Data Rights
Single vs. Multiple Ownership?

Organizations that are interested in accurate and dependable decision support systems; accurate metrics; descriptive, predictive and prescriptive analytics; cause-and-effect analysis; and/or if the enterprise or business process is data driven should value data quality, in conjunction with data confidentiality, integrity and availability. If data quality is important to an organization, then it should be concerned with the data sources. And if it is concerned with the data sources, then understanding data ownership rights and data governance should be of the utmost importance to an organization. The ability to identify rightful data ownership in relation to data quality is necessary for any organization that values data quality.

Data have multiple owners, as opposed to a single owner, most of the time, and that data ownership type and data source could directly relate to the quality of the available personal and/or enterprise data themselves. The following discusses the types, categories and data use today in determining the rightful ownership of data in this digital age.[1] One question that comes to mind immediately when discussing rightful data ownership is whether one can logically and legally claim that personal or enterprise data belong to one or more of the following:

- The person or object from which the data are generated
- The originator or generator of the data
- The owner or custodian of the system(s) through which data are processed and/or stored
- The data's external custodian in the case of cloud computing or otherwise

Attempts to answer the aforementioned question leads one to see that the issue of determining the proper ownership of data is as complex as the data themselves.

Definitions

Determining the rightful owner of data requires an understanding of data. Moreover, data have taken many definitions in recent years, especially with the advent of big data. Many have difficulty separating its true definition from the meaning of information. Thus, data can be defined as a distinctive fact or set of facts or distinctive pieces of information. In general, data can also be defined as information collected by researchers in the form of responses to survey instruments, questionnaires, interviews, videotapes and the like.[2] Other definitions of data are rooted in their role in the representation of facts about the world.[3] In a more complex sense, data can be structured, semi-structured and unstructured.

Personal Data

Personal information or data is described as a person's first or last name that is combined with

Patrick I. Offor, Ph. D., SAP Application Associate
Is an associate faculty member at the City University of Seattle (Washington, USA) and the chief of operations at the Production Support Branch, Logistics Data Analysis Center, US Army Materiel Command, Huntsville, Alabama, USA. He is also an active duty chief warrant officer five in the US Army. His previous assignments include working as a logistics staff and liaison officer with the Headquarters, US Department of the Army for Logistics (HQDA G4); chief, Supply and Mobility Division, White House Communication Agency (WHCA); capabilities developer at the US Combined Arms Support Command (USCASCOM) for the Global Combat Support System-Army (GCSS-Army); accountable officer and supply and services officer at the 4th Brigade Combat Team and 101st Combat Aviation Brigade, 101st Airborne Division (Air Assault), Fort Campbell, Kentucky, USA; and as a system administrator (Unix Solaris) at the 16th Corps Support Group, Hanau, Germany.

either the person's social security number (SSN) or national identification number, driver's license number, and credit and/or debit card number or their associated security codes.[4] The EU General Data Protection Regulation (GDPR)[5] suggests that EU citizens have the right to the protection of their personal data, noting nonetheless that the right is not absolute because personal data protection rights must be balanced against other fundamental rights and must be considered based on their relationship to the data's function in society.

Enterprise Data

Enterprise data comprise all data being accumulated at an enterprise, including the enterprise's financial and operational data, its intellectual property (i.e., patents, trademarks, copyrights, trade secrets), and the data used by the enterprise to transact with its internal and external users, suppliers, buyers and the like. Enterprise data can be described as the "central characteristics of the organizations, their internal structure and processes as well as their behavior as corporate actors in different social and economic contexts."[6]

Data Collection Dimensions

To determine data's rightful owner, an analysis of the ways in which data are collected, amassed, stored and transmitted was articulated and evaluated. The methodology of evaluating the data's criticality an enterprise transmits in a given transaction was not considered because such consideration is beyond the scope of this discussion. Therefore, the primary purpose was to evaluate the main activities from which data are generated, the ones that generate the data and data transmission.

Data can be categorized as master data, transactional data, reference data, metadata, historical data and temporal data. Regardless of the category, there are three states of data: data at rest, data in use, and data in motion or transit. An element of data from an individual or enterprise can be at rest in one system, in use or in process in another system, and in motion in yet another. Therefore, the argument is that if data can be in different states and in different systems simultaneously, then it is plausible for each entity in which the data reside, whether in-process in or in-route in, to claim ownership of that data element. For instance, a person's SSN can be at rest in the US Social Security Administration office or bank database, while the same SSN is being used at a US state Department of Motor Vehicle (DMV) to process a driver's license, and in motion while a mortgage company is sending the SSN to US credit reporting agencies for credit scores. At an instance, passport information belonging to any international traveler around the world could simultaneously be at rest at the issuing authority's database, in use by the traveler at an airport for identification and security checks, and in motion as the traveler's airline transmits its passenger manifest or checked-in bag information.

Online Activities

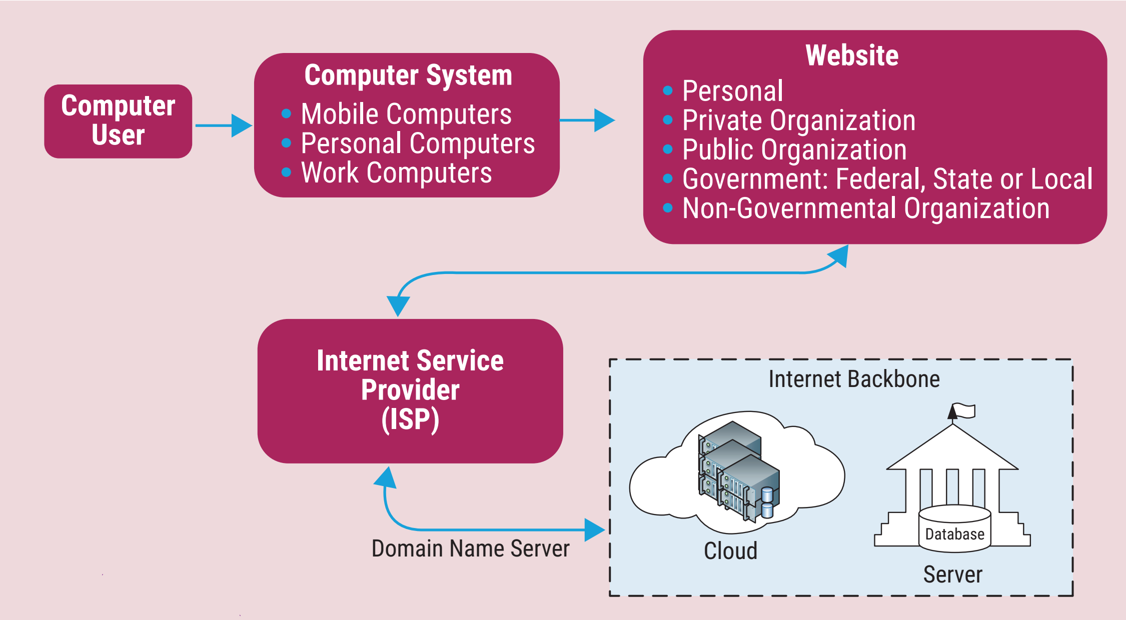
Through online activities in ecommerce, egovernment, emedical, emarketplace or the like, users pass data to people and across systems, organizations or entities. **Figure 1** shows an Internet connection for any electronic activity mentioned here. Computer users log on to their computers, click on a web browser to log on to Google, Amazon, Walmart, a healthcare system, a state/local government agency or a private individual's website to register or purchase goods or services. By doing so, users leave their pertinent data on the computer, website, Internet service provider (ISP) and the database servers involved. This means that for one transaction, one or more data elements from a single individual or enterprise user are made available to one or more systems or entities.

> "TYPICALLY, A SINGLE COMPUTER TRANSACTION IN A STORE INVOLVES THE PROVISION OF DATA TO THE MERCHANT, A BANK, A WEB SERVICE (CLOUD) HOST AND AN ISP."

Point of Sale

Terabytes and petabytes or more of data are constantly being collected from individuals and organizations around the world at points of sale (POS). To complete a transaction, the data are also

Figure 1—An Illustration of Users' Connections to the Internet



used, stored and transmitted; sold to third-party organizations; and/or transferred to trading partners at will, with little or no notification to the persons or entities from whom the data were generated.

Typically, a single computer transaction in a store involves the provision of data to the merchant, a bank, a web service (cloud) host and an ISP. Each person or organization involved in processing the transaction or purchase could claim the data it collected, and the original owner of the data could still lay claim. It should be noted that POS is usually associated with a brick-and-mortar store where customers go into the store, select item(s) they want or need, checkout from an onsite or cloud-based POS that is manually operated or automated, and exit the store.

Surveillance

Massive amounts of data are being captured through surveillance around the world, every second of every minute, every minute of every hour and every hour of every day. Private and public organizations, law enforcement, military, government agencies, and others use various surveillance tools to capture data from the citizenry, including the use of the stingray surveillance camera, traffic cameras, drones, Radio Frequency Identification (RFID) chips, body cameras, closed-circuit television (CCTV) cameras, Range-R radar devices, nanny cameras, facial recognition systems and others. The question here is, again, whether the data being collected belong to those from whom the data is collected or the people or organizations collecting the data. When properties are transferred, the transferrer physically loses the property, at the very least. Unlike properties, the transferrer of personal or enterprise data does not lose the data or information (i.e., the capturing of biometrics from individuals does not cause them to lose the physical characteristics of their eyes [retina or iris], facial bone, hand geometry, fingerprints, voice or value of the data).

Social and Media Activities

With the advent of social media platforms such as Facebook, YouTube, Snapchat, Instagram, Twitter, WhatsApp, Pinterest or LinkedIn, individuals and groups constantly leave behind pieces of their data. In 2018, 73 percent of adults in the United States used YouTube, and 68 percent used Facebook.[7] Instagram and Facebook[8] collect content; communication; product-related information users provide; and network, connection, usage pattern and transaction information. A review of Instagram's Terms of Use indicates that although it does not claim ownership of users' content, it requires users to grant it a "non-exclusive, royalty-free, transferable, sub-licensable, worldwide license to host, use, distribute, modify, run, copy, publicly perform or display, translate, and create derivative works of your content."[9] If this is not an ownership claim of content, then it may be difficult to ascertain what could be a claim of ownership.

Medical Facilities

Hospitals and clinics around the country and the globe constantly collect data from their patients and medical staff. The irony is that the patients rarely question why their data are being collected even when some of the data being requested could be deemed invalid or irrelevant to the medical service being provided. Personal data "is deemed irrelevant if the data is not necessary for a transaction."[10] Some data being collected from medical facilities are demographics, clinical, diagnosis, wellness, treatment, laboratory tests, patients' insurance, prescription, hospitalization and administrative information.[11] Moreover, the data could be aggregated for medical research, statistics and disease controls. Medical staff includes, but is not limited to, the primary care physicians, specialists, pharmacists, nurse practitioners, nurses, and administrative and support staffs. In addition, patients are tracked throughout their hospital stay, and their data are being collected during admission, diagnosis, prognosis, treatment and discharge. The question nonetheless is whether the hospitals and clinics own the data or whether the data belong to the patients and medical staff from whom the data is being collected.

Analysis

Designating the most accurate approach in measuring the real value of data is not a settled matter. Data valuation is still subjective. Microsoft acquired LinkedIn Corp. in 2016 for US$26.2 billion, at US$196 per share, which also amounted to paying approximately US$260 per monthly active user at the time.[12] Instagram attracted 30 million new users following Facebook's announcement of its acquisition in 2012 for US$1 billion.[13] The common thread between the two acquisitions is data. Therefore, the following is a closer examination of the value of data.

Data as an Asset and Commodity

Technological advancements in information systems have changed society's perspective and how data are valued. Users' information is being collected daily at every point of human activities, whether or not people are aware. Many billion-dollar enterprises such as Google and Facebook thrive on big data that they regard as an asset, because data have formed the core revenue generation for them and are deemed a cash cow. Big data "not only support(s) online transactions, it helps organizations in recording, relating, comparing, understanding, predicting, and prescribing their online users' behaviors based on all or a combination of the five business measurement variables, i.e., cost, time, quantity, quality, and human reaction."[14] Today, many organizations trade on personal and enterprise data for their business interests, production, sales, forecasting, credit reporting, election activities, business intelligence, decision support systems, predictive and prescriptive analyses, etc. Therefore, the value of data cannot be overstated.

> "TECHNOLOGICAL ADVANCEMENTS IN INFORMATION SYSTEMS HAVE CHANGED SOCIETY'S PERSPECTIVE AND HOW DATA ARE VALUED."

Data Value

One of the most important characteristics of data is that the sharing of their ownership by one or more people, systems or entities does not necessarily diminish their value. Thus, if persons or entities A, B, C, D and E are involved in a transaction to obtain an information security book online, where A is the user who purchased the book, B is the information system used to make the purchase, C is the online merchant's website, D is the Internet service provider (ISP) that provided the throughput for the transaction, E is the organization that hosts the cloud service for the merchant. Anecdotal and empirical evidence have shown that, in this situation, entities A-E would retain some data from A that was collected willingly or unwillingly (nonconsensual technology-assisted collection and obligatory passage point-induced or cohesive disclosure).[15] While there is nothing in the extant literature that indicates that the value of the data is diminished because of the five touch points listed, the issue of rightful ownership is still not settled, especially in the absence of a clear data ownership regulation.[16] While B-E could profit from such data, the value of A's data has not shown to have diminished.

Discernment of Data Ownership

Data ownership constitutes having authority or rights and responsibility or control over data. Thus, data are said to have a single owner when one person, individual, entity or organization has the sole right and control over the data. Similarly, data are said to have multiple owners when either people, groups, individuals, enterprises or entities have rights and control over them.

Understanding data ownership and/or knowing the originator of the data in the data cube, warehouse or lake is crucial in understanding the data's authenticity and quality. In addition, understanding data ownership helps navigate data ownership rules, authority, control, format, generation, processing and assigning credit or blame for the data and their inventory.[17] Imagine a web portal where users can enter information about themselves directly (single ownership) and another that gathers the same users' information from their social media page (multiple ownership). Data users will mostly receive unadulterated information from the single data owners, except for false data, which users may enter intentionally or otherwise. However, the users' data received from social media sources is subject to the same possible false information from users and the ownership rules and assumptions the social media entities use in gathering information. Therefore, the importance of knowing whether data have a single or multiple owners or simply knowing who owns the data cannot be overemphasized.

However, it is worth noting that the nature and approach of assigning data ownership in organizations is different from the tenets of the argument herein. In the enterprise world, data governance assigns data ownership to either a person, business resource, technical resource or a federated responsibility.[18] This discussion aims to establish that data have multiple owners because each system or entity through which data are propagated or to whom data are syndicated could claim ownership of the data, including the data source itself.

Indicators of Data Ownership

There are several indicators or precursors and paradigms to identify whether data have one or more owners. The indicators are listed as questions to help navigate through the complexities of evaluating who owns data. Hence, answering one or more of these questions helps determine whether data are from a single owner or multiple owners.

- Are the data coming from a primary or a secondary source? While primary data may come from a single owner, secondary data might not. Primary data[19] are original data and factual in nature. They constitute data that come directly from a source, e.g., the information submitted by an individual or enterprise, either on a form or online, to receive services or to purchase products. On the other hand, secondary data are data that were previously collected by someone else[20] or by another entity.
- Are the data novel? New data may come from a single owner, while existing data may have undergone transmission, transformation and may reside across multiple databases.
- Are the data an intellectual property (IP)? IP is the "rights given to persons over the creations of the minds"[21] or the "assignment of property rights through patents, copyrights and trademarks. These property rights allow the holder to exercise a monopoly on the use of the item for a specified period."[22] The notion here is that IP data belong to either a person(s), a group(s), an enterprise(s), or a group of people, enterprises or a combination thereof. Therefore, understanding the ownership of the IP helps in determining the singularity or plurality of data ownership.
- Have the data been processed or transmitted before? If the data have been processed or transmitted before, the likelihood that more than one person, group or entity may claim ownership to the data is high; as such, the data could reasonably be said to have multiple owners.
- Can the data be found in one or more databases? Considering that data ownership is based on having authority and control of the data, it is reasonable to say that the person, group or enterprise in whose system the data reside could claim the data ownership. Thus, a reasonable

person can infer that the data's existence in multiple databases means that the data have multiple owners.

In addition, **figure 2** shows an identification, assessment and assignment of data ownership rights based on the analysis of the 11 data ownership paradigms adapted from *Enterprise Knowledge Management*.[23]

Analysis of the Ownership Paradigms

The following is the assessment and discussion on the *Enterprise Knowledge Management* data ownership paradigms, which informed, in part, the assignment of data ownership rights listed in **figure 2**:

1. **The creator of the data is the owner of the data**—This represents the investment of the creator in the creation or generation of the data. The creator or generator of the data can be an individual, group or organization. If an organization, the data could be generated organically from its various departments, divisions, branches or sections—indicative of single ownership.
2. **The consumer of the data is the owner of the data**—This refers to broader ownership of data by multiple people, groups or entities, i.e., financial and sales enterprises such as Merrill Lynch, Fidelity, Vanguard, Walmart, Lowes, Home Depot and others that pull stocks, sales, distribution, and customer and vendor master data from various sources and aggregate the data for metrics, predictive and/or prescriptive analysis, especially if an external source exists among the data sources—indicative of multiple data ownership.
3. **The compiler of the data is the owner of the data**—This refers to entities such as Google, Facebook, LinkedIn, Yahoo and others that gather data from various sources, compile them, and make them available online via search engines or otherwise to their users—indicative of multiple data ownership.

> “IT IS COMMON FOR LARGE ORGANIZATIONS TO CLAIM OWNERSHIP OF DATA THEY HAVE GENERATED AND ACCUMULATED OVER TIME.”

4. **Enterprise as the owner of the data it generates and accumulates**—It is common for large organizations to claim ownership of data they have generated and accumulated over time through their business operations, including data from their trading partners—indicative of multiple data ownership.
5. **The funding organization of the data as the owner of the data**—This refers to situations in which the person, group or entity that

Figure 2—Data Ownership Rights

	Data Ownership Paradigms	Single Ownership	Multiple Ownership
1.	The creator of the data is the owner of the data.	X	
2.	The consumer of the data is the owner of the data.		X
3.	The compiler of the data is the owner of the data.		X
4.	Enterprise as the owner of the data it generates and accumulates		X
5.	The funding organization of the data as the owner of the data	X	
6.	The decoder of the data as the owner of the data	X	
7.	The packager of the data as the owner of the data		X
8.	The reader of the data as the owner of the data		X
9.	The data subject of the data as the owner of the data		X
10.	The purchaser/licenser of the data as the owner of the data		X
11.	Everyone is a data owner.		X

commissioned the data creation claims data ownership, e.g., paying a person or other enterprise to write a business intelligence report or research grant. In the case of technology change implementation, the project/program sponsor would be the rightful owner of the data, unless stated otherwise in the contract—indicative of single data ownership.

6. **The decoder of the data as the owner of the data**—When the data are locked or encoded, the person, group or entity that decodes the data is the owner of the data, because there is a cost associated with the data decoding, which is an investment in the value of the data, e.g., decoding genetics information or discovery of pharmaceutical products—indicative of single data ownership.

7. **The packager of the data as the owner of the data**—The focus is on the person, group or entity that formats data for a particular use because of the value the formatting engenders, i.e., formatting an author(s) manuscript by a publisher or packaging an individual, group, or entity's product or services by a marketer for a particular market. The data could have multiple ownership because of the value the packager added to the publishing, products or services—indicative of multiple data ownership.

8. **The reader of the data as the owner of the data**—This "implies that the value of any data that can be read is subsumed by the readers, and therefore the reader gains value through adding that information to an information repository."[24] It is possible for readers to absorb the data they have read and use the data to form the basis for their future work, i.e., research using the principles inherent in a given theory to propose a solution to a research problem. Here, the author and the reader may claim the data—indicative of multiple data ownership.

9. **The data subject of the data as the owner of the data**—A third party claims the data collected from a data subject, e.g., in a hospital, patients may claim their protected health information (PHI), yet the medical staff (doctors, nurses and others) tied to patients' diagnoses and prognoses and pharmacy may claim patients' prescription/refill information to ensure the patients' compliance with the prescriptions, and the drug manufacturer may identify the doctors who prescribe its medication. Hence, although the data are generated from a patient, the hospital may have a claim to some of the data. In the case of pure or applied research, the researcher or practitioner may lay claim to the data generated from a sample subject—indicative of multiple data ownership.

10. **The purchaser/licenser of the data as the owner of the data**—A buyer or licenser who purchases or licenses data, respectively, could claim data ownership because of the investment (e.g., customer or vendor master data, transactional data, purchasing mailing lists, email addresses) and the expectation of a return in investment (ROI) therein—indicative of multiple data ownership.

11. **Everyone is a data owner**—This refers to the concept of global data ownership, where individuals, groups or entities may have limited to no data ownership, i.e., open source, scientific data, or published works or results—indicative of multiple data ownership.

Conclusion

The idea that data mostly have multiple owners is supported based on the assessment and review of the extant literature and looking at practical events around the dimensions of data collection and ownership paradigms. The fact that an element of data can be at rest, in use and in motion simultaneously in separate systems is indicative. If multiple persons and/or systems could simultaneously claim ownership of an element of data, implicitly or explicitly, then it is reasonable to argue and conclude that data mostly have multiple owners.

Endnotes

1 Vermaat, M. E.; S. L. Sebok; S. M. Freund; J. T. Campbell; M. Frydenber; A. Ly; *Discovering Computers 2016: Digital Technology, Data, and Devices*, Cengage Learning, USA, 2016

2 Vogt, W. P.; R. B. Johnson; *The SAGE Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*, Sage Publication, USA, 2016

3 DAMA International, *DAMA–DMBOK Data Management Body of Knowledge, 2nd Edition*, Technics Publications, USA, 2017

[4] Tennessee Code 47-18-2107, “Breaches of Security Systems; Definitions; Notice,” 2018, *https://www.wyattfirm.com/uploads/1566/doc/47-18-2107_Breaches_of_security_systems_definitions_notice.pdf*

[5] Intersoft Consulting, General Data Protection Regulation GDPR, 2018, *https://gdpr-info.eu/*

[6] Liebig, S.; “Organizational Data,” Working Paper Series of the German Council for Social and Economic Data 67, German Council for Social and Economic Data (RatSWD), 2009, *https://ideas.repec.org/p/rsw/rswwps/rswwps67.html*

[7] Smith, A.; M. Anderson; “Social Media Use in 2018,” Pew Research Center, 1 March 2018, *https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/*

[8] Facebook, Data Policy, 19 April 2018, *https://www.facebook.com/full_data_use_policy*

[9] Instagram, Terms of Use, 19 April 2018, *https://help.instagram.com/581066165581870/?helpref=hc_fnav&bc[0]=Instagram%20Help&bc[1]=Privacy%20and%20Safety%20Center*

[10] Offor, P. I.; “Why We Disclose Personal Information Despite Cybersecurity Risks and Vulnerabilities: Obligatory Passage Point Perspective,” *International Journal of Smart Education and Urban Society (IJSEUS)*, vol. 9, iss. 4, 2018, *https://www.igi-global.com/article/why-we-disclose-personal-information-despite-cybersecurity-risks-and-vulnerabilities/214053*

[11] University of Washington Library, “Data Resources in the Health Sciences,” USA, 2019, *https://guides.lib.uw.edu/hsl/data/findclin#s-lg-box-1908467*

[12] Short, J. E.; S. Todd; “What’s Your Data Worth?” *MIT Sloan Management Review*, 3 March 2017, *https://sloanreview.mit.edu/article/whats-your-data-worth/*

[13] Bosker, B.; “Instagram Acquired by Facebook for $1 Billion,” *HuffPost*, 9 June 2012, *https://www.huffpost.com/entry/instagram-facebook-acquisition_n_1412623*

[14] *Op cit* Offor, p. 41

[15] *Op cit* Offor, p. 39–40

[16] Kostkova P.; H. Brewer; S. de Lusignan; E. Fottrell; B. Goldacre; G. Hart; P. Koczan; P. Knight; C. Marsolier; R. A. McKendry; E. Ross; A. Sasse; R. Sullivan; S. Chaytor; O. Stevenson; R. Velho; J. Tooke; “Who Owns the Data? Open Data for Healthcare,” *Frontiers in Public Health*, 17 February 2016, *https://www.frontiersin.org/articles/10.3389/fpubh.2016.00007/full*

[17] McDonald, D. D.; “Why ‘Data Ownership’ Matters,” CTO Vision, 16 February 2017, *https://ctovision.com/data-ownership-matters/*

[18] Thomas, G.; “Assigning Data Ownership,” The Data Governance Institute, 28 September 2013, *www.datagovernance.com/assigning-data-ownership/*

[19] Surbhi, S.; “Difference Between Primary and Secondary Data,” Key Differences, 26 August 2017, *https://keydifferences.com/difference-between-primary-and-secondary-data.html#KeyDifferences*

[20] *Ibid*.

[21] Yang, D.; *Understanding and Profiting From Intellectual Property: Strategies Across Borders*, Palgrave Macmillan, USA, 2013

[22] Organisation for Economic Co-Operation and Development, Glossary of Industrial Organisation Economics and Competition Law, France, 1990, *www.oecd.org/regreform/sectors/2376087.pdf*

[23] Loshin, D.; *Enterprise Knowledge Management: The Data Quality Approach*, Academic Press, USA, 2001

[24] *Ibid*.