DOI:10.1145/2398356.2398368

Deven Desai

Law and Technology

Beyond Location: Data Security in the 21st Century

Viewing evolving data security issues as engineering problems to be solved.

The continued attention to data protection and the growth of cloud computing highlight tensions among data protection regulators, businesses, and the computer science communities. As new data protection laws are proposed, these groups have the chance to share insights and achieve their respective goals; but right now, with respect to data security, they may be passing by each other.

Data protection laws seek to protect user rights and rely in part on a certain view of data location and related security practices to ensure those rights are maintained. In simplified terms, data protection laws tend to focus on data not leaving a country or region as part of a given data protection regime. Businesses apply cloud techniques for a range of purposes. Some ends are internal such as improved network operations; some are external such as selling storage and services to other businesses. In either case, advances in, and the future of, cloud computing rely on moving data on an almost continuous basis. Thus, the political and business interests seem to be set to collide. That collision is not, however, inevitable. Does data protection require keeping data in one place? Is data security enhanced or harmed by such an approach? Does jurisdiction have to turn on data location? By parsing what is at stake for location and jurisdiction and what cloud computing may offer for security, we should be able to fashion laws that respect the political interests in data protection and that draw on the best insights of computer science to achieve heightened data security.

Possibly Competing Interests

Governments and businesses have legitimate, competing interests in data management. For example, they disagree about what to do with cloud computing. Sometimes the debates devolve into accusations that the other side "doesn't get it." Yet, if we start by stating what those interests are, we should be able see where the interests intersect or diverge. Once that is done, we can see whether there is a way to bridge remaining gaps.

Although there are many different data protection laws, the European Union's approach provides a way to understand government interests and possible mistakes on the horizon. Unfortunately, mandated data location serves two, conflicting purposes. On the one hand, it allows for an exercise of jurisdiction based on the idea that data stored in a particular jurisdiction is subject to the laws of that place. The need for jurisdiction is real. Governments want to be able to reach out and touch our data. They also want to enforce laws to protect their citizens and their data. On the other hand, the EU's previous Data Protection Directive (DPD) and current, proposed General Data Protection Regulation (GDPR) seek to prevent unauthorized access to and, by extension, use of data. For example, Article 30 of the GDPR requires that those responsible for data processing take "appropriate technical and organizational measures to ensure a level of security appropriate to the risks represented by the processing and the nature of the personal data to be protected." It also requires those responsible for data "protect personal data against accidental or unlawful destruction or accidental loss and to prevent any unlawful forms of processing, in particular any unauthorized disclosure, dissemination or access, or alteration of personal data."

The location problem arises because the current and proposed approaches employ complicated rules about data location, storage, and movement to achieve the protection goals. In addition, the laudable goals of Article 30 inadvertently run into the realities of the latest security advances in cloud computing. For example, in a recent decision in the EU, data location requirements interfered with a city's ability to use modern cloud computing services. [a]

What then is cloud computing and how does it relate to data security? First, one can think of the security problem as the ways in which someone could gain unauthorized access to data. That view comports with Article 30. A perhaps somewhat misunder-

a See Notification of decision—New email solution within Narvik local authority (Narvik kommune)—Google Apps, Norwegian Data Inspectorate, reference number 11/00593-7/SEV, January 16, 2012 (denying a request to use Google Apps for email and other service based in part on location of data and methods for data storage concerns).

stood issue is from where the threat of unauthorized access comes. Given some recent stories about security breaches, some may think the largest threat for unauthorized access is at data centers which, like banks, might be targeted and then breached. Yet, losing computers and thumb drives is a major way that data is lost, perhaps more so than through a data center.[b] It is the fact of walking data—that is, data on a portable device—that leads to cloud computing as a step forward in security practices. Cloud services address many data security issues, especially by reducing threats from loss of a personal device. But not all cloud services are the same. A big problem for any data protection law is that the type of cloud service and the way data is managed varies greatly depending on how the provider manages the backend *and* what the customer is doing.

Some cloud computing is distributed computing. The data may be sharded across many servers; copies are made of documents and then split across servers. Instead of a single point of failure where all your data happens to be on the server, your data is spread out on many servers. That model can apply to in-house data management or when providing services to others. In simplest terms the provider may manage all service in-house or use other parties as well, but pinning down exactly what mix of data management is in play requires knowing the configuration for a given service.

New work on optimization means that the best cloud services will likely *move data often, if not all the time*, to address issues such as overheating that threatens servers, bandwidth, packet loss, power, resources, and other "failure modes" as well as to be more efficient with their networks as cycles of computing fluctuate. Thus moving data is part of securing the data against loss, which is another goal of Article 30. In addition, customers may use a cloud service on an ongoing or a temporary basis, but others run the service. For example, with Amazon's E2C offering, the cloud is a commodity service, which allows a company to buy services for a short period to accomplish a large task. When the *New York Times* converted 11 million articles to .pdf, it used a rented cloud for a day to finish the work and at a much lower cost than having its own data center for the task. All of these variables challenge any attempt to mandate security protocols based on location, because companies, customers, and even governments will not know ahead of time what option they want to pursue or they will be forced to choose older, slower, and costlier approaches to data for the sake of compliance.

Governments want to be able to reach out and touch our data. They also want to enforce laws to protect their citizens and their data.

Global cloud services present additional problems. With multiple jurisdictions involved, anyone offering or using cloud services could face location requirements for each country in which they operate. To the EU's credit, it is trying to address that criticism as it applies to the current DPD. The proposed GDPR will be binding on all members. The current state-by-state approach under the DPD would go away in favor of a harmonized approach to data with the regulation being implemented in its entirety and taking effect even without a member state taking action to put the regulation into national law.

b See, for example, 2010 Annual Study: Global Cost of a Data Breach, May 2011; http://www.symantec.com/content/en/us/about/media/pdfs/symantec_cost_of_data_breach_global_2010.pdf.

Even if the EU harmonizes its data laws, the focus, however, is still on data as residing in one place or within a region. The EU is big enough that one might think location problem is not an issue. One could set up data centers across the EU and move the data within the system. With one law to govern, it will all work out. That view misses the fact that, like a power grid, data networks have cycles of demand and manage that demand dynamically. If there is idle capacity in a region, it may be useful for service outside the region. If there is a demand spike in the region, capacity from outside the region may be used to meet the demand. Location-based data rules clash with these realities. Furthermore, many countries are copying the EU approach to data protection. The EU is a large region and market; Singapore is not. Nor is Vietnam, Costa Rica, Egypt, Peru, Ghana, or most single countries. From a security perspective, location-based rules falter in a large area such as the EU; they fail in smaller markets. Nonetheless, governments have a real need to protect their citizens' data and to have a legal process to obtain data. Location models, however, do not achieve these goals well.

From a security perspective, location-based rules falter in a large area such as the EU; they fail in smaller markets.

Possible Ways Forward

Neither mandated location compliance by governments nor claims of inability to comply with laws by businesses provides satisfactory results; but some options are available. Part of a possible solution is to abandon the location-based aspects of data security laws. Indeed, forcing companies to limit data movement ignores the reality of data management and can increase the security threat rather than reduce it. Unraveling the jurisdiction question is more difficult. As Jack Goldsmith and Tim Wu point out, governments will always find a way to exert power to shape the world as they see fit, so places and borders still matter.[1] This point is already seen in current laws in the U.S. and Europe, which allow governments to access data when investigating national security or terrorism crimes. Thus, countries or regions may fashion new data laws that move beyond location to determine jurisdiction and demand data for areas beyond national security and terrorism. If so, companies will have to find ways to comply. The time when a company could stick its data-head in the sand of one country and reject other countries' laws may be over precisely because of government needs, global computing services, and advances in data security and networking. If companies wish to have the flexibility to employ different data management methods and especially ones that involve continual movement of data, they cannot simultaneously argue that no law or method covers how and when a government may gain access to data. Yet, it is this need to comply that may undermine the trust of one country over another. For example, Country X may be comfortable with data moving to Country Y, but not Country Z, because Country Z has a history of forcing companies to divulge data. All of which presents an opportunity.

As data security laws evolve, governments, companies, and computer scientists will have to work together to create a data security system for the 21st century. A key hurdle is identifying when any government may demand data. Transparent policies and possibly treaties could help better identify and govern under what circumstances a country may demand data from another. Countries might work with local industry to create data security and data breach laws with real teeth as a way to signal that poor data security has consequences. Countries should also provide more room for companies to challenge requests and reveal them so the global market has a better sense of what is being sought, which countries respect data protection laws, and which do not. Such changes would allow companies to compete based not only on their security systems but their willingness to defend customer interests. In return companies and computer scientists will likely have to design systems with an eye toward the ability to respond to government requests when those requests are proper. Such solutions may involve ways to tag data as coming from a citizen of a particular country. Here, issues of privacy and freedom arise, because the more one can tag and trace data, the more one can use it for surveillance. This possibility shows why increased transparency is needed, for at the very least it would allow citizens to object to pacts between governments and companies that tread on individual rights. Nonetheless, distributed computing techniques and encryption even with some type of tracing system may be better protection than relying on data residing in one place. After all, if data is stored in one place and a government knows where the data is, a government can simply walk off with the server. So far, however, the one group that could explain whether these ideas or others are viable—computer scientists—is missing from this interplay.

The nuances and possibilities of how we choose to manage data present computer science with the chance to inform policymakers about how best to meet their interests while simultaneously meeting the needs of business and individuals. Providing a better understanding of how cloud computing, in its range of manifestations, operates can only improve how the government shapes data policy. Governments should also explain their needs to computer scientists. By presenting issues as engineering problems to be solved, governments would likely stimulate the desire to meet a challenge. Together, governments, businesses, and computer scientists ought to be able to leverage advances in technology so all may benefit. Removing the focus on data location as way to increase data security is hopefully just one step in that direction. C

Reference

1. Goldsmith, J. and Wu, T. *Who Controls the Internet? Illusions of a Borderless World.* Oxford University Press, 2006, 180–181.

Deven Desai (devenrdesai@gmail.com) is a law professor at the Thomas Jefferson School of Law in San Diego, CA, and recently completed serving as the first Academic Research Counsel at Google, Inc.