उपयोग को रोकने का कोई कारण नहीं है। निस्संदेह, यहाँ पर परिकल्पित डाटा क्रांति में काफी लागत आएगी। हालांकि, सामाजिक लाभ सरकार की लागत से कहीं अधिक होगा, लेकिन सरकारी वित्त पर पड़ने वाले दबाव को कम करने के लिए उत्पन्न आंकड़ों के कम से कम एक हिस्से को मुद्रिकृत किया जाना चाहिए। यह देखते हुए कि निजी क्षेत्र के पास इस डाटा से बड़े पैमाने पर लाभ प्राप्त करने की क्षमता है, उनसे इसके उपयोग के लिए शुल्क लेना उचित है।

4.57 उदाहरण के लिए, निजी क्षेत्र को किसी जिले के छात्रों के परीक्षा में प्राप्त अंक संबंधी डाटा (व्यक्तिगत जानकारी गुप्त रखते हुए) लेने की अनुमति देना। छात्रों के परीक्षा के अंकों, प्रत्येक जिले की जनसांख्यिकीय विशेषताओं और सार्वजनिक शिक्षा योजनाओं की प्रभावकारिता के संबंध में सार्वजनिक रूप से उपलब्ध

डाटा का उपयोग करके निजी फर्म शिक्षा में पूरी नहीं की जा सकी जरूरतों को उजागर करने में सक्षम हो सकती है और संबंधित जिलों की विशिष्ट आवश्यकताओं के अनुरूप अभिनव ट्यूटोरिंग उत्पादों को विकसित करके इन जरूरतों को पूरा कर सकती है। इन उत्पादों से न केवल निजी क्षेत्र को लाभ प्राप्त होगा, बल्कि इससे शिक्षा के स्तर और सामाजिक कल्याण में सुधार होने के अलावा, डाटा का मुद्रिकरण भी हो सकेगा और सरकार के लिए राजस्व भी उत्पन्न होगा।

4.58 वैकल्पिक रूप से, डाटासेट को विश्लेषण एजेंसियों को बेचा जा सकता है, जो डाटा का प्रसंस्करण करती हैं, अंतर्दृष्टि विकसित करते हैं और इन अंतर्दृष्टियों को आगे कॉर्पोरेट सेक्टर को बेच देती हैं, जो बदले में इस अंतर्दृष्टि का उपयोग करके मांग की भविष्यवाणी कर सकते हैं, अप्रयुक्त बाजारों की खोज कर सकते हैं या नए उत्पादों में नयापन ला सकते हैं। हर तरह से निजी क्षेत्र को डाटा से लाभान्वित होने की बहुत गुंजाइश है और उन्हें शुल्क के आधार पर ऐसा करने की अनुमति दी जानी चाहिए। सौभाग्य की बात है कि निजी क्षेत्र को डाटा से लाभ प्राप्त करने की अनुमति देने के बावजूद भी निजता और गोपनीयता की रक्षा के लिए सख्त तकनीकी तंत्र मौजूद है।

## लाभार्थी के रूप में नागरिक

4.59 नागरिक प्रस्तावित डाटा क्रांति के लाभार्थियों का सबसे बड़ा समूह हैं। डिजिटल लॉकर के मामले पर विचार करें। यह कई महत्वपूर्ण रूप में उस योजना के समान है जो हमने ऊपर उल्लिखित किए हैं, लेकिन यह राज्य द्वारा जारी किए जाने वाले कुछ दस्तावेजों तक की सीमित है। नागरिकों को अब राज्य से “मूल” दस्तावेजों जैसे कि उनका ड्राइविंग लाइसेंस, आधार कार्ड, पैन कार्ड

### बॉक्स 8 : राष्ट्रीय छात्रवृत्ति पोर्टल

भारत सरकार ने पहले ही छात्रवृत्ति संबंधी डाटा को एकीकृत करने की दिशा में आगे कदम बढ़ा लिया है। केंद्रीय और राज्य स्तरों पर विभिन्न मंत्रालयों द्वारा कार्यान्वित सभी छात्रवृत्ति योजनाओं के समांजस्य के लिए राष्ट्रीय छात्रवृत्ति पोर्टल शुरू की गई थी। यह पोर्टल, छात्र आवेदन, आवेदन रसीद, प्रसंस्करण और निधि के वितरण के लिए मंजूरी से लेकर छात्रवृत्ति से संबंधित सभी सेवाओं के लिए एक व्यापक आधार के रूप में कार्य करता है। यह पोर्टल, विभिन्न स्तरों पर समस्त सरकारी छात्रवृत्ति योजनाओं के सभी लाभार्थियों का एक पारदर्शी डाटाबेस बनाने के अलावा, छात्रवृत्ति का पता लगाने में आने वाली अड़चनों को कम करता है और प्रत्यक्ष लाभ अंतरण (डीबीटी) को सुविधाजनक बनाता है।

सीबीएसई परिणाम आदि प्राप्त करने के लिए जगह-जगह भाग-दौड़ नहीं करनी होगी। ये दस्तवेज भारत के हर निवासी के जीवन में महत्वपूर्ण हैं। इन दस्तावेजों की सबसे ज्यादा जरूरत उन लोगों को होती है जो कल्याण के लिए राज्य पर निर्भर होते हैं। प्रायः इन्हे प्राप्त करना, कमजोर वर्गों के लिए अत्यंत कठिन हो जाता है।

4.60 डिजिटल क्लाउड में एक ही स्थान पर एक सत्यापित प्रारूप में इनके समस्त प्रलेखों को उपलब्ध कराता है। इस लॉकर में एक्सेस करने के लिए नागरिकों को केवल इंटरनेट से जुड़े किसी उपकरण, स्मार्टफोन या कम्प्यूटर की जरूरत होती है। यह डाउनस्ट्रीम प्रक्रिया को डिजिटल बनाने, जैसे कॉलेजों में एडमिशन में सहायता करता है। ऐसे किसी भी व्यक्ति के लिए, जिसे सरकार से किसी गुम या लापता प्रलेख को पुनः प्राप्त करना हो अथवा छायाप्रतियां “अनुप्रमाणित” करानी हों, डिजिटल सुविधा का अनुभव अत्यधिक समय बचाने वाला और प्रयोक्तानुकूल होता है। उन्हें इस भय के साये में रहना नहीं पड़ता कि उनके कीमती कागजात गुम होकर गलत हाथों में पड़ जाएंगे या कोई अन्य गड़बड़ी हो जाएगी।

4.61 इसी दिशा में अनुरूप, भारतीय रिजर्व बैंक ने नॉन-बैंकिंग फाइनेंशियल कंपनी-अकाउंट एग्रीगेटर (एनबीएफसीए) की घोषणा की है। वित्तीय उद्योग के अंतर्गत भी ऐसा कोई भी व्यक्ति, जिसका डाटा अनेक प्रदाताओं के पास हो, वित्तीय जीवन की जटिलताओं में उलझा रहता है। हो सकता है कि किसी व्यक्ति का भारतीय स्टेट बैंक में खाता हो किंतु वह ऋण एचडीएफसी से लेता है। उसके म्यूचुअल फंड निवेश

एक्सिस बैंक के माध्यम से हों, किंतु उसका बीमा एलआईसी के माध्यम से किया जाए। हो सकता है कि उसके पास स्टैंडर्ड चार्टर्ड का क्रेडिट कार्ड हो, किंतु कुछ स्टॉक्स में प्रत्यक्ष निवेश के लिए वह मोतीलाल ओसवाल का इस्तेमाल करता हो। ऐसे मामले में उस व्यक्ति द्वारा अपने वित्तीय जीवन से जुड़े और संयुक्त किए जाने वाले आंकड़े अनेक डाटा प्रदाताओं के पास प्रकीर्ण रूप में होते हैं। एनबीएफसी-ए नागरिकों द्वारा अपेक्षित किसी भी प्रयोजन के लिए, प्रयोक्ताओं को एक साथ आंकड़े प्राप्त करने की सुविधा प्रदान करता है। यह वैयक्तिक वित्तीय प्रबंधन के लिए हो सकता है। अथवा किसी आवास ऋण के लिए डिजिटल रूप में आवेदन करने के लिए भी हो सकता है। एनबीएफसी-ए न तो डाटा को रीड करता है और न ही किसी दुरुपयोग (घुसपैठिए) के लिए संपूर्ण डाटा-सेट का सृजन करता है। यह नागरिकों को इन संस्थाओं से केवल एक मशीनी रूप में पठनीय प्रारूप में उनके आंकड़ों को प्राप्त करने की सुविधा प्रदान करता है ताकि उनके द्वारा इनका उपयोग सार्थक रूप में किया जा सके।

### भावी परिदृश्य

आधार के माध्यम से भारत डाटा एवं प्रौद्योगिकी क्रांति की गुणात्मक वृद्धि के साथ अग्रणी पंक्ति में रहा है। चूंकि सामाजिक कल्याण हेतु डाटा का सृजन इष्टतम परिमाण में निजी क्षेत्र द्वारा नहीं किया जा सकता अतः सरकार को लोगों की भलाई के साधन के रूप में डाटा को देखना चाहिए और आवश्यक निवेश कार्य करना चाहिए। लोक कल्याण के रूप में डाटा के सृजन के लाभों की प्राप्ति डाटा की निजता की कानूनी रूपरेखा के भीतर की

### बॉक्स 9: राष्ट्रीय स्वास्थ्य रजिस्ट्री का विचार

स्वच्छ भारत से स्वस्थ भारत में और उसके बाद फिर सुंदर भारत में रूपांतरण के लिए नागरिकों का स्वास्थ्य सर्वाधिक महत्व रखता है। इस कवायद में इलाज के लिए प्रतीक्षा करने की बजाय एहतियात बरतना अधिक महत्वपूर्ण होगा। एक राष्ट्रीय स्वास्थ्य रजिस्टर, जिसमें निजता संबंधी समस्त आवश्यक सुरक्षाओं के साथ नागरिकों के स्वास्थ्य से संबंधित अभिलेख रखे जाएं, भविष्य सूचक एवं निर्देशात्मक प्रयोजनों के लिए स्वास्थ्य संबंधी विश्लेषण में समर्थ बनाने की दिशा में दूरगामी परिणाम प्रदान करेगा। ऐसे राष्ट्रीय स्वास्थ्य रजिस्टर का कार्यान्वयन नागरिकों के आधार नंबर के प्रयोग के साथ किया जाएगा। जब कोई चिकित्सक इस राष्ट्रीय स्वास्थ्य रजिस्टर से किसी रोगी के चिकित्सा इतिवृत्त को जान सकेगा तो यह सुविधा आपातकालीन/ट्रॉमा मामलों में विशेष रूप से उपयोगी होगी और संभवतः इसके माध्यम से अनेक लोगों का जीवन बचाया जा सकेगा। इस रजिस्टर के विभिन्न संघटकों के अंतर्गत (प) अस्पतालों और सार्वजनिक स्वास्थ्य केंद्रों, (पप) सिंडोम की गिरानी, (पपप) टीकाकरण सूचना प्रणालियों, (पअ) इलेक्ट्रॉनिक लैबोरेट्री रिपोर्टिंग, और मधुमेह, उच्च रक्तचाप, कैंसर, एड्स जैसे, दखल की अपेक्षा रखने वाले मुख्य रोगों के लिए उप-रजिस्ट्रों के डाटाबेस को सम्मिलित किया जा सकता है। इस रजिस्टर से छद्मनाम आंकड़ों को निजी पक्षकारों को विश्लेषणार्थ बेचा जा सकता है, जिसके फलस्वरूप भविष्यसूचक और निर्देशात्मक ज्ञान की उपलब्धता से निवारण उपायों का संवर्धन हो सकता है।



जा सकती है। आगे बढ़ते हुए, डाटा और सूचना हाइवे को भौतिक राजमार्गों की भांति समान रूप से महत्वपूर्ण अवसंरचना के रूप में देखा जाए। ऐसे कदम से भारत को अपने लोगों के कल्याण हेतु प्रौद्योगिकीय प्रगति

के लाभों का गुणात्मक उपयोग करने में सहायता मिल सकती है। इसीलिए, भारत के संविधान की मूल-भावना के अनुसार डाटा (आंकड़े) “लोगों के, लोगों द्वारा लोगों के लिए”, सरकार का मूलमंत्र होना चाहिए।

### अध्याय पर एक नजर

- डाटा के एकीकरण और संचयन में प्रौद्योगिकीय प्रगति को देखते हुए समाज का इष्टतम डाटा उपयोग पहले से अधिक हुआ है।
- चूंकि हो सकता है कि निजी क्षेत्र लाभकारी डाटा का लाभ प्राप्त करने में निवेश न कर सके अतः सरकार को लोगों, विशेषकर गरीबों और देश में सामाजिक क्षेत्र के कल्याण के लिए डाटा का सृजन करने में हस्तक्षेप करना चाहिए।
- सरकार के पास पहले से ही प्रशासनिक, सर्वेक्षण, संस्थागत और नागरिकों के बारे में लेन-देन संबंधी डाटा का एक समृद्ध भंडार उपलब्ध है, परंतु यह डाटा अनेक सरकारी निकायों में प्रकीर्ण रूप से बिखरा हुआ है। इन पृथक-पृथक विशिष्ट डाटा-समूहों को एकरूप करने के गुणात्मक लाभ होंगे और इनका असीमित उपयोग किया जा सकेगा।
- गोपनीय जानकारी के संरक्षण और उसे सांझा करने के लिए अतिविकसित प्रौद्योगिकी पहले ही उपलब्ध है। अतः डेटा निजता के परिप्रेक्ष्य से डेटा को सार्वजनिक पदार्थ बनाया जा सकता है। इस प्रकार इसे सार्वजनिक स्वरूप प्रदान करते समय यह ध्यान रखना आवश्यक है कि निजता की अभिजात्य वरीयताएं समाज के बरीब वर्ग पर थोपी नहीं जाएं गरीब वर्ग को तो अपनी जीवन दशा में सुधार ही महत्वपूर्ण लगता है।
- चूंकि सामाजिक हित के डाटा का सृजन लोगों द्वारा किया जाता है अतः यह “लोगों का, लोगों के द्वारा और लोगों के लिए” होना चाहिए।

### संदर्भ

Brook, E., Rosman, D., Holman, C. 2008. “Public good through data linkage: measuring research outputs from the Western Australian data linkage system.” *Australian and New Zealand journal of public health* 32(1): 19-23.

Dresner Advisory Services, LLC. 2017. *Big Data Analytics Market Study*. [https://www.microstrategy.com/getmedia/cd052225-be60-49fd-ab1c-4984ebc3cde9/Dresner-Report-Big\\_Data\\_Analytic\\_Market\\_Study-WisdomofCrowdsSeries-2017.pdf](https://www.microstrategy.com/getmedia/cd052225-be60-49fd-ab1c-4984ebc3cde9/Dresner-Report-Big_Data_Analytic_Market_Study-WisdomofCrowdsSeries-2017.pdf)

Hilbert, Martin. 2016. “Big Data for Development: A Review of Promises and Challenges.” *Development Policy Review* 34:135-174.

James M., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., and Angela Hung Byers.

2011. *Big Data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

[https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI\\_big\\_data\\_exec\\_summary.ashx](https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.ashx)

Lane, J., Stodden, V., Bender, S. and H. Nissenbaum. 2014. *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. New York: Cambridge University Press. [http://assets.cambridge.org/97811070/67356/frontmatter/9781107067356\\_frontmatter.pdf](http://assets.cambridge.org/97811070/67356/frontmatter/9781107067356_frontmatter.pdf)

Ministry of Electronics & Information Technology. 2019. “Open Government Data Portal: An Overview.” Accessed June 20. [https://www.meity.gov.in/writereaddata/files/OGD\\_Overview%20v\\_2.pdf](https://www.meity.gov.in/writereaddata/files/OGD_Overview%20v_2.pdf).

Nilekani, Nandan. 2018. “Data to the People. India's Inclusive Internet.” *Foreign Affairs* 97(5).

Ritchie F, and Richard Welpton. 2011. “Sharing risks, sharing benefits: Data as a public good.” Presentation at Joint UNECE/Eurostat work session on statistical data confidentiality, Tarragona, Spain, October 26-28. [http://eprints.uwe.ac.uk/22460/1/21\\_Ritchie-Welpton.pdf](http://eprints.uwe.ac.uk/22460/1/21_Ritchie-Welpton.pdf)

Rodwin MA, and JD. Abramson. 2012. “Clinical trial data as a public good.” *JAMA* 308: 871-2.

Taylor, L. 2016. “The ethics of big data as a public good: Which public? Whose good?” *Philosophical Transactions of the Royal Society* 374: 1–13.