# The Future of Data Centers



A Survey Design

By:

Glen Cooper 214-914-6756

For:

EPPS 6302

Fall 2021

Dr. Ho

#### Introduction

Mass storage of information began as early as 283 BCE with the founding of the Library of Alexandria. Today great collections are stored in huge buildings such as the Library of Congress, but the core of today's data storage resides in buildings dispersed around the world known as data centers (Berson, 2015). Data centers have their beginnings in the huge computer rooms of the 1940s, such as the room that housed the ENIAC, the world's first large computer. Data centers represent the apogee of a process that began with the first industrial revolution and Charles Babbage's idea of an analytical engine (Bowker, 2019). The first industrial revolution started around 1800. During this revolution, for the first time, goods and services were produced by machines. "Hard" capital in the form of railways, heavy industry, and the steam engine were essential ingredients of this phase. The second revolution began at the end of the 19th century with electrification and the "soft" process innovation (i.e., the assembly line) (Lischka, 2011). The third industrial revolution began in the 1970s and was distinguished by further automation through electronics (Lischka, 2011). Much of the early advancements during this period came through the installation of mainframe computers. These were large complex computers costing millions of dollars. They were not only expensive in their own right, but required institutions to invest in expensive environmental controls, staffing to operate them, costly vendor maintenance agreements, upgrades to their telecom bandwidth, and time-consuming periodic upgrades (Bowman, 2012). Later this period saw the development of the personal computer. During this later period, especially during the 1980s, users started deploying Personal Computers (PC) everywhere and in many cases without regard to operating requirements or connectivity. However, as technology operations grow in complexity, organizations became aware of the need to control and integrate these resources (Tatnall, A. (Ed.), 2012). PCs began to be connected to

more powerful computers known as servers. Servers were then connected to other servers. Servers began to be located in central locations known as server rooms. Servers were connected to other external servers resulting in the birth of the internet and allowing for global access to information (Lischka, 2011). Today by subscribing to your local internet, a person can instantaneously plug into information located in thousands of computers located throughout the globe. Servers began to replace mainframes. Groups of servers were then accumulated together in large buildings. This centralization gave birth to the data center.

Today's modern data centers range from 1,000 square feet to over 1 million square feet (Taneja, 2020). It is difficult to determine a precise date in history when the first data center was created (Tatnall, A. (Ed.), 2012). However, the concept of a shared facility that delivered computer services was originally developed by the Japanese. The benefits of shared data center facilities were lower barriers to entry, faster to market delivery, lower maintenance costs, and concentration of expertise (Bowman, 2012). A data center is a physical facility housing computer systems and associated components. Data centers comprise the computer systems and staff and include power connections, backup power supplies, communication equipment, redundant communication systems, air conditioning, fire suppression equipment, and physical security (Tatnall, A. (Ed.), 2012).

Currently, there are three favors of data centers: stand-alone data center, shared infrastructure, and collocation-caged environments (Bowman, 2012). The Stand-alone data centers are built by single owners to house an enterprise's computer needs. Shared infrastructure or multi-tenanted centers are single-purpose buildings that offer enhancements such as emergency power service generators, cooling systems, security, monitoring, and maintenance. Tenants take largely unimproved space from the service provider. The tenants install their computer equipment within

the rented facility. Finally, collocation-caged environments are multi-tenant facilities in which there are shared improvements and servers. The owners of the centers sell services by the server cabinet (Bowman, 2012).

According to Reportlinker. (2020). the data center market is expected to grow at a compound annual growth rate (CAGR) of over 1% from 2019 to 2025. Understanding that growth will be the topic area of this study.

#### **Research Statements**

## Statement of the Research Problem

Data centers have become an integral part of modern life. Unlike a public library, which represents the public's understanding of a centralized physical store of knowledge, data centers are an agora where most of today's information is stored. Invisible to the public but critical to first-world modern life. While pundits and researchers attempt to predict the future direction of this vital infrastructure, the opinion of actual participates in the industry's development is less understood.

#### Statement of the Research Purpose

The purpose of this study will be to uncover opinions on the future direction of data center development from those participating in the industry's management and creation. A survey methodology will be followed in order to capture and evaluate that knowledge.

## Statement of the Research Question

Two basic research questions will be analyzed in this project:

- 1. What factors do data center industry participants believe are the most significant in expanding data center services?
- 2. What limitations do data center professionals believe are the most significant in restricting the expansion of data center services?

## **Policy Relevance**

There are several potential policy-related outcomes from this research. Likely policies will relate to either for-profit organizations, government, or education. From a for-profit organizational perspective, the results of this survey can have directional consequences. These consequences can be strategic by either guiding an organization in a new direction or reinforcing the directional decisions already made. From a governmental perspective, privacy and environmental laws might be reviewed depending upon survey results. Artificial Intelligence policies may need to be developed. Governmental decisions regarding infrastructure development, especially telecommunications infrastructure, might be evaluated and revised. Additionally, universities might gain knowledge as to possible curriculum development to meet data center industry requirements. All suggestions should ultimately be assessed through a framework that considers the goal, problem, competing solutions, implementation, enforcement, and monitoring aspects of policy.

#### Literature Review

A survey of recent literature identified two data center surveys like the one proposed in this study. Uptime Institute conducted the most recent survey. Uptime Institute is a data center standard-setting organization. This survey was designed to "reveal what [data center] operators around the world are thinking, doing and planning in the areas of efficiency, resiliency, workload placement, staffing and new technology adoption." (Ascierto & Lawrence, 2020). Uptime's survey is comprehensive and well-funded. It is international in scope covering organizational details (e.g., what percentage of your organization's total IT run in different types of environments, current power usage, outage levels, etc.), operational issues (e.g., what factors drive decisions within the organization, causes of outages, etc.), and industry issues (e.g., how difficult is it to hire needed staff, etc.). However, the survey is more focused on recent history than opinions about the future direction of data centers (Ascierto & Lawrence, 2020). The second survey was conducted by Gartner, Inc. Gartner is a well-respected research organization. The specific questions asked in the Gartner survey were not available; however, the conclusion drawn by Gartner included that hybrid IT (i.e., using multiple infrastructure platforms) are becoming more common, that enterprise data centers are being simplified, and that technology is allowing data centers to increase resiliency (Gartner Reprint, n.d.). From these conclusions, again, this survey appeared more focused on the current status of data centers rather than an attempt to seek industry opinions about the future directions of the industry. Therefore, the literature of data center surveys appears to have a gap in asking data center professionals directly their opinion on the direction of the industry. This research project will attempt to fill that gap. In developing this research project's survey, a review of the current literature was conducted to evaluate the more important topic areas within the industry. One of the primary concerns within

the industry is the level of availability or uptime a given data center can maintain without interruption (Uptime Institute, 2018; Kleyman; 2012, Ajmera, n.d.). As a corollary to uptime, uninterruptible power supply, or the ability to switch to alternative energy sources if the primary utility service becomes unavailable was of interest (Uptime Institute, 2018; Ascierto & Lawrence, 2020). Of course, the power availability from utility services is also of concern (Uptime Institute, 2018; Ascierto & Lawrence, 2020; Gartner Reprint, n.d.). Operating data centers require a significant amount of power; therefore, energy efficiencies within those operations is a focus (Reddy, Setz, Gangadharan & Aiello, 2017; Covas, Silva, & Dias, 2013; Ascierto & Lawrence, 2020; Gartner Reprint, n.d.; Ajmera, n.d.). These large power draws produce a considerable level of heat which must be removed from operations meaning that cooling technologies are of great importance (Reddy, Setz, Gangadharan & Aiello, 2017; Covas, Silva, & Dias, 2013; ASHRAE, 2008; Kleyman, 2012). Other operational concerns include data load capacity and growth management, operations monitoring, management, and security (Reddy, Setz, Gangadharan & Aiello, 2017; Digital Realty, 2020; Technavio, 2020; Ascierto & Lawrence, 2020; Kleyman, 2012; Gartner Reprint, n.d.; Cole, 2012; Ajmera, n.d.). External factors are also of industry concern. Items such as space availability, labor supply, laws, and new technologies are all being monitored by the industry (Digital Realty, 2020; Technavio, 2020; Gartner Reprint, n.d.; Ascierto & Lawrence, 2020; Reddy, Setz, Gangadharan & Aiello, 2017; Giampietro & Mayumi, 2018). Finally, competitive opportunities are important (Digital Realty, 2020; Uptime Institute, 2018; Ajmera, n.d.) All these issues will be included within this research project's investigation.

#### **Research Design**

## **Theory**

Understanding an industry is susceptible to both quantitative and qualitative analyses.

Quantitative methodologies are successful when sufficient historical data is available. However, projecting the future course of an industry using historical data has limitations because such data mostly explains what has occurred rather than what will happen. Generally, while constrained by physical, technological, and social conditions, the future of an industry is more a function of human perceptions and decision-making. Understanding human perceptions and decision-making are fertile ground for qualitative research. As such, to address the research question defined above, this research project will incorporate a qualitative approach.

"Despite the interdisciplinary recognition of the value of 'qualitative research' ... qualitative research is not a unified field of theory and practice" (Guest, Namey, & Mitchell, 2012). There may be no one correct way to conduct qualitative research, but Guest, et al., suggest that when selecting a method, the researcher should first inquire how well the topic of interest is established. If it is not well established, they suggest generalized inquiries into available literature, secondary data, or discussions with experts/stakeholders. In the current case, the data center industry is well established, so these approaches would likely yield little new information. Second, Guest, et al. suggest that if the topic is established but the boundaries of the subject are not well established, the researcher should engage in a literature review followed by participator observation. In this current study, the topic boundaries are also well established. Finally, Guest, et al. suggest that if both the topic and boundaries are established, the researcher should employ

focus groups, semi-structured interviews, or surveys. As this researcher has a limited financial budget and wishes to use more fixed-response categories, a survey methodology appears optimal.

## Study Population and Sample Frame

"Surveys ... are never interested in the characteristics of individual respondents per se. They are interested in statistics that combine those answers to summarize the characteristics of groups of persons. Sample surveys combine the answers of individual respondents in statistical computing steps ... to construct statistics describing all persons in the sample." (Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009). This survey will seek to draw an inference as to the opinions of the data center industry population. Thus, the target populations for this research are persons employed within the data center industry. The base unit for this study will be individual persons. The sampling frame will be industry persons that have membership in data center industry groups. This researcher has identified five potential industry groups to participate in this research (see Appendix 1- Data Center Trade Organizations). The members of these combined groups, or a subset thereof, form the base of the sampling frame of this study. The element of the population to be sampled will be the list of email addresses of the base unit. This strategy is open to the risk of under-coverage of the entire group of persons within the data center industry because the frame only includes those motivated to be in industry groups. In addition, some email addresses may be out of date. These two factors may result in nonobservation error. It is also possible that there are foreign (i.e., ineligible) units within the population. Persons who are not members of the data center industry yet chose to participate in industry groups. There is also the risk of duplication. Duplication may occur because a person

may have their email entered twice in the same organization. The same person may have different email listings within the same industry group, or a single person may participate in separate industry groups. If a participant is listed twice, there is a risk of causing those persons to be overrepresented. Clustering may also be a problem when a given email is provided that is not matched with a single individual but with a group of individuals, for example, a department or entire company. Potential bias in the sample frame is also possible as senior executives may be less willing to supply business emails than those involved in consulting to the data center industry.

Including all these possible errors, it is still reasonable to assume that an industry group should represent a cross-section of the data center industry population even though a well-defined sampling frame coverage error using email listings may not be completely unknowable under this sampling approach. However, an attempt will be made to calculate coverage error a best as possible by calculating the actual frame covered and not covered using estimates of the total data center industry workforce and the sample surveyed. Additionally, efforts will be made to understand the level of industry trade groups' representation within the industry.

Post-survey adjustment analysis will be considered. For example, decisions must be made about including a data record with less than complete information, also known as "item missing data." Nonresponse patterns over different subgroups will be analyzed. Weighting up underrepresentation will be considered. Potential weightings might include an adjustment for company size, employment position, geographical location, etc. Nonresponse error (nonresponse rate) will be evaluated for its potential impact on the survey results

#### Methodological Approach

As mentioned above, a survey methodology will be utilized for this research project. The survey will be computer-based, email broadcasted, website presented, and, in general, close-end questioning. The most commonly cited advantages of computer-based email surveys are that they provide "quick access to information, wide geographical scope, a larger sample size, ... reduced cost, ... no need for face-to-face interaction, [and] no paper wasted" (Fricker & Schonlau, 2002). The commonly cited disadvantages of computer-based email surveys are the limited access to or high costs of some databases, duplication of email accounts, the potential for the collection of irrelevant data, low motivation of the targeted respondents, and need for precise formulation of the questions to avoid unreliable answers (Fricker & Schonlau, 2002). The limited access and high cost of databases can be overcome by obtaining data from several different resources. A well-crafted request for participation from those organizations should also help increase the availability of data. In addition, the potential for duplicate email accounts can be partially reduced through deduplication processes. Finally, the collection of irrelevant data can be alleviated by carefully targeting respondents through data source selection.

The low motivation of the targeted respondents presents a unique challenge. Well-crafted communications that appeal to a communal spirit, namely, advancement of industry knowledge and other assurances, can help motivate respondents. For example, Almeda, 2009 suggests that the survey introduction should include: identification of the persons or organization conducting the study, explanation of the importance of the research and reason/s why it is being undertaken, the reason why the respondent should be the one to answer the questionnaire, assurance that there are no right or wrong answers, and guarantee the confidentiality of responses.

Additionally, Almeda, 2009 suggests that surveys begin with neutral questions to draw the

respondent into the questionnaire and encourage completion. Almeda, 2009 also recommends that question length and type vary to keep respondents interested and that difficult questions be placed deeper within the questionnaire to avoid having the respondent quit early out of frustration. Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009 suggest that sensitive questions, which may be seen as intrusive or embarrassing by the respondent, be placed at the end of the survey. Additionally, making multiple distributions to initial nonrespondents has been shown to be effective in reducing nonresponses (Goyder, 1985; Heberlein and Baumgartner, 1978).

The need for precise question formulation mainly arises because the respondent has limited options for response. If the correction option is not offered or the participant misinterprets the question, they will be unable to answer as the researcher intended. Thus, pretesting of the survey can help reduce this risk (Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009). In addition, guidelines detailed below have also been shown to reduce this risk (Almeda, 2009):

- Questionnaire appearance should be simple and well-spaced.
- Type font and size should be readable.
- Questions should be ordered in such a manner as to provide a logical sequence.
- Question flow should be respondent-friendly.
- Response categories should be mutually exclusive.
- Technical terms should be kept to a minimum and defined as required.
- Response categories should include the "Other" option with sufficient space for writing answers.
- Verbally rating scales used an even number of response categories so that no neutral answers were permitted.

Finally, Tourangeau, Rips, & Rasinski, 2000 recommend that all questions be carefully screened for grammatical ambiguity, excessive complexity, faulty presupposition, vague concepts, vague quantifiers, unfamiliar terms, or false inferences.

Using the above guidelines, this research project will request participation from several data center industry groups (see Appendix 1 - Data Center Trade Organizations) using a well-crafted email solicitation (see Appendix 2 - Data Center Trade Organizations Participation Email Draft). Once data is received, it will be examined for duplicate email addresses. Respondent participation will be encouraged through a well-created solicitation email and survey introduction (see Appendix 3 – Data Survey, Introduction). The overall survey will begin with neutral demographic questions (see Appendix 3 – Data Survey, Individual Organization Demographics) before moving into more in-depth questions. Questions will vary in form (e.g., multiple-choice, fill in the blank, and Likert scale). Sensitive questions will be left to the end of the survey (see Appendix 3 – Data Survey, Personal Employment Demographics, and Personal Demographics). All questions will be carefully reviewed to ensure they satisfy the aboveoutlined design suggestions. In addition to those guidelines, this research will keep survey instructions to a minimum and allow the questionnaire flow to do the navigational work. The survey will use repeat wording and timeframes to avoid confusion or misinterpretations. Timeframe responses will be designed around normal business planning ranges (e.g., 1-5 years, 5-10 years, etc.). Appropriate definitions will be supplied both at the outset of the questionnaire and within questions. Questioning will also be designed to offer clues to encourage the participants to think about high-level strategic industry issues.

Before final distribution, a pretest survey distribution will be conducted. This pretest will help assure that the distribution methodology is well understood and operationally functional. In

addition, a survey of the pretest survey participates will be conducted to gain feedback on the survey design. Other approaches to validate the survey questions, such as expert review, face-to-face administration, and focus group discussions, may be considered.

Finally, during the actual distribution of the survey, unresponsive recipients will have follow-up second request emails sent.

# Method of Data Collection

This research project will be conducted using a web-based computer-administered email survey process utilizing the Qualtrics XM (<a href="www.qualtrics.com">www.qualtrics.com</a>) online survey system. This system distributes survey questionnaires to participants based on email addresses. There is no interviewer involvement in the data collection process. All survey responses are collected and returned electronically by the Qualtrics XM system to the researcher. This system represents a low-cost, efficient alternative to either a direct mail survey or face-to-face interviewing.

Depending on the adequacy of the responses and overall data set, the survey may also be distributed through self-selected web participation methods offered by social media sources such as Twitter, Facebook, LinkedIn, etc. Also, certain industry discussion groups or voluntary participation websites might be used to survey distribution.

#### Role of the Researcher

The role of the researcher is central to any qualitative study. In this study, the researcher will attempt to maintain an epistemologically neutral position and lessen the distance between the researcher and the researched to present an objective as possible analysis.

#### **Ethical Issues**

Maintaining appropriate levels of confidentiality is likely the most significant ethical issue to be encountered in this survey. Fortunately, Qualtrics omits Response ID for privacy purposes. Omitting Response IDs limits the researcher's ability to link specific responses to individual email addresses and names. That, in turn, reduces the possibility that privacy will be accidentally violated. In general, the topic area is non-sensitive. However, some questions relating to personal demographics (e.g., income) may be considered sensitive. Appropriate record-keeping procedures will need to be designed to comply with ethical and legal requirements. None of this research involves any deception, so that should not be a concern. It is also possible that some survey recipients might be international. In that case, international guidelines and laws for research on human subjects will have to be reviewed and addressed or those subjects removed. Finally, the survey will be submitted to the University of Texas at Dallas' Institutional Review Board (IRB) for review and approval.

## **Findings**

## Result of Survey Pretest and Survey of Survey

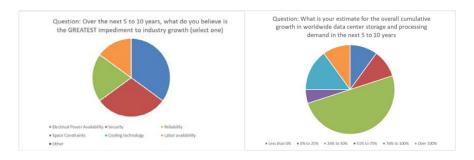
## Survey Pretest

A pretest of the data survey was conducted between October 30 and November 5, 2021, using the survey located in Appendix 3. The cover email for this distribution is found in Appendix 4. The recipients of this pretest survey were judgmentally selected. The objective of this pretest was to evaluate the mechanics of the distribution process and provide data to begin designing survey analyses. Below is a reconciliation of that distribution is below:

Reconciliation to Final Responses per Qualtrics	
Total Email Distribution	23
Issues with Email addresses	(3)
Subtotal	20
Correcting for email address issues distributions	3
No responses	(2)
Final Total Responses	21

In general, the mechanics of the Qualtrics distribution process was straightforward. There were a few lessons learned from this pretest distribution. First, Qualtrics produces sample distributions that test the survey before it is distributed. These sample distributions should be deleted before general distribution not to have them blended with the overall distribution results. Second, it is advisable to make sure your question numbering is sequential before distribution which facilitates later reconciliations and analyses.

The pretest survey allowed the researcher some data to begin to formulate a final distribution analysis. Here are a couple of draft analyses:



## Survey of Survey

A survey of survey questionnaire was conducted between November 2 and November 5, 2021, using the survey located in Appendix 6. The cover email for this distribution is found in Appendix 5. The survey's object was to address opinions about the pretest survey length, question flow, and content. A total of 17 survey of survey questionnaires were distributed. The total distribution was four fewer than the pretest distribution receipts because some pretest responses were received after the cutoff date for the survey of survey distribution. All survey of survey questionnaires were returned. Below are the results of this second survey:

Length of survey	of survey Feeling about survey length			gth		
_	Count	%		Count	%	
1-5 minutes	0	0%	Too Short	0	0%	
6-10 minutes	10	59%	Too Long	6	35%	
11-15 minutes	4	24%	About Right	11	65%	
more than 15 minute	3	18%		17	100%	
Total Respondents	17	100%				
Feeling about survey	flow		Degree of Gra	Degree of Grammatical Errors		
_	Count	%		Count	%	
Positive	15	88%	Low	16	94%	
Negative	2	12%	High	1	6%	
	17	100%		17	100%	
Readablity			Comfort with (	Comfort with Questions		
_	Count	%		Count	%	
Positive	16	94%	Positive	15	88%	
Negative	1	6%	Negative	2	12%	
_	17	100%		17	100%	

As is demonstrated, the results of this second survey were generally positive.

## **Future Analysis**

In addition to the analysis drafted above, Qualtrics also provides a measure of respondents' survey duration. This statistic could be useful in the propriety of the responses. Another measure of completeness might be the degree to which open-ended questions are responded to and the level of words used in longer response opportunities.

# **Interpretation of Findings**

Pending outcome of the research.

## **Progress Report**

Many features of this research project remain to be completed. These include:

- Obtaining all necessary university faculty and Institutional Review Board (IRB) approvals.
- Distribution of data request email to industry groups
- Cleaning of email lists obtain
- Planning and conducting the actual survey data
- Response analyses
- Draft final report

#### **Bibliography / Citations**

## References

- Ajmera, H. (n.d.). A Complete Guide to Cloud Computing (p. 11). www.harshajmera.com
- Almeda, Josefina V., (2009, August), Introduction to Questionnaire Design, University of the Philippines.
- Ascierto, R., & Lawrence, A. (2020). Uptime Institute Global Data Center Survey 2020 (UII-38 v1.1; p. 32).
- ASHRAE, (2008). High density data centers: Case studies and best practices. (p. 193). American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.
- Berson, Josh. (2015). Computable Bodies: Instrumented Life and the Human Somatic Niche.
- Bowker, G. C. (2019). The Time of Computers: From Babbage and the 1830s to the Present. In W. Aspray (Ed.), Historical Studies in Computing, Information, and Society: Insights from the Flatiron Lectures (pp. 1–15). Springer International Publishing. <a href="https://doi.org/10.1007/978-3-030-18955-6\_1">https://doi.org/10.1007/978-3-030-18955-6\_1</a>
- Bowman, Ronald H., (2012), How We Got Here: History of Data Centers and Current Choices. In Business Continuity Planning for Data Centers and Systems (pp. 1–8). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470428405.ch1
- Cole, D. (2012). Data Center Infrastructure Management. 19.
- Covas, M. T., Silva, C. A., & Dias, L. C. (2013). Multicriteria decision analysis for sustainable data centers location. International Transactions in Operational Research, 20(3), 269–299. https://doi.org/10.1111/j.1475-3995.2012.00874.x
- Digital Realty. (2020). Digital Realty, Data Gravity Index (p. 41). Digital Reality. https://go2.digitalrealty.com/rs/087-YZJ-646/images/Report Digital Realty 2009Data Gravity Index Report.pdf
- Fricker, R.D., & Schonlau, M. (2002). Advantages and Disadvantages of Internet Research Surveys: Evidence from the Literature. Field Methods, 14, 347 367.
- Gartner Reprint. (n.d.), The Future of Enterprise Data Centers—What's Next, Retrieved July 26, 2021, from https://www.gartner.com/doc/reprints?id=1-25O1471R&ct=210329&st=sb&mkt\_tok=MDg3LVlaSi02NDYAAAF-QDqsR36QH19CFU6VluDgTj5NcAB8F%E2%80%A6
- Giampietro, M., & Mayumi, K. (2018). Unraveling the Complexity of the Jevons Paradox: The Link Between Innovation, Efficiency, and Sustainability. Frontiers in Energy Research, 6, 26. https://doi.org/10.3389/fenrg.2018.00026
- Groves, Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). Survey methodology (2nd ed). Wiley.
- Guest, G., Namey, E. E., & Mitchell, M. L. (2012). Collecting Qualitative Data: A Field Manual for Applied Research. SAGE Publications.

- Heberlein, T., and Baumgartner, R. (1978), "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature," American Sociological Review, 43, pp. 447-462.
- Kleyman, B. (2012). Colocation Selection: Best Practices and Critical Considerations for Choosing the Right Data Center Colocation Solution (p. 22). RagingWire Data Centers.
- Lischka, K. (2011). John McCarthy: Der Vater der Rechner-Cloud ist tot DER SPIEGEL Netzwelt. Retrieved May 5, 2020, from https://www.spiegel.de/netzwelt/web/john-mccarthy-der-vater-der-rechner-cloud-ist-tot-a-793795.html
- Reddy, V. D., Setz, B., Rao, G. S. V. R. K., Gangadharan, G. R., & Aiello, M. (2017). Metrics for Sustainable Data Centers. IEEE Transactions on Sustainable Computing, 2(3), 290–303. https://doi.org/10.1109/TSUSC.2017.2701883
- Reportlinker. (2020). The US data center market is expected to grow at a CAGR of over 1% during the period 2019-2025. Retrieved September 28, 2021, from https://www.prnewswire.com/news-releases/the-us-data-center-market-is-expected-to-grow-at-a-cagr-of-over-1-during-the-period-20192025-301046940.html
- Taneja, M. (2020, December 29). Largest data centers in the world in 2020: Data Centers for Data Science. Analytics Vidhya. Retrieved November 10, 2021, from https://www.analyticsvidhya.com/blog/2020/09/8-largest-data-centers-world-2020/.
- Tatnall, A. (Ed.). (2012). Reflections on the History of Computing (Vol. 387). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33899-1
- Technavio. (2020, October 6), Data Center Server Market—Post Pandemic Business Strategies and Processes AP NEWS. https://apnews.com/press-release/business-wire/virus-outbreak-technology-latin-america-and-caribbean-data-storage-cloud-computing-1cc9745e2aad4ebf9ba56601eeab50a7
- Technavio. (2020, October 6). Data center server market Post Pandemic business strategies and PROCESSES: Enterprise Server refresh cycles to BOOST Growth: Technavio. AP NEWS. Retrieved September 28, 2021, from https://apnews.com/press-release/business-wire/virus-outbreak-technology-latin-america-and-caribbean-data-storage-cloud-computing-1cc9745e2aad4ebf9ba56601eeab50a7.
- Tourangeau, R., Rips, L., and Rasinski, K. (2000), The Psychology of Survey Response, Cambridge: Cambridge University Press.
- Uptime Institute. (2018). Uptime Tier Standard: Topology. https://uptimeinstitute.com/publications/asset/tier-standard-topology

## Appendix 1

## **Data Center Trade Organizations**

Below is a listing of organizations targeted for participation in the research survey:

#### **AFCOM**

Description: AFCOM is the leading and longest-standing association in the data center

industry.

Website: https://www.afcom.com/

Contact Information: CustomerSupport@afcom.com

## Data Center World (paired with AFCOM)

Description: Data Center World is the global conference for data center, facilities, and IT

professionals.

Website: <a href="https://www.datacenterworld.com/">https://www.datacenterworld.com/</a>

Contact Information: Public Relations, Meryl Franzman, Meryl.Franzman@informa.com

## DCA GLOBAL (Data Centre Alliance)

Description: Data center trade association.

Website: <a href="https://dca-global.org/">https://dca-global.org/</a>

Contact Information: info@dca-global.org

#### 7X24eXChange International

Description: 7×24 Exchange is the leading knowledge exchange for those who design,

build, operate and maintain information infrastructures.

Website: https://www.7x24exchange.org/

Contact Information: Local chapter: Texas South Leader: Jack W. Kolar 832 759-4227

jackoolair@outlook.com

#### Data Center Coalition

Description: The Data Center Coalition (DCC) is the trade association for the data center

industry.

Website: <a href="https://www.datacentercoalition.org/">https://www.datacentercoalition.org/</a> Contact Information: info@datacentercoalition.org

21

Appendix 2

**Data Center Trade Organizations Participation Email Draft** 

Dear < COMPANY NAME>,

I am a graduate student in The University of Texas at Dallas' Public Policy and Political

Economy doctoral program. My dissertation research is focused on U.S. data centers. As such, I

want to invite your organization's participation in a short 10 to 15-minute email survey of your

members. At no cost to you, the mutually beneficial objective of the survey will be to gain a

better understanding of the issues and concerns faced by the data center industry. I am

particularly interested in surveying opinions on data center growth and development but would

be willing to incorporate any specific survey questions your organization may wish to explore.

All data, analyses, and reports will be shared with your organization. You will be authorized to

publish them freely. The membership email listing and individual responses would be held in

strict confidentiality. The survey will be prepared under appropriate professional and ethical

standards. These standards include an evaluation by the University's Institutional Review Board

(IRB) which monitors ethical adherence in social study research.

I look forward to your reply.

Sincerely,

Glen Cooper

The University of Texas at Dallas

Office: 214-914-6756

Gxc162030@utdallas.edu.

#### Appendix 3

## **Data Survey**

Note: The logical format of the below is as follows: General sections are presented as underlined text, subheadings are in plain text title format, question or statement are presented as plain text, question logic in defined in bracketed text, and potential responses are presented in plain text.

#### Introduction

Purpose: Thank you for taking part in this survey. I am Glen Cooper a PhD student at the University of Texas at Dallas (UTD). This questionnaire is part of a project designed to obtain opinions on the future of data centers. Your responses will be part of a larger scientific evaluation of the data center industry. You as a member of the data center industry are in a position to provide unique and valuable insights for this project. There are no right, or wrong answers and all individual responses will be kept strictly confidential. Your name will in no way be connected to the findings of this study. Please answer every question giving your opinion to the best of your knowledge. Again, thank you for your participation.

Time Estimate: This survey is estimated to take around 10 minutes.

Data Center Definition: For the remainder of this survey, the definition of a Data Center will be a building or portion of a building whose primary purpose is to contain a computer room and related support areas. Generally, data center facilities include computer systems, telecommunications and storage systems, backup power supplies, redundancy systems, security devices and controls, and environmental systems (e.g., air conditioning, fire suppression, etc.).

[Question Logics: Response requirements = None / Display Logic = None]
Individual Organization Demographics

Opening Message

Please answer the below questions as they relate to your most current data centerrelated employment. Answer all questions based upon your best current knowledge.

[Question Logics: Response requirements = None / Display Logic = None]

## Questions

Which best describes your company:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Local

Regional

National

International

Other [fill in]

I do not know

Approximately how many years has your company been in business:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

0-5 years

5-10 years

10-20 years

```
20-30 years
       >30 years
       I have no estimate
Approximately how many people are employed by your company:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       0-50 people
       51-250 people
       251-500 people
       501-1000 people
       1001-2000 people
       >2000 people
       I have no estimate
Approximately how many people are employed by your company in DATA
CENTER OPERATIONS:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       0-50
      51-250
       251-500
       501-1000
       1001-2000
```

>2000

Don't know

How best would you describe the MAJORITY of your company's data center(s) operations:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Managed Hosting Platform (i.e., businesses rent IT equipment and infrastructure from your company)

Colocation Data Centers (i.e., enterprises rent space within your facility and install their computer equipment)

Enterprise Data Centers (i.e., the facility is wholly owned and operated by the business)

Other [fill in]

I do not know

Please indicate your opinion regarding the below statements about your company's data center operations (Strongly Agree, Agree, Disagree, Strongly Disagree)

[Question Logics: Response requirements = Request response / Display Logic =
None / Add statement group = Mobile friendly / Question Type = Matrix table /
Matrix type = Likert / Answer type = Allow one answer / Randomization = Yes]
Growth Questions

My organization's goals are aligned with industry growth potential

My organization's culture is conducive to growth

Strategic Planning Questions

My organization has a clearly defined vision for the future

My organization has an effective strategic planning process

My company effectively monitors industry changes in technology,
government regulations, and business practices

#### Change Management

My organization has strategies in place to accommodate change My company has a successful approach to implementing operational strategies

A lack of employee empowerment reduces my company's ability to execute initiatives

## **Innovation Questions**

My enterprise fosters a culture of innovation

My company continually experiments with new operational and business strategies

My organization places a higher budgetary priority on maintenance than innovation

Based on the following scale (Not Concerned, Somewhat Concerned, Concerned, Very Concerned), how concerned is your organization about the following industry issues:

[Question Logics: Response requirements = Request response / Display Logic =
None / Add statement group = Mobile friendly / Question Type = Matrix table /
Matrix type = Likert / Answer type = Allow one answer / Randomization = Yes]
Uptime availability

Operational monitoring

Management and organizational structure

Security

Energy efficiency

Space availability

[Question Logics: Response requirements = Request response / Display Logic = None]

Overall, in the next 5 to 10 years what percentage increase in load capacity do you expect your company's data center(s) to experience (select the highest applicable percentage range):

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

0%-25%

26%-50%

51%-75%76%-100%

over 100%

I have no estimate

In the next 5 to 10 years, in your opinion, how will growth MOSTLY be supported within your organization (select one):

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Addition of new servers and supporting hardware

Server upgrade and replacement

Virtualization

Other \_\_\_\_\_

I have no opinion

You expect Artificial Intelligence to have a major impact on your company's operations in the next (select one):

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

5 years

10 years

20 or more years

Will not have an effect

I have no opinion?

## **Data Center Industry**

## Opening Message

Thank you for your time so far. You are about half of the way through the survey.

Please answer the below questions as they relate to the data center industry as a whole.

[Question Logics: Response requirements = None / Display Logic = None]

## Questions

Do you see a current supply gap in data center labor?

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Yes

In your opinion, what is causing the current data center labor shortage (select all that apply):

[Question Logics: Response requirements = Request response /
Display Logic = None / Question Type = Multiple choice / Answer
type = Allow one answer / Conditioned on "Do you see a current
supply gap in data center labor" answer = Yes]

The sector is relatively invisible to labor, causing too few candidates entering the market.

Labor recognizes the industry potential but market demand is increasing faster than supply.

Working conditions, such as shift scheduling, is undesirable to potential candidates.

Candidates pay is below similar technology jobs.

Positions lack upward mobility.

Position knowledge lacks transferability to other technology jobs.

Positions lack social glamor.

Retirement of existing management and technicians.

How many years into the future do you expect the labor supply shortage to continue:

[Question Logics: Response requirements = Request response /
Display Logic = None / Question Type = Multiple choice / Answer
type = Allow one answer / Conditioned on "Do you see a current
supply gap in data center labor" answer = Yes]

5-10 years

11-20 years

21-30 years

>30 years

Do you expect that Artificial Intelligence and other automation will have a major industry impact in the next:

[Question Logics: Response requirements = Request response / Display Logic = None / Question Type = Multiple choice / Answer type = Allow one answer]

5 years

10 years

20 years

No impact

Within the next 5 to 10 years, how do you believe MOST data center growth will be accommodated:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Addition of servers

Replacement of latent servers (i.e., hardware refreshes)

Improvements in operational efficiencies

Other
Do you anticipate utility power supply will impede industry growth in the within
the next:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
5 years
10 years
20 years
No impediment
Over the next 10 to 20 years, what do you see as the PREDOMINATE future for
data center's Uninterruptible Power Supply (UPS):
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
Static UPS Systems (batteries)
Rotary UPS Systems (kinetic energy)
Other
In the next 5 to 10 years, where do you see the largest data center growth
potential:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
Public clouds
Private clouds
Other

Over the next 5 to 10 years, what do you believe is the GREATEST impediment to industry growth (select one):

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

**Electrical Power Availability** 

Security, Reliability

**Space Constraints** 

Cooling technology

Labor availability

Other \_\_\_\_\_

There are no impediments

In the next 5 to 10 years, do you believe new environmental laws will have a [fill in blank] impact on data center operations.

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Small

Medium

Large

In the next 5 to 10 years, do you believe new privacy laws will have a [fill in blank] impact on data center operations.

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Small

Medium Large In the next 5 to 10 years, which of the following do you believe will be the BIGGEST driver of data center growth: [Question Logics: Response requirements = Request response / Display Logic = None / Question Type = Multiple choice / Answer type = Allow one answer] Modular computing Cloud-based services Internet of Things (IoT) Big Data / Data Lake Storage Other \_\_\_\_\_ What do you believe is currently the fastest growing service provided by the data center industry [Question Logics: Response requirements = Request response / Display Logic = None / Question Type = Multiple choice / Answer type = Allow one answer] Software as a Service (SaaS) Platform as a Service (PaaS) Infrastructure as a Service (IaaS) Cloud Storage Other \_\_\_\_\_ In the next 5 to 10 years, what do you envision will be the bigger issue with

public cloud storage:

```
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Architecture transparency
       Access security
       Management processes
       Reliability of service
       Other _____
In the next 5 to 10 years, where do you see Power Usage Effectiveness (PUE)
values headed
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Higher (less efficient)
       Lower (more efficient)
       About the Same
In the next 5 to 10 years, do you see industry Power Density:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Improving rapidly
       Increasing slowly
       Remaining about the same
       Decreasing
```

What is your estimate for the overall cumulative growth in worldwide data center storage and processing demand in the next: 5 to 10 years:

```
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Less than 0%
       0-25%
       26-50%
       51-75%
       76-100%
       76-100%
       Over 100%
       76-100%
       Over 100%
During the next 5 to 10 years, what industry do you expect will show the greatest
growth in data center usage:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Banking, financial services and insurance (BFSI)
       Manufacturing
       IT and Telecom
       Energy
       Healthcare
       Government
       Entertainment and Media
       Retail
```

Other \_\_\_\_\_

What is your estimate for the overall cumulative growth in worldwide data center storage and processing demand in the next 10 to 20 years:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Less than 0%

0-25%

26-50%

51-75%

76-100%

## Personal Employment Demographics

# Opening Message

The survey is almost done. All that remains are the demographic questions.

Please answer the below questions as they relate to your most current data center employment.

 $[Question\ Logics:\ Response\ requirements = None\ /\ Display\ Logic = None]$ 

Questions

How best would you describe your position at your company:

[Question Logics: Response requirements = Request response / Display

Logic = None / Question Type = Multiple choice / Answer type = Allow

one answer]

Senior Executive

Executive

Middle Management

Technician

**Outside Consultant** 

Other \_\_\_\_\_

How many years have you held that position at your company:

[Question Logics: Response requirements = Request response / Display

Logic = None / Question Type = Multiple choice / Answer type = Allow

one answer]

1-5 years

6-10 years

11-15 years

over 15 years

# Personal Demographics

Opening Message

Please answer the below questions as they relate to your current status.

[Question Logics: Response requirements = None / Display Logic = None]

Questions

What gender do you identify as:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Male

Female

Non-binary / third gender

Prefer not to say.

```
What is your age:
```

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Less than 21 years old

22 - 30 years old

31 - 45 years old

45 years or older

Prefer not to say

Please specify your ethnicity.

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Caucasian

African-American

Latino or Hispanic

Asian

Native American

Native Hawaiian or Pacific Islander

Two or More

Other/Unknown

Prefer not to say

Where is your home located

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

North America/Central America
South America
Europe
Africa
Asia
Australia
Caribbean Islands
Pacific Islands
Other:
Prefer not to say
What is the highest degree or level of education you have completed:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
Some High School
High School
Bachelor's Degree
Master's Degree
Doctorial
Trade School
Prefer not to say
What is your current employment status:
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type – Multiple choice / Answer type – Allow one answer!

**Employed Full-Time** 

**Employed Part-Time** 

Seeking opportunities

Retired

Prefer not to say

What is your annual household income:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Less than \$25,000

\$25,000 - \$50,000

\$50,000 - \$100,000

\$100,000 - \$200,000

More than \$200,000

Prefer not to say

#### **Other**

Please use the space below to provide any additional comments about this questionnaire, your organization, or the data center industry. All comments remain strictly confidential.

[Question Logics: Response requirements = None / Display Logic = None]

#### Conclusion

We thank you for your time spent taking this survey. Your response has been recorded

[Question Logics: Response requirements = None / Display Logic = None]

#### Appendix 4

### **Pretest Survey Email**

Subject: Glen's Survey

Hello,

I have developed a survey and would like your help in conducting a pretest of that survey. I am in a doctoral program at the University of Texas at Dallas, as many of you know. As part of this program, I am researching the development of United States' data centers. Much like the train stations of the 19th century represented the central connecting points for continental travel, data centers represent the connecting points for much of our current digital world. Today's data centers are buildings whose primary purpose is to contain a computer room and related support areas. They represent the core of the internet and electronic commerce.

YOU NEED TO KNOW NOTHING ABOUT DATA CENTERS. I just what to know how the survey works and your feelings about it. If you answered the question with whatever knowledge you have about the topic and with whatever you think might be an appropriate answer, that would be great. AGAIN, YOU ARE HELPING ME EVALUATE THE SURVEY, NOT DATA CENTERS.

If you would be willing, please click on the below link and take the survey; if you could do so in the next couple of days, that would be best.

After you take the survey, I will follow up with another very short survey to get your feelings about this first survey.

I appreciate any help you can provide.

Glen

#### Appendix 5

# **Survey on Survey Email**

Subject: Glen's Survey about the Survey

Hello,

Thank you very much for taking my survey. The mechanical part of the survey went reasonably well, and I did learn several things which I can use during my larger distribution. Also, thank you for any comments you made. They have been noted.

I have ONE MORE FAVOR to ask. Would you mind taking this new survey about the survey? It should only take a couple of minutes but will help me gather your impressions about the first survey. Your thoughts will help me refine the final version. I want to capture your ideas while they are still fresh, so If you could complete this new survey in the next couple of days, that would be best.

I once again appreciate your help.

Glen

## Appendix 6

# Survey on Survey Data Center Survey of Survey

Note: The logical format of the below is as follows: General sections are presented as underlined text, subheadings are in plain text title format, question or statement are presented as plain text, question logic in defined in bracketed text, and potential responses are presented in plain text.

#### **Introduction**

Thank you so much for taking my Data Center Survey pretest. The pretest's purpose was to evaluate any problems in executing the survey and with the formatting of questions in the study that might impact a much larger deployment. This follow-up survey is designed to solicit your feedback about the Data Center Survey.

If you previously sent me an email/text with comments, please ALSO COMPLETE this survey. However, there is no need to repeat your previous comments here as I have already captured them.

Time Estimate: This survey is estimated to take around 1-5 minutes.

[Question Logics: Response requirements = None / Display Logic = None]

Opening Message

Please answer the below questions.

[Question Logics: Response requirements = None / Display Logic = None]

#### Questions

How long did you feel the survey took:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer

1-5 minutes

6-10 minutes

11-15 minutes

more than 15 minutes

Did you feel that the survey was:

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Too Short

Too Long

About Right

Did you feel that the questions had a natural flow?

[Question Logics: Response requirements = Request response / Display Logic =

None / Question Type = Multiple choice / Answer type = Allow one answer]

Yes

No

Please comment on how you believe the flow might be made better:

 $[Question\ Logics:\ Response\ requirements = Request\ response\ /\ Display$ 

Logic = None / Question Type = Multiple choice / Answer type = Allow

one answer / Conditioned on "Did you feel that the questions had a natural

flow?" = No

Did you notice any grammatical errors?

[Question Logics: Response requirements = Request response / Display Logic = None / Question Type = Multiple choice / Answer type = Allow one answer]

Yes

No

Were the grammatical errors so numerous as to be distracting?

[Question Logics: Response requirements = Request response / Display

Logic = None / Question Type = Multiple choice / Answer type = Allow

one answer / Conditioned on "Did you notice any grammatical errors?" =

Yes]

Yes

No

Were there particular questions you found nonsensical?

[Question Logics: Response requirements = Request response / Display Logic = None / Question Type = Multiple choice / Answer type = Allow one answer]

Yes

No

Please comment as to generally which questions you found nonsensical

[Question Logics: Response requirements = Request response / Display

Logic = None / Question Type = Multiple choice / Answer type = Allow

one answer / Conditioned on "Were there particular questions you found

nonsensical?" = Yes]

Did you feel you had to go back and reread or repeat questions?

```
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Yes
       No
              Was the rereading process distracting?
              [Question Logics: Response requirements = Request response /
              Display Logic = None / Question Type = Multiple choice / Answer
              type = Allow one answer / Conditioned on "Did you feel you had
              to go back and reread or repeat questions?" = Yes]
                     Yes
                     No
Were there questions that made you feel uncomfortable?
[Question Logics: Response requirements = Request response / Display Logic =
None / Question Type = Multiple choice / Answer type = Allow one answer]
       Yes
       No
       Please comment as to generally which questions made you
       uncomfortable____
       [Question Logics: Response requirements = Request response / Display
       Logic = None / Question Type = Multiple choice / Answer type = Allow
       one answer / Conditioned on "Were there questions that made you feel
       uncomfortable?" = Yes]
```

Please use the space below to provide any additional comments about the Data
Center Survey questionnaire