Big Data

Mining a national resource

Since the Obama Administration announced the "Big Data Research and Development Initiative" in 2012, a number of nations across the globe have set in motion their national Big Data plans. Improving the government's ability to extract insights from a large volume of information and complex data sets is also increasingly topping the national agendas of governments in the Middle East.

The regulatory debate

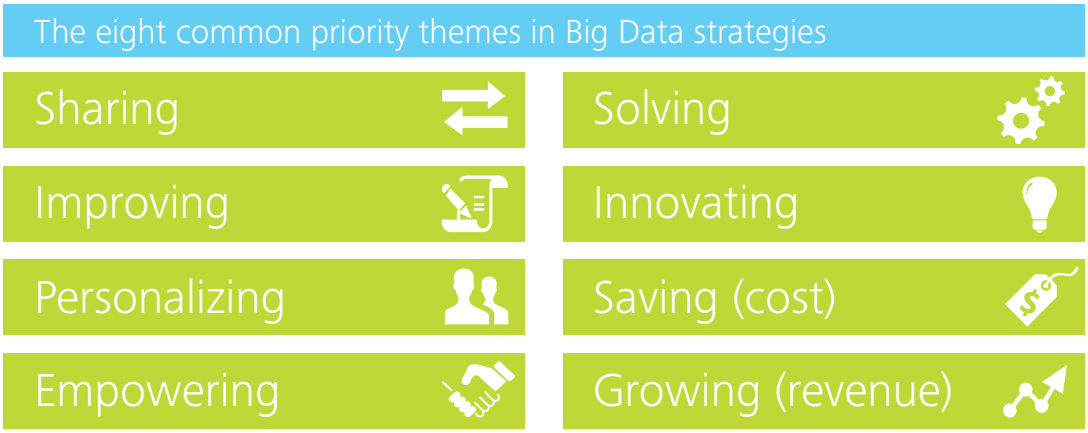
Big Data usage, by private and public entities, is coming under increasing scrutiny from regulators and policy makers, with data privacy clearly top of mind in terms of key concerns. While there are limited instances of "Big Data regulation" internationally it has become an increasingly important theme for regulators within and without the Middle East region.

The challenge is twofold. First, how can the government leverage the massive amounts of data available within its different entities to deliver better service, formulate more predictive policies and reduce operational inefficiencies? Secondly, how can regulators set up a regulatory framework that will encourage usage of Big Data within the private sector and at the same time provide enough assurance for the citizens that their data is not being misused?

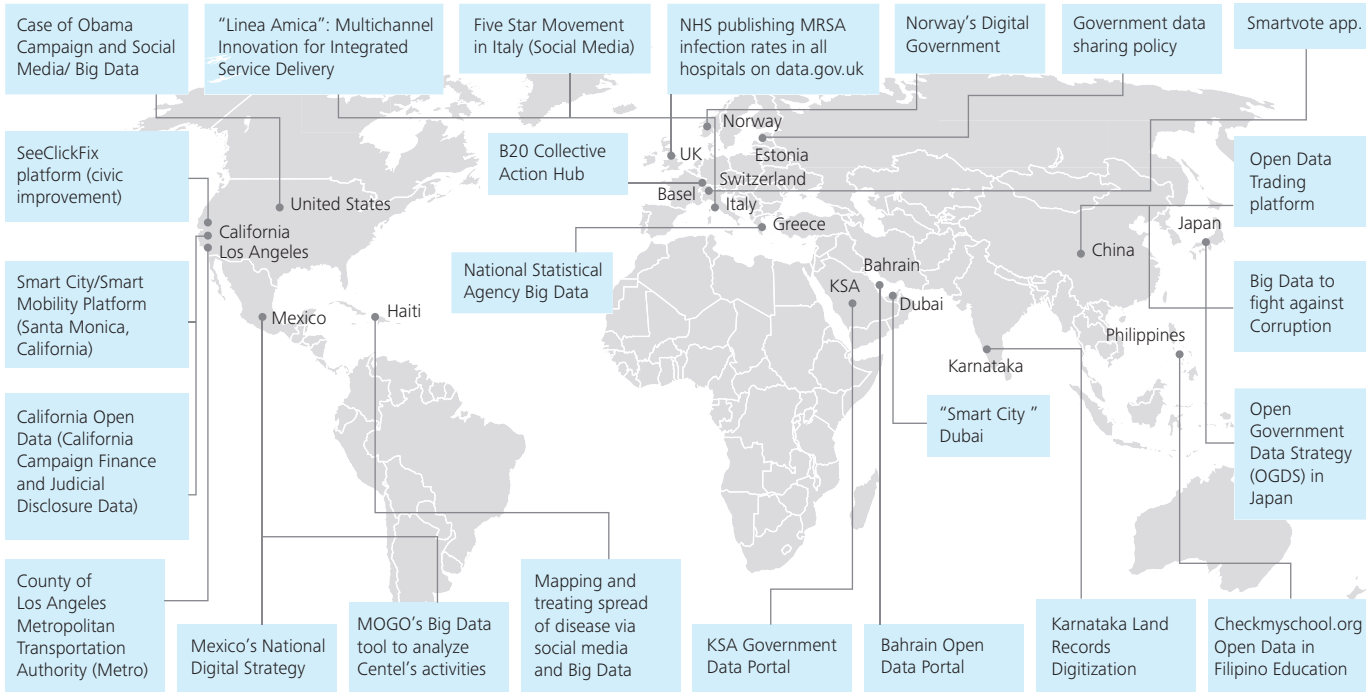
The stars are aligned – Big Data strategy to support the national vision

Globally, governments have started implementing Big Data programs, either as sectorial initiatives, or in some cases, as national plans. Eight broad themes and objectives seem common across the various programs we have surveyed:

1. Sharing data sets among government institutions;
2. Improving policy-making, service delivery and operational efficiency via the use of Big Data tools;
3. Personalizing government services to the needs of citizens;
4. Empowering citizens to make more informed decisions;
5. Solving complex public policy issues via the analysis of large, multi-dimensional data sets;
6. Promoting innovation via the emergence of new business models based on large government data sets;
7. Reducing the transactional costs and redundant expenditures across government agencies; and more occasionally,
8. Benefiting from Big and Open data to increase revenues (e.g. tax income.)



Select national Big and Open Data governmental initiatives



Source: Various government websites, Monitor Deloitte analysis

In some cases, a Big Data Strategy at national level has been developed to serve as a support for the implementation of national plans. For instance in Singapore, the Big Data national strategy focuses on infrastructure building to position Singapore as a "Smart Nation" and a Big Data hub for other countries. Estonia has been using ICT (Internet and Communications Technology) for a variety of services. It was the first country to allow online voting–now 25 percent of voting takes place electronically and 99.6 percent of banking transactions happen online. Its "Digital Agenda 2020" outlines the need to study innovations like Big Data to deliver services to its citizens.

In the Middle East region, we are seeing the emergence of national Big Data strategies principally focused on three needs of government:

a. To improve the efficiency of existing services by ensuring that data is not requested multiple times for one individual;
b. To support policy making by data analysis; and
c. To use Open Data to enhance engagement with the public and businesses.

In the Middle East region, we are seeing the emergence of national Big Data strategies principally focused on three needs of government

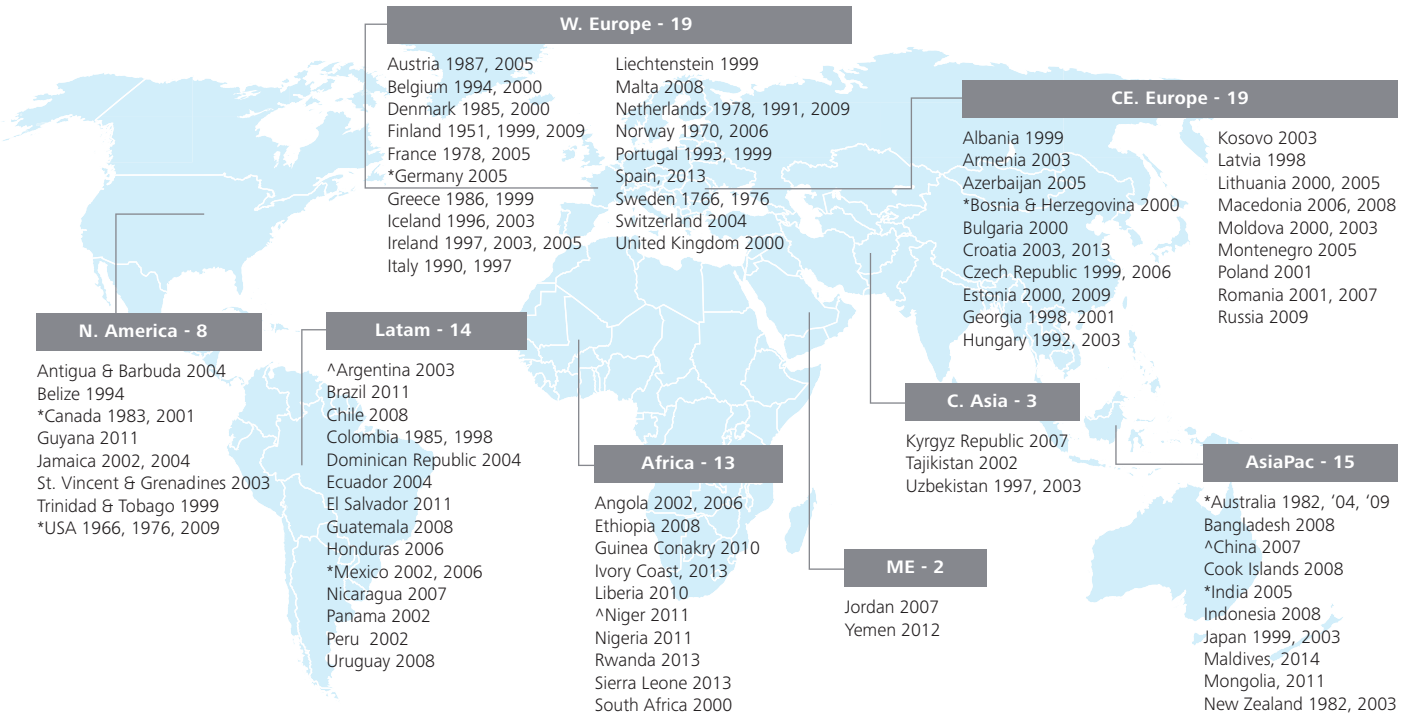
In the Middle East North Africa region, the first challenge to Big and Open Data regulation has been the lack of underlying legal instruments

Open vs. Private – the challenge of opening up data

As for other national resources, the exploitation of Big and Open Data needs to be carefully governed, particularly in the early days of adoption. The primary legal, ethical or sometime national security challenges come from the degree of openness of data: which data sets can or should be shared and under which terms?

Historically, in many markets, there have been two sets of policies defining the perimeters of this issue: privacy laws and freedom of access to information legislation. Privacy laws are becoming increasingly widespread. The number of privacy laws grew from 20 in the 1990s to over 100 now. However, not all countries are equal in their data privacy law maturity. Examples of countries with a comprehensive privacy law include all European member states (27 countries), Singapore, and Canada. Countries such as South Africa and Australia are also in advanced stages of creating legislation. A number of countries such as the United States have sectorial laws governing data privacy in particular areas (such as children's medical records or financial information.) Freedom of access to information (FoAI) laws on the other hand, allow access by citizens to data held by national governments. They establish a "right-to-know" legal process by which requests may be made for government-held information. This type of legislation is as widespread as privacy laws although their prevalence in the Middle East is very low (only two countries: Jordan and Yemen.)

Freedom of Access To Information (ATI) laws



° Law used primarily to restrict ATI and freedom of expression.
^Actionable ATI regulation rather than a law.
* Federal system: several states (also) have ATI laws.

Source: Various government websites, Monitor Deloitte analysis

In contrast with data privacy laws, Open Data laws are still relatively few as their origins are more recent (post-2010) and include laws such as the 2014 Digital accountability and transparency act in the United States. In many cases, instead of full fledge laws, Open Data policies were developed, for instance the EC directive 2003/98/EC on the re-use of public sector information, which took root in the FoAI laws mentioned above.

In the Middle East North Africa region, the first challenge to Big and Open data regulation has been the lack of underlying legal instruments. Privacy laws are being developed but do not always have a track record of implementation and the limited number of countries with FoAI frameworks makes it even more difficult to have Open Data laws or data sharing policies in place. Overall the region still has a long way to go on data sharing as demonstrated by the current low ranking of most Middle East nations, with all Gulf Cooperation Council (GCC) countries ranked below the 50th global rank in the Open Data Barometer. However, some interesting initiatives are underway across the region–the UAE and Qatar have plans to launch or revamp their open data portals and Saudi Arabia has recently launched a new Open Data portal (http://www.data.gov.sa/).

Managing data access

Open data is however only one component in the vast realm of applications of government Big Data. Sharing information among government institutions can also lead to significant value creation across one of the eight themes discussed above. Nevertheless, with large amounts of data available but also scattered among different entities, it is critical to establish a structured framework to collect, analyze and share the data.

Open data is however only one component in the vast realm of applications of government Big Data. Sharing information among government institutions can also lead to significant value creation.

Sharing data within government entities or between the public and private sectors has promoted the creation of innovative products and services. This becomes increasingly of value when developing smart cities, since smart city services will require automated interactions between public and private sector companies through standard data interfaces. For example when London released its maps of the tube to the public, new online and mobile applications were created by entrepreneurs to improve the usage of the tube.

The flipside to sharing data this freely is security. Countries are often struggling to strike the balance between improving services and exposing the country to threats unintentionally. It is well-known that multiple streams of open data can be overlaid to deduce potentially sensitive information. For example, it may be possible to deduce the destination and movement of oil supplies via ships from oil production data and port vessel movement data.

Another issue that needs to be considered when addressing the regulatory framework of Big and Open Data is the price tag of data

Prevention, as the old adage recommends, is better than a cure. Governments can mitigate these risks by adequately whetting the data through a consistent and coherent framework. It starts with putting the data broadly in buckets that mark it as open or closed depending on the national or institutional policy. Buckets that are classified as potential for Open Data then need to be assessed for consumption by the public by ensuring that sensitive private or national data is rendered anonymous, or removed. The data should ideally then go through at least one more round of checks by qualified data scientists that look at the risks of releasing the data sets. The broad selection framework should be tailored to the government's policies.

Regulating data monetization

Another issue that needs to be considered when addressing the regulatory framework of Big and Open Data is the price tag of data. The general consensus is that Open Data is free to access and use, but not all organizations are willing to give away their data for free. Organizations often consider their data to be part of their Intellectual Property and are reluctant to forego additional revenues for the cause of open data. Telecom companies are a case in point as they typically sit on a variety of structured (call data, geo-location data) and unstructured data (web browsing logs, social media posts, app usage patterns, etc.) that has high value when analyzed. For example, retailers are usually willing to pay for information that tells them the density and profile of customers in the vicinity of their location during different times of the day. However, these and other types of geo-location data are very useful in providing public services as well. In Sierra Leone, telecom data has been used to predict day-to-day outbreaks of Ebola and take preventative action to contain the disease. This natural tension between opening data and monetizing it will need to be carefully thought through by governments in the region as they weigh the potential costs and benefits of a free versus a pay model for data access.

Capability building: data scientists are the new unicorns!

With an increasing role as Big Data regulators and promoters, a key challenge for governments in the region will be capability building. Big Data is bringing about a new breed of roles and job positions both in the core business and in the Information Technology (IT) departments of public and private organizations.

The analysis of real-time structured and unstructured data requires not only sophisticated systems but also new resources that did not exist several years ago. The presence of data owners, governors and stewards among others is as much part of the sophisticated

system as the mining tools, business intelligence applications, and server and network infrastructure. The emerging positions and roles are becoming more difficult to fill due to the increasing need to have both business and technical knowledge and skills. Moreover, the limited existence of skilled professionals currently in the industry's workforce hinders effective knowledge and skill transfer. This is eventually leading to increased reliance on contractors and consultants, which in turn challenges organizations in terms of skill retention and future capability building.

Going forward, it will be crucial for governments to invest in building their capabilities around Big Data if they want to achieve tangible results from their Big Data strategy. Although this will involve significant investment in capturing and retaining key talents, it will also demand stronger awareness of Big Data and its operations among the government leadership as well as Human Resource managers and talent specialists. The profiles required for the realization of national Big Data strategies are not only sophisticated but also often a unique blend of skills (e.g. Physician crossed with Data specialist.) Data Scientists are the new unicorns of the talent pool.

by **Emmanuel Durou**, Partner, Technology, Media and Telecommunications, Monitor Deloitte

Going forward, it will be crucial for governments to invest in building their capabilities around Big Data if they want to achieve tangible results from their Big Data strategy