Journal of International Technology and Information Management

Volume 22 | Issue 3 Article 4

11-6-2013

Implementing Cloud Computing In Small & Mid-Market Life-Sciences

Thomas Sommer
Quinnipiac University

Ramesh Subramanian
Quinnipiac University

Recommended Citation
Sommer, Thomas and Subramanian, Ramesh (2013) "Implementing Cloud Computing In Small & Mid-Market Life-Sciences," *Journal of International Technology and Information Management*: Vol. 22: Iss. 3, Article 4.
DOI: https://doi.org/10.58729/1941-6679.1015
Available at: https://scholarworks.lib.csusb.edu/jitim/vol22/iss3/4

Implementing Cloud Computing in Small and Mid-Market Life-Sciences: A Mixed-Method Study

Thomas Sommer
Quinnipiac University
USA

Ramesh Subramanian
Quinnipiac University
USA

ABSTRACT

This paper uses a mixed-method approach to study cloud computing implementation in emerging biotech and pharmaceutical companies. The study investigated four small biotech and pharmaceutical companies. The objective was to determine the positive and negative impacts of cloud computing and the impact of such implementation – especially the business impacts in an increasing global and competitive environment. The unique benefits, drawbacks, and various drivers of cloud implementation in these companies were identified. The research techniques were mixed qualitative methods that included action research, observations, and review of other case studies. The research indicated that small biotech and pharmaceutical companies found cloud computing to be very attractive albeit with some drawbacks. The paper provides a detailed discussion of the rationale in opting for cloud implementation by the emerging life-sciences companies; a comprehensive literature review of cloud implementation; the processes involved during the planning and implementation stages. The paper concludes by providing detailed recommendations on cloud implementation to organizations in the life-sciences domain.

INTRODUCTION

Life-sciences organizations such as bio-tech and pharmaceutical companies are natural candidates for cloud-computing implementations. These companies, whether they are start-ups or well established market players, are typically characterized by their intense research and development (R&D) efforts that constantly generate massive volumes of data. The data emanate from various phases that include the pre-clinical trials phase, clinical trials phase, drug approval phase and finally the production, marketing and post-sales phases. At each phase, the large volumes of data generated require to be analyzed, categorized and securely stored in accordance with regulatory mandates and corporate directives. The data analysis and processing can be time consuming, yet have to be accomplished expeditiously, given the intensely competitive global environment under which these companies operate. The need for cost and time controls is critical, as life-sciences companies need to adhere to increasingly aggressive development timelines and comply with changing global regulations in a timely manner. The cost and security requirements alone would deter most start-ups in the life-sciences.

Recent developments in cloud computing and cloud implementations offer the promise of efficient and cost-effective data analysis and processing as well as multi-layered security and controls to small biotech and pharmaceutical organizations that often cannot achieve these on

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

their own. Given the constraints that small life-sciences companies face, and the promise that cloud computing offers, it is useful to study the experiences of small to medium sized life-sciences companies that have moved to implement cloud-based computing. This paper discusses the results of a study of cloud computing implementation in the context of four emerging life-sciences companies in North America. The study identifies the common themes shared by the companies in their move towards cloud implementation. The focus is on the processes involved, perceptions of the various stakeholders, benefits and challenges, and lessons learned.

Our primary focus is the experiential aspects of cloud implementation and the realistic challenges that such implementations pose on small and emerging life sciences companies. As a result of our study we are able to provide a comprehensive set of recommendations pertaining to cloud implementation that would be useful to other similar small and medium life-sciences (and possibly other) companies. The study used a mixed-method approach, employing library and media research, public and confidential documents, interviews, surveys and case-studies. We believe that our results and recommendations are a valuable addition to the available literature in the field, in addition to serving as a comprehensive guide to potential cloud implementers.

This paper is organized as follows. In the section that follows immediately, we provide a background on cloud computing and various cloud computing models. Following that, we provide a detailed literature review and show how our work differs from past work in this area. We then list the questions that we hope to provide answers for. We follow that with a discussion on the methodology used in our study along with our justification and the data collection methods we employ. This is followed by a detailed analysis of the data collected. The data analysis includes cases of representative companies that were studied. Finally, we provide a comprehensive list of recommendations for small life-sciences companies that intend to move towards cloud computing. Following that we offer our conclusions and propose some future work in this area.

BACKGROUND

Despite the present-day hype that surrounds it, the concept of cloud computing mystifies many organizations – especially those dealing with the deluge of data being generated, such as small life-sciences companies. Multiple terms are often used to describe cloud computing, e.g. grid, distributed, on-demand, cluster, utility, virtualization, and software-as-a-service. The U.S. National Institute of Standards and Technology (NIST) provides a "Working Definition of Cloud Computing" in the document NIST 800-145. NIST's definition describes five crucial characteristics of cloud computing (i.e. broad network access, rapid elasticity, measured service, on-demand self-service, and resource pooling), three cloud service models (i.e. Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS)), and four cloud deployment models (i.e. public, private, hybrid, and community). Simply stated, cloud computing refers to end-users connecting, on demand, with applications or services running on shared servers, often hosted and virtualized, rather than traditional, in-house, dedicated servers. Thus cloud computing dramatically differs from the decades-old client-server computing model where applications were assigned to specific hardware, often residing on-location in data centers. On-demand cloud computing also empowers its end-users by allowing them to use their choice of Internet-connected devices, anywhere or at any time (Knorr & Gruman, 2009).

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

An important factor in favor of cloud computing is its low cost, pay-as-you-grow business model. This model puts sophisticated and complex computing capabilities within reach of even small companies, thereby enabling these companies to use and transform information technology into an engine that drives business (Cognizant, 2010). Companies could use cloud computing to economically scale their businesses as needed, while rapidly completing complex research-to-market tasks they simply could not accomplish on their own (Bowers, 2011). Thus cloud computing promises cost-savings, increased agility, and the type of scalability that responds to the rapid changes in both technology and business needs.

REVIEW OF LITERATURE

A review of research on cloud computing in life sciences reveals that this is an active area of research. The prior research has generally focused on the following areas:

- **Cost reduction and increased speed:** According to Proffitt (2009), early adopters of cloud computing such as Pfizer, Johnson & Johnson, and Eli Lily all used Amazon Web Services (AWS) and Amazon EC2 (Elastic Compute Cloud). These pharmaceutical companies were able to perform R&D using the cloud, and process proteomics, bioinformatics, statistics and adaptive trial design more rapidly with predictable time and costs. BioTeam co-founder and technology director Chris Dagdigian noted that Amazon's cloud computing rates started at 10 cents per hour, and calculated that a traditional 100 CPU-hour research problem could be solved using EC2 in 1 hour for $40 (Davies, 2009). Heritage (2012) argued that small pharma and biotech companies which do not have the resources to support a large informatics infrastructure, could use SaaS based informatics that save time, costs, which are more effective with better collaborative workflows. Another study showed how BT in England partnered with Accelrys to create a life sciences research and development cloud, or On Demand Compute service. The partnership helped scientists working in the pharmaceutical and biotech industries reduce the costs associated with R&D when bringing new drugs to market (Nguyen, 2012).
- **Effects of enhanced connectivity:** Effective connectivity is a major factor in life-sciences research and development. Kubick (2011) argued that cloud computing, using a solitary Internet connection, could reduce the effort of individually integrating each research system at various locations yet provide availability to all. Bowers (2011) suggested that cloud service providers (CSPs) could provide SMB life-sciences organizations with "best practices" such as comprehensive data protection, 24x7 disaster recovery, multi-site replication, real-time monitoring, and state-of-the-art emergency response systems – services they generally could not afford.
- **Scalability and performance:** According to May (2010), most life science companies lack the necessary computing infrastructure required to analyze and store their research data. In order to increase their computational power, many life science researchers are searching beyond their own organizations and turning to decentralized systems, like supercomputers or grids of many smaller computers working together. Along the same lines, Taylor (2010) noted that bioinformatics researchers, using ultra large data sets, achieved better performance with respect to management of failures,

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

data analysis, and computational jobs using cloud based systems. Afgan et al. (2010) illustrated how to deploy a compute cluster using the Amazon EC2 cloud and Galaxy CloudMan. Their system was simple to use and enabled small groups of researchers to deploy heavy computational resources without requiring support from skilled bioinformatics personnel. Barbadora (2012) noted that an increasing number of SMB life sciences companies are choosing cloud-based CRMs that give them enterprise-class performance within a reasonably priced model. These systems historically are cost-prohibitive for most small life sciences companies because traditional in-house CRM applications, ongoing maintenance, and upgrade costs are normally beyond their budgets.

- **Collaborative drug discovery and data analysis:** CSPs are now providing cloud-based drug discovery software platforms that increase the power of collaboration. These private, secure cloud architectures create a barrier-free virtual world that permits researchers in remote areas, sometimes with few technological resources, to fully participate in research projects (McCarthy, 2012). Vandeweyer, Reyniers, Wuyts, Rooms, and Kooy (2011) discussed on the collaborative advantages of cloud computing and noted that open-source web based platforms such as CNV-WebStore are being used in clinical practice by both lab technicians and clinicians to compare results against clinical information without producing overwhelming amounts of data (Vandeweyer et al, 2011). Chidambaram studied how a small life sciences company can effectively manage their collaborations and document sharing, CRM, and ERP all in the cloud (Chidambaram, 2011).
- **Cloud database technologies and their efficacy:** Taylor (2010) noted that new open source database technologies such as Apache Hadoop, Hbase, and MapReduce are cost-effective, reliable, scalable, and distributed. These technologies in the cloud provide SMB life science companies with distributed processing of large data sets across clusters of computers using a simple programming model (Taylor, 2010). Do, Esteves, Karten, and Bier (Do et al, 2010) researched Booly, a similar cloud-based relational database that runs on multiple load balanced servers and can easily be accessed through a graphical user interface using a web browser. They noted that Booly is a comprehensive platform for the creation, storage, and integration of biological databases that can assist researchers in developing novel discoveries in the laboratory. Qiu et al. maintained a hybrid cloud that utilized MapReduce combined with Message-Passing Interface (MPI) standard, for programming parallel computers – technology that offers an appealing production environment for life sciences applications. They also used three cloud-based computational infrastructures in their study (Azure, Amazon and FutureGrid) and showed how life science organizations with few resources could successfully create this environment (Qiu et al., 2010).
- **Cloud security issues:** Sansom (2010) argued that while cloud computing may offer compelling solutions for small companies that struggle with very large data sets, only "precompetitive or non-confidential" data would be used in the cloud due to security issues. However, CSPs are currently helping small life science start-up companies manage their public and private clouds using Amazon and Google. Kubick (2011) noted the security and loss of control issues resulting from cloud computing but argued that in most cases a CSP provides much higher degrees of disaster recovery and auditable security than most internal IT departments. Bowers (2011) expanded on

ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

these ideas and suggested data security in the cloud must be strong, extensive, and dependable, and adhere to regulatory requirements, and went on to suggest a variety of codes, regulations and certifications that pertained to life-sciences companies. A 2008 study by Gartner Research (Gartner Research, 2008) suggested that security delivered as a cloud-based service, would more than triple in life science organizations by 2013.

- **Regulatory issues:** Besides cloud security, another concern of many life science organizations is the typical regulatory concerns. Kubick (2011) noted that CSPs should comply with Health Insurance Portability and Accountability Act (HIPAA), Good Clinical Practice (GCP) standards, and have their cloud services regularly audited. However he also suggested that regulatory agencies and their lack of definitive regulatory positions can negatively influence attitudes regarding cloud adoption, thus making small life science companies wary of cloud computing. Gorban (2012) maintained regulated environments that utilize cloud computing can manage compliance with the use of strong controls and auditable documentation, thereby mitigating associated risks.
- **Efficiency issues:** Clinical trials at life science companies have become more global and regulatory examination has continued to climb, making paper-based processes even more difficult. Clinical trials, in particular, require the tracking of documents being sent between multiple sites, sponsors, contract research organizations (CROs), and stakeholders. To counter this, Shurell (2010) argued that by using a SaaS-based solution in the cloud, electronic documents can be securely tracked globally, which can accelerate contract negotiations, patient recruitment, protocol design, clinical trials, and other activities of pharmaceutical companies. Bowers (2011) also cited examples of life science organizations that use cloud computing in fundamental research to speed up the development process. He noted that companies were able to bring down fixed IT costs without undermining service levels, while significantly expanding computing and storage capabilities.

The literature survey above shows that cloud computing is currently an active area of research. Various research projects have addressed a variety of issues pertaining to cloud computing, in various contexts. In summarizing the above, the following observations can be made:

- A variety of new technologies (including several open source technologies) have emerged over recent years that enable and enhance cloud computing
- Several large CSPs have emerged in recent years, providing a variety of cloud computing services.
- SMB life science companies could experience remarkable economic advantages and time-saving by migrating their computing to a cloud service provider.
- The process of implementing a cloud solution could take a much shorter time when compared to traditional in-house solutions, without interrupting normal business.
- Cloud-based systems would require minimal in-house IT support, which liberates infrastructure and resources for other activities.
- Cloud computing requires smaller upfront investment and provides predictable cost management that is based on operating budgets instead of capital budgets, which is particularly attractive to start-up companies.

- Established pharmaceutical and biotechnology companies have adopted IaaS and PaaS for computationally heavy research such as molecular modeling, proteomics, and bioinformatics. These companies thus had immediate access to on-demand processing and storage services along with hosted environments for developing custom applications.
- These types of services are especially attractive to start-ups, considering they can avoid capital investments and instead rent the computing infrastructure needed during product development.
- Cloud computing also reduces start-up times and can be made available much faster than it takes to build infrastructure internally, without bureaucratic hurdles or delays.
- Cloud computing enhances global collaborations at substantially low costs.
- Cloud implementations should carefully consider the security and regulatory aspects and implications while planning to move to the cloud.

Some of the literature also describes results accruing from cloud implementation. However, there is not much in the above that studies describe the actual *experience* of cloud implementation in small and medium sized life-sciences companies, and the perceptions, questions and strategic decision-making that accompany such implementations. Nor does the literature discuss the benefits and challenges that accrue from cloud implementation in small life-sciences companies. Our study addresses this gap in the above literature.

As noted above, the main objective of our study was to examine the various aspects of cloud computing implementation as it relates to emerging life-sciences companies. We hoped to identify the unique benefits, drawbacks, and various drivers of cloud implementation in these companies. Given this main objective, the study then looked at the following questions:

1. What were the perceived benefits in a company's move towards cloud implementation?
2. What were the observed challenges in a company's cloud implementation strategy?
3. What were the cloud implementation strategies used by the companies studied (including models, vendors and applications)
4. What were the management and implementation issues associated with cloud implementation (including change management, IT loss of control, avoiding vendor lock-in)?
5. What was the security, legal and regulatory issues associated with cloud implementation?
6. How can the cloud implementations plan for future technology developments (e.g. virtualization)?
7. What would be the recommendations to emerging life-sciences companies with regards to cloud implementation?

METHODOLOGY

Our study primarily followed a mixed-methods approach using techniques such as first-hand observations, interviews, case studies, and peer-reviewed published literature, augmented by surveys. It was thought that the mixed-methods approach would garner multiple points of view for comparative purposes. Bias was minimized during the analysis process through the use of electronic data gathering and validated responses.

ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

Sample Population Selection

The population in this research was four emerging biotech and pharmaceutical companies, each with less than 100 employees. Three of those are private companies and one is publicly listed, with current market value of less than $25 per share. All four companies are considered SMB life science companies with market capitalization of less than $100 million. All organizations and participants utilized cloud computing prior to the beginning of this study, for periods ranging from one to seven years. The sample population was chosen using purposeful sampling based on the participant or stakeholder's experience. Prior permission was obtained whenever gathering data within the organizations, and all participation was voluntary. Specific sample groups included: managers and senior management - with an understanding of the strategic goals and corporate business plan, technical staff - with an understanding of cloud mechanics, and cloud end-users (also known as 'internal customers') within the organization. The sample population consisted of 47 total participants within the four companies: 11 participants were managers or senior managers, 17 participants were technical staff, and 19 participants were cloud end-users.

Data Collection

Qualitative data was collected through detailed interviews, as well as from internal documents, trade journals, media reports and by deploying a survey. Long interviews[1] were conducted on selected representatives from the companies as mentioned above. Interviews were electronically recorded and later transcribed, then given to participants for verification of content and meaning. Validity was determined using interviewer corroboration and member check, or respondent validation, by verifying the quality of the researcher's data and conclusions when compared with the experiences of the research participants. Notes from first hand observations, case studies, and peer-reviewed literature were captured electronically and verified by repetitive review and cross-examination. Additionally, a concept map was created, using Wordle.net, with the electronic data that was collected from the various methods.

Survey data was collected electronically first-hand and remotely by using SurveyMonkey survey software, one laptop, one iPhone, and one iPad. Ten survey questions (please see Appendix Figure A-1) were designed to reveal the advantages and disadvantages of cloud computing and to illicit clusters of opinions and overall themes from research participants. The questions were posed to sample population participants electronically in survey form as well verbally during structured interviews. The surveys allowed the researcher to reach a large majority of the sample population, while the interviews were held independently with approximately 10 percent of the sample population. Survey responses were given to respondents for verification of their accuracy.

[1] A 'long interview' is 'a sharply focused, rapid, highly intensive interview process that seeks to diminish the indeterminacy and redundancy that attends more unstructured research processes. The long interview calls for special kinds of preparation and structure, including the use of an open-ended questionnaire, so that the investigator can maximize the value of the time spent with the respondent' (McCracken 1998). The researchers felt that this type of data collection method would provide a subject's critical interpretation of an event or development as experienced by him or her.

ANALYSIS

To achieve the objectives of the study we needed analyze the data and identify themes that were common to all of the companies studied, especially with respect to perceptions, questions and strategic decision-making among the stakeholders. The various data types collected (text, audio, images, survey, etc.) were input into the TAMS Analyzer, a computer-assisted qualitative data analysis software (CAQDAS) system, which aided in coding and the identification of themes.

A meta-analysis was conducted from reading and reviewing research data, the literature review, observations, notes, interviews, surveys, transcripts, case studies, and research documentation provided by the participant companies. Obvious patterns that reflected the advantages and disadvantages of cloud computing were gathered. The concept map was regularly updated, which helped visually classify, code, and identify meaningful clusters. Through this exercise and additional sorting and organizing, patterns and themes began to emerge.

These patterns and themes can be broadly categorized as *benefits*, *challenges*, and the associated *ramifications*.

Benefits

Based on the data aggregated from the interviews and surveys, we identified several benefits, the most significant of which (as noted by the participants) were reduced cost, the ability to avoid future IT expenses. Other benefits listed were:

- Improved scalability of IT resources – access to extensive file storage
- Quick set-up, more flexible, 24x7 connectivity and management.
- Enhanced security and privacy features.
- Increased collaboration on drug discovery.
- Availability of on-demand, redundant cloud databases.
- Better life sciences efficiency and regulatory compliance.
- The availability of multiple CSPs, thus providing more comparison and choices.

Overall, the participants were positive about cloud experience. The following case vignette of one of the companies studied corroborates these positive aspects.

> *Vignette 1: Company A found its business development and strategic partnerships suddenly required extensive file storage and file sharing. The company wanted extra storage available for both internal users and third-party partners, regardless of location or time of day. With limited resources and no full-time IT staff, the company chose cloud file servers that were quickly setup and then managed 24x7 by cloud service providers. The company easily added several terabytes of cloud storage for secure file sharing and collaboration via Windows, Mac, and iOS/Android mobile applications. The company found user management was simple and included auditable privacy/authentication controls (AD/LDAP), file versioning, file locking (check in/check out), and virtual dropbox (drag-n-drop) services. These cloud file servers included redundant real-time data center mirrors that were SSAE 16 compliant colocation*

facilities and SAS70 Type II compliant data centers, with 99.999% up-time guarantees. For less than $5 per user/per month the company's cloud file servers provided enterprise class features with continuous 256-bit AES encryption (in transit and at rest), and multi-level access controls that were HIPAA and Safe Harbor compliant. The company found its initial file sharing experience so positive that it identified new uses for its cloud storage. The company added real-time disaster recovery of its critical IT systems, desktop, laptop, and mobile users. With persistent file versioning these new backups complied with the company's data retention policy and permitted litigation holds of off-site files in the event of future legal actions. Additionally the company's project teams found cloud storage easier to collaborate with regardless of file size or local of worldwide partners, and it provided limited risk of violating HIPAA compliance. Overall the company's experience with cloud storage and cloud file sharing was very positive.

Challenges

The biggest challenge noted by the participants was cloud security. Other challenges that were explicitly mentioned or became apparent during the study were:

- Maintaining Privacy and confidentiality of corporate data.
- Loss of control within the IT function (e.g. managing multiple platforms and devices)
- Maintaining reliability of key systems and availability of services or data.
- Lack of organizational control over services or data.
- Legal ramifications (government regulations, compliance and auditing)
- Concerns over cloud vendor lock-in.

The existence of these challenges was corroborated by another illustrative case vignette:

Vignette 2: Almost every employee at Company B frequently used some type of mobile device (smartphone or tablet) for both work-related and personal computing. All of these devices connected to the company's cloud environment. Managing or controlling that process quickly became a challenge. Mistakes were made early on, ranging from not researching (ahead of time) how workers could best use the devices, to underestimating the costs and the additional security challenges mobile devices present. The use of multiple devices rapidly changed what the company felt was the best practice of its cloud infrastructure. As the company grew and the staff worked longer hours, nights and weekends, personally owned BOYD (bring-your-own-device) mobile devices quickly became a low-cost (owned by the employee) necessity that required the organization to adapt. Initially the company did not grasp or plan for the full impact of mobile cloud computing and soon became overwhelmed by its failure to adequately manage and support mobile devices. BYOD also quickly exposed the company to a major security issue as employees learned they could bypass security controls on their company-owned device(s) via "jailbreak" (Apple iOS) or "root" (Google Android). This process allowed users to run pirated applications on their mobile devices and to bypass built-in security controls, in order to fully customize mobile devices more than device manufacturers allow. These workarounds and loopholes also created data leakages and

security holes that could be used by undesired spyware or malware to steal or corrupt sensitive data. In addition, these created the possibility of sensitive corporate data being transferred to unsecured personal devices without the ability to remote wipe the hard drives if the mobile device was stolen. In order to mitigate these issues and protect its cloud data the company was compelled to make several operational changes (discussed in more detail later in the Recommendations section). The changes included: (a) Developing a comprehensive mobile security plan after reassessing its overall security policies and procedures, and (b) Investing in proper management tools i.e., deploying Mobile Device Management (MDM) system that secured, monitored, and managed all mobile devices that connected to the Company's cloud resources. The MDM also prevented jailbreaking devices and defined applications that were required, permitted, or banned.

Ramifications

The most common ramifications felt by the companies pertained to managing system changes security, regulatory compliance, legal compliance, as well as the emergence of threats such as loss of IT control and vendor lock-in.

Managing System Changes

Participants acknowledged they depended on automatic change management with no additional expenditures for future updates, in terms of software and hardware. Participants expressed concerns about reliability in terms of changes made by CSPs and how these changes affected their business needs and/or might negatively impact their production environments. Therefore the companies required advance notice from CSPs prior to the application of software patches or updates. They took efforts to have redundancies in place to ensure that risks from patches and upgrades were minimal. Appropriate personnel (e.g. IT staff, scientists, and project management staff) monitored, scheduled and approved changes made by the CSPs. Due to these change management efforts, the entire process – from initial cloud deployment to on-going change management proceeded with no significant issues in terms of down-time or additional costs. The participants generally indicated that they had high-availability (HA) systems (Sommer, 2013).

Security Issues

Participants in this study cited, most often, confidentiality and security, and the associated complexity, as the greatest concerns of cloud computing. However, security protection offered by their cloud service providers (CSPs) was considered an advantage by most organizations that utilized strong Service-Level Agreements (SLAs) and appropriate security controls. Most organizations also viewed security controls in cloud computing as no different from security controls in other IT environments. All organizations in this study mitigated risks by requiring CSPs (or themselves) to use strong encryption and privately controlled encryption keys, both during data transit and storage inside the cloud infrastructure. Participants frequently cited their desire for enhanced security in the following areas: corporate data security, application security, process security, infrastructure security, R&D security, and personnel security. A majority of these companies also maintained they required comprehensive security standards be used by their CSPs, such as: HIPAA, SAS70 Type II or SSAE 16, Safe Harbor Compliance, FIPS 200 / SP 800-53, ISO 27001, ISO 27002, SOC 2 & 3, WebTrust and SysTrust, and Certificate of

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

Cloud Security Knowledge (CCSK) (Sommer, 2013). A participant at a company stated that they had strong controls in place and performed frequent audits based on industry security standards. They also used a comprehensive framework provided by the Cloud Security Alliance that helped them evaluate cloud computing risks and informed their security decisions.

Regulatory compliance

Organizations in this case study were required to comply with various regulatory agencies and auditors both inside and outside the United States. Participants suggested that their use of cloud computing provided them with regulatory proficiency that was compliant, scalable, and on-demand. One participant explained that every CSP that the company used complied with their audit and compliance requirements. For example, the participant at the company indicated that its compliance with HIPAA laws required that patient data be kept safe, and therefore the company only used CSPs that provided HIPAA compliance. Participants also frequently noted that their cloud-based systems had full disaster recovery capabilities and underwent frequent audits. One major issue for some of these organizations was the U.S. Food and Drig Administration's (FDA) "21 CFR Part-11" compliance that requires regular validation of systems hardware. These participants required that CSPs provide evidence of "21 CFR Part-11" compliance, such as validated e-signatures, system controls, hardware/software version and revision control, plus reporting and auditing abilities that verified record integrity and reliability.

Legal Ramifications

A majority of participants cited legal issues as a major concern of cloud computing. All participant organizations required clear legal definitions in their CSP agreements on what was/was not being provided by CSPs, ownership of information/system, as well as what should happen in case a vendor filed for bankruptcy. A majority of organizations indicated that when CSPs were clearly aware of the consequences for violating these policies, it motivated them to successfully execute their agreements. Furthermore, these organizations avoided SLAs or contracts that limited, ignored, or glossed over potential data loss, privacy, security and e-discovery issues. These organizations expected CSPs to assume responsibility and liability in case of network outages and data loss (Sommer, 2013). One participant indicated that the company had requested indemnification clauses in some of their service-level agreements that penalized CSPs when agreements were violated. Additional legal concerns that require to be specifically addressed with cloud-based implementations were related to laws such as: Sarbanes-Oxley, the Gramm-Leach-Bliley Act, and the Patient Protection and Affordable Care Act.

Cloud Vendor Lock-in

Another major concern from a majority of organizations cited was their perceived inability to move to another cloud offering or to another CSP. Few organizations had undergone significant moves, therefore much of this fear was speculation. However, in order to alleviate this issue, most organizations viewed data portability as a crucial aspect as it chose CSPs. Deployments that utilized different cloud provider solutions, e.g. for disaster recovery or global presence, were considered the best solutions in terms of portability and risk management (Sommer, 2013). As these organizations grew and expanded to several disparate CSPs, with different infrastructures, operational practices, and security expertise, they expected the levels of complexity would inevitably increase. Organizations realized this requires a pervasive and highly trustworthy method of securing organizational data as it is securely transported data to and from the cloud.

ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

Lack of IS Control

A majority of participants in this study indicated they believed their organizations lacked complete control over their data, which they saw as an ongoing issue. Although most CSPs used by participant organizations deployed fully automated management platforms that maintained IT control and transparency, they could not provide specific instances where lack of control resulted in negative outcomes since cloud adoption. Several participants described experiences prior to cloud adoption when their in-house controls failed, mainly due to their small staffs and inadequate support (Sommer, 2013). Participant from one company explained they quickly overcame their hesitation and turned-over control to a CSP after their own in-house IT systems failed. Overall lack of IS control became less of an issue as time passed and organizations relied more heavily on the scalability and reliability of their CSP provided systems.

RECOMMENDATIONS

Throughout the lifespan of life science organizations – from start-up to R&D, from the pre-clinical phase to clinical trial work, and from drug approval to market – the massive volumes of data constantly generated must be analyzed and securely stored in accordance with regulatory agencies and corporate directives, all while improving cost and time efficiencies. Cloud computing, when properly implemented, has the promise of adding multiple security layers and controls that small biotech and pharmaceutical organizations often cannot accomplish with their scarce resources.

Organizations can start by first introducing cloud computing into routine processes, without large capital expenditures, and then increase usage as necessary. However, cloud implementation also poses many challenges and concerns. Our comprehensive analysis above forms the basis for certain specific recommendations that we feel are major concerns in cloud implementations, especially in emerging life-sciences companies.

Cloud Security & Privacy

Cloud security and privacy in life-science organizations are one of the most critical considerations when considering cloud implementation. These are complex issues, as evidenced by their use of public, private, and hybrid cloud models and varied off-site infrastructure and physical locations. Organizations concerned about overall cloud security can significantly improve data protection and the associated infrastructure with proper planning, evaluation, and monitoring of CSPs along with these key security elements:

- *Application Security*: verify strong encryption and authentication controls are used.
- *Data Security*: verify auditable security checks and best practice cryptography, that prevent breaches, are used.
- *Infrastructure Security*: verify redundancy of infrastructure and uninterruptible service are tested and used.
- *Process Security*: verify industry best practices are used, and managed by certified security professionals.
- *Personnel Security*: verify background checks and strong confidentiality agreements,

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

with all personnel exposed to data, are used.
- *Product Development Security*: verify secure development lifecycle processes are used, that protect applications in production and in development (Cloud Security Alliance, 2011).

Additionally, organizations should enhance their cloud security and privacy by implementing 2048-bit SSL certificates in SaaS systems. Organizations should add more comprehensive cloud security guidelines to their enterprise architecture. The Cloud Security Alliance provides security guidance in 14 domains that cover operation and governance of cloud services (Cloud Security Alliance, 2011). These domains emphasize security and privacy in a multitenant environment, for example:

- Domain 1: Cloud Computing Architectural Framework
- Domain 2: Governance and Enterprise Risk Management
- Domain 3: Legal Issues: Contracts and Electronic Discovery
- Domain 4: Compliance and Audit Management
- Domain 5: Information Management and Data Security
- Domain 6: Interoperability and Portability
- Domain 7: Traditional Security, Business Continuity, and Disaster Recovery
- Domain 8: Data Center Operations
- Domain 9: Incident Response
- Domain 10: Application Security
- Domain 11: Encryption and Key Management
- Domain 12: Identity, Entitlement, and Access Management
- Domain 13: Virtualization
- Domain 14: Security as a Service

Identity Management

An important part of cloud computing is the process of managing users, controlling their access to applications and services, and the authentication of users in a Web Services world. For years two of the most broadly adopted authentication and authorization standards have been Kerberos and Security Assertion Markup Language (SAML). However, cloud service providers are beginning to embrace newer standards such as OAuth 2.0, OpenID Connect, and Simple Cloud Identity Management (SCIM) in order to more easily exchange authentication information between multiple cloud providers and companies from web, mobile and desktop applications. According to the Cloud Security Alliance, SCIM is a newly emerging standard that makes the management of identities inexpensive, with simpler and faster implementation, while making it easy for organizations to migrate user identities into and out of the cloud. Furthermore SCIM is considered cloud-friendly because its RESTful API is supported by many cloud service providers and it works with existing authentication protocols like OAuth 2.0 and OpenID connect (Cloud Security Alliance, 2011). According to analyst Sean Deuby, OAuth 2.0 and OpenID Connect are two newer identity frameworks that support the next generation of web single sign-on (Deuby, 2013). Case Vignette 2 in the Analysis section showcased the challenges brought out by a multiplicity of devices. The main issue is authentication and authorization. Organizations considering cloud adoption should evaluate these newer identity management frameworks and

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

seek cloud service providers that support them, not only for easier interoperability with other cloud service providers, but for use with hybrid cloud environments, partners, and future growth as small companies progress into enterprises. The OpenID Connect Protocol Suite is illustrated in the Appendix A-2.

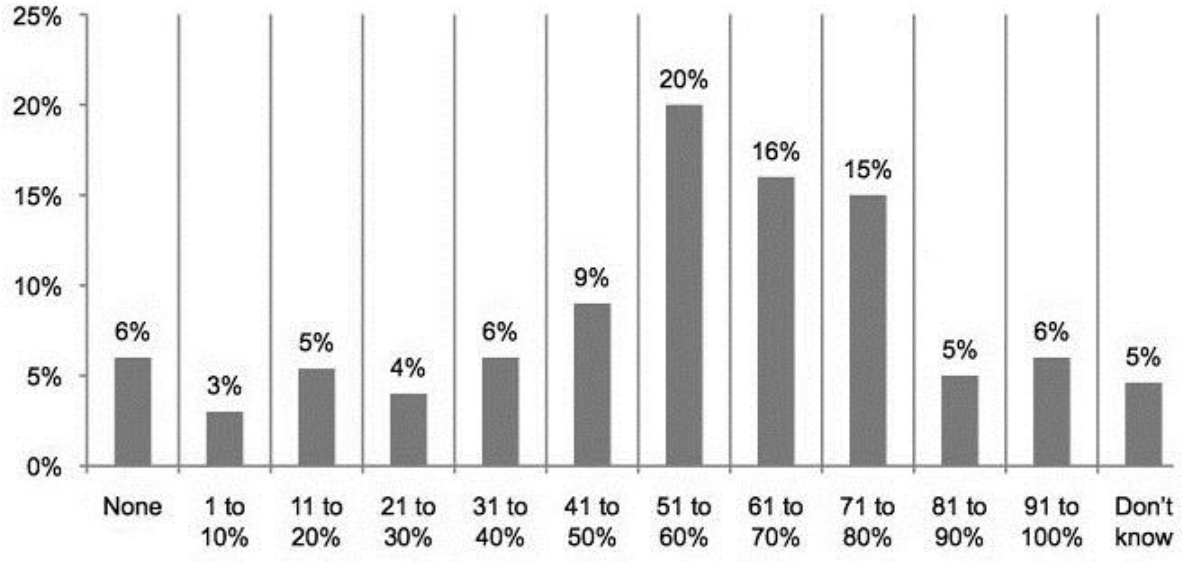
Mobile Security

Case Vignette 2 in the Analysis section also brought out the reality of BYOD in organizations. Cloud computing and the popularity of powerful consumer-branded mobile devices has resulted in employees in wanting to use smartphones or tablets at work and home to share files and data, regardless of time of day. This raises security and privacy issues. According to a recent study titled "The Risk of Regulated Data on Mobile Devices" by Ponemon Institute many companies are not taking the necessary steps to protect sensitive data on mobile devices. Fifty-four percent of respondents in the study had on average five data breach events, from either theft or loss of mobile devices that held regulated, sensitive data (Ponemon Institute, 2013).

Given this situation, small organizations that utilize cloud computing, like the participants in this case study, should develop a mobile device security program that builds on existing network security and Mobile Device Management (MDM) system, which can help address a variety of mobile security concerns, using the following recommendations:

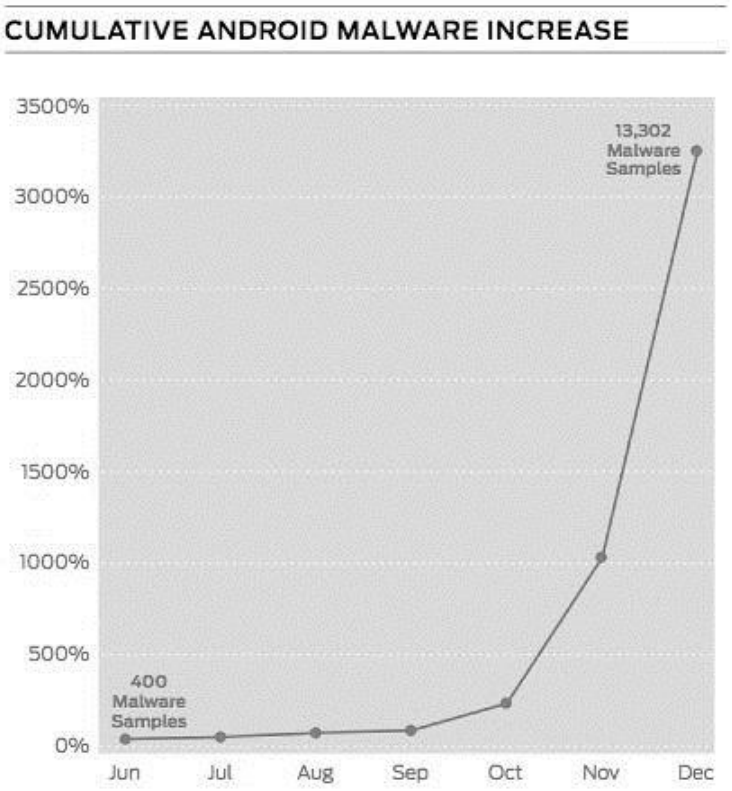
Research The Mobile Threatscape. Considering Google Android and Apple iOS are dominant consumer platforms and BYOD has compelling advantages for business, comparison of these operating systems would be necessary. According to the Ponemon study 59 percent of organizations allow their employees to use their personally owned mobile devices (BYOD) at work, as seen in Figure 1 below (Ponemon Institute, 2013).

Figure 1: Percentage of employees using personally owned mobile devices in the workplace (Ponemon Institute, 2013).



Reisinger (2012) noted that the overall security of Android devices is a major threat and most security experts now view Android as the most attacked mobile OS, the easiest to target because of its open architecture. According to Juniper Networks "2011 Internet Security Threat Intelligence Report" (Juniper, 2012) more than 25 times more Android Malware was identified in 2011 than in 2010. In fact in 2011 alone, malware targeting of the Google's Android platform rose 3,325 percent, Figure 2 below (Juniper Networks, 2012).

Figure 2: Cumulative Android Malware Increase in 2011. (Juniper Networks, 2012).



Develop a Mobile Device Security Plan. Using a written plan, organizations should clearly indicate how company and personal mobile devices are allowed to connect to corporate resources, with defined levels of permitted access. An MDM system (discussed below) should be deployed to secure, monitor, manage and support all mobile devices that connect to company resources. The organization's mobile risks, regulatory compliance issues, and governance issues should be identified, and appropriate security and access controls that should be in place. Document backups and disaster recovery plans, as well as remote wiping procedures, and mobile applications policies should be developed. End-users should be trained, based on device and/or operating systems as well as the acceptable use policies of the organization.

Deploy Mobile Device Management (MDM) system. This software secures, monitors, manages and supports all mobile devices on multiple operating systems. It usually includes over-the-air distribution of applications, data and configuration settings for all types of mobile devices. It can be made applicable to both company-owned and personal devices. An MDM system can also deploy and manage third-party or in-house mobile applications.

ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

Create Centralized Network Access Control for Mobile Devices. Develop mobile authentication practices that provide endpoint integrity checks and authorization of resources, e.g. VPN via Cisco IPSec, SSL VPN, SSL/TLS with X.509 certificates, WPA2 Enterprise with 802.1X, certificate-based authentication, and RSA SecurID or CRYPTOCard.

Mobile Device Authentication Policy. Establish strong device security policies to protect corporate information, e.g. strong passcodes, passcode expirations, passcode reuse history, maximum failed attempts, over-the-air passcode enforcement, progressive passcode timeout, etc.

Mobile Device Configuration and Restrictions. Devices should be configured using MDM systems with encrypted configuration profiles, e.g. CMS (Cryptographic Message Syntax) supporting 3DES and AES 128. Device restrictions determine which features users can access on the device, like passcode policies, cameras, or web-browsing restrictions.

Mobile Data security. Devices should include hardware encryption (like 256-bit AES) and data protection (unique device passcodes) to generate strong encryption keys. Additionally, remote wipe and local wipe features set by an MDM server are important.

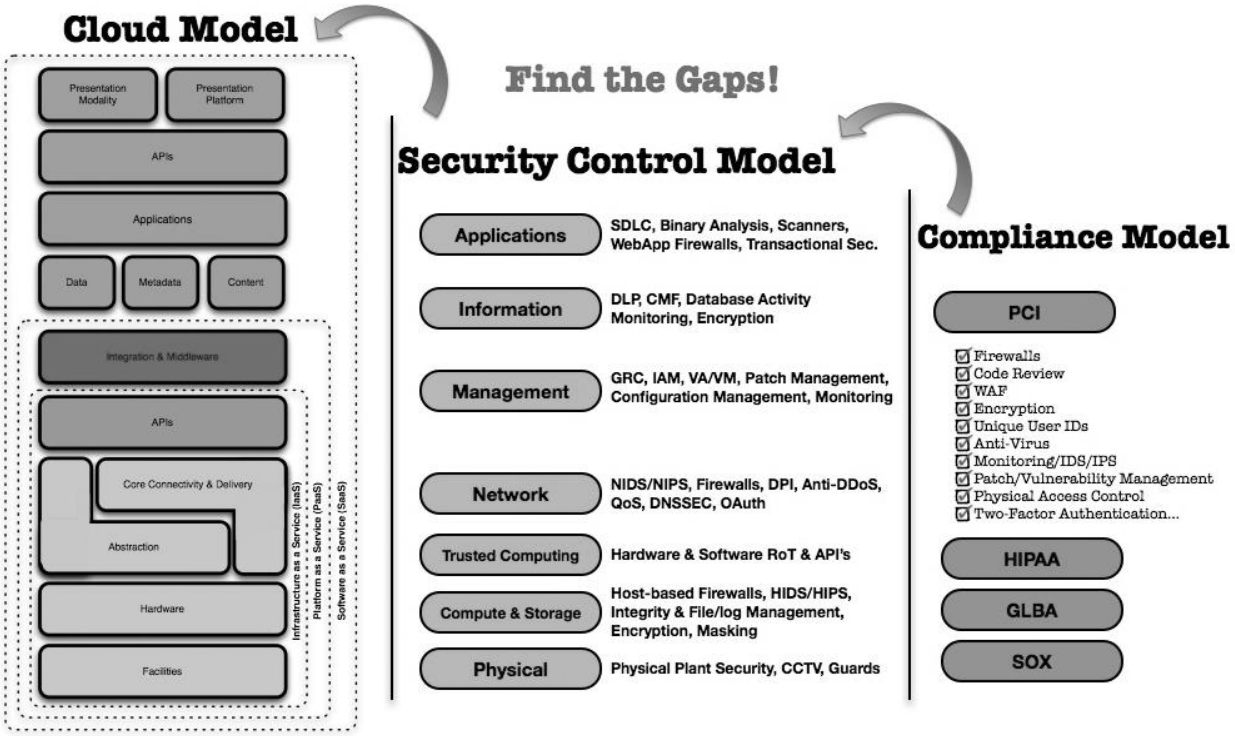
Mobile Application Security. Ensure runtime protection so that applications are "sandboxed" and cannot access data stored by other applications. Mandatory code signing - third-party applications that are signed by the developer using certificates issued by a device manufacturer, e.g. Google or Apple. Application encryption APIs that allow software developers to symmetrically encrypt data using, for example, AES or 3DES.

Regulatory Proficiency

Every organization in this study was required to comply with various regulatory agencies and auditors, both inside and outside the United States. The cloud systems used by these types of organizations present ongoing regulatory challenges, which have very high audit and data retention requirements. Despite the challenges, cloud computing can provide these organizations with regulatory compliance mechanisms that are scalable and available on-demand. This can be accomplished by performing a gap analysis and then mapping the chosen cloud service model to security controls and necessary compliance models. This process is illustrated in Figure 3.

ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

Figure 3: Mapping Cloud Model to Security Controls & Compliance Models (Cloud Security Alliance, 2011).



Improved Reliability and Access

Another major concern of organizations in this case study was the reliability of their key systems. Organizations preferred "high availability" (HA) cloud systems that provided "99.999%" availability of services (downtime less than 5.26 minutes per year). In order to achieve HA reliability organizations should design or require CSPs to provide: replicated servers across multiple zones, automatic failover or mirroring, and comprehensive disaster recovery.

CONCLUSIONS

Cloud adoption is poised to grow in the future. According to Gardner Research, by 2016 at least 50 percent of enterprise email users will rely primarily on a web browser, mobile tablet or mobile device instead of a desktop client. Gartner Research also predicts that by 2017 more than 50 percent of Global 1000 companies will store customer-sensitive data in the public cloud (Gartner Research, 2008; Savitz, 2011). Mullin (2009) notes drug companies like Pfizer, Eli Lilly & Co., Johnson & Johnson, and Genentech that have adopted cloud computing found benefits with large amounts of data storage, lower costs, and faster data processing. The companies in this study indicated they plan to continue using various types of cloud computing.

The main objective of our study was to examine the various aspects of cloud computing implementation as it relates to emerging life-sciences companies. We hoped to identify the

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

unique benefits, drawbacks, and various drivers of cloud implementation in these companies. Interpretations from this research indicate that small life-sciences companies found cloud computing very attractive in general. There were some relatively minor drawbacks, which could be mitigated with adequate planning and proper implementation.

Our study identified many common themes apparent in the companies that were studied. The advantages of cloud computing in the emerging biotech and pharmaceutical organizations studied were identified as: reduced cost and greater R&D speed, improved efficiency, enhanced agility, superior storage and data analysis, improved change management, superior collaboration and connectivity, enhanced security, faster drug discovery, better performance, appreciable regulatory proficiency, and much greater scalability and flexibility of IT resources. The study challenges of cloud computing in the emerging biotech and pharmaceutical organizations, as evident from the organizations studied were: concerns about security, confidentiality of corporate data, legal ramifications, cloud vendor lock-in, and lack of information systems control.

Security concerns and regulatory issues were identified as the predominant challenges of cloud computing in this study. However, with limited budgets and few, if any, onsite security professionals, the SMB life-science companies in this study considered the overall security and control features provided by CSPs as superior and more comprehensive than what could be achieved by their limited in-house staff. There would be cost efficiencies as well. However, it is clear that as cloud computing grows, CSPs must maintain the highest levels of security in order to retain this advantage and true business value for these organizations. The CSPs helped the companies satisfy their regulatory challenges. Regarding the regulatory aspect of cloud computing, the scenario is bleaker. Participants pointed out that a lack of clear-cut regulations regarding cloud computing from the regulatory agencies was a disincentive to further cloud adoption.

Companies in this study initially lacked the adequate computational infrastructure to meet their future needs. In order to gain such abilities they often partnered with larger academic institutions, biotech, or pharmaceutical companies. The participants and organizations in this study seemed to be ideal candidates for larger-scale participation in cloud computing. Those organizations that have embraced cloud computing were able to efficiently grow and more quickly build competitive advantages, while simultaneously reducing IT expenditures – no longer having to procure, maintain, and update systems or support all end-users.

Smaller organizations that lacked adequate computational or data management infrastructure were ideally poised to take advantage of cloud computing's pay-as-you-grow structure. The organizations in this study found cloud computing met their needs for voluminous internal computer power without additional IT overhead. The study also found that cloud computing offered a major business advantage to SMB life sciences companies – with its faster, cheaper, more scalable model – thereby helping these companies create a competitive parity with much larger organizations and at least a competitive advantage over their peers.

ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

FUTURE WORK

In the future, this research could be expanded to include significantly more participants and organizations of different types, beyond just life science companies and their business processes. That would help determine and uncover additional opportunities or challenges that cloud computing would pose to organizations. This research could also be extended to involve significantly more participants and companies to perhaps reveal comprehensive cloud taxonomy or enterprise architecture for other types of emerging organizations.

REFERENCES

Afgan, E., Baker, D., Coraor, N., Chapman, B., Nekrutenko, A., & Taylor, J. (2010). Galaxy CloudMan: Delivering cloud compute clusters. BMC Bioinformatics, 11 Suppl 12(Suppl 12), S4-S4. doi:10.1186/1471-2105-11-S12-S4

Barbadora, L. (2012, June 14). Rising Number of Small- to Mid-Sized Pharmaceutical Companies Switching to Cloud CRM. Retrieved from http://bloom.bg/YG1EIv

Bowers, L. (2011). Cloud Computing Efficiency. *Applied Clinical Trials,* 20(7), 45-46,48-51. Retrieved June 6, 2012, from *ProQuest Health and Medical Complete.* (Document ID: 2410599311).

Chidambaram, V. (2011, December). A Case Study on Cloud Computing: Genotypic Technology Puts its ERP on the Cloud. Retrieved from http://bit.ly/GMQzgg

Cloud Security Alliance. (2011). *Security guidance for critical areas of focus in cloud computing V3.0* [White paper]. Retrieved from https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf

Cognizant. (2010). Cognizanti - Next-Generation Services in a Reset Economy. Retrieved from http://cogniz.at/XLoXkR

Davies, K.. (2009, November). The 'C' Word. *Bio - IT World,* 8(6), 24,26,42. Retrieved June 5, 2012, from ProQuest Computing. (Document ID: 1955425591).

Deuby, S. (2013, July 10). What Are OAuth 2.0 and OpenID Connect? Retrieved from windowsitpro.com/identity-management/what-are-oauth-20-and-openid-connect

Do, L. H., Esteves, F. F., Karten, H. J., & Bier, E. (2010). Booly: A new data integration platform. *BMC Bioinformatics,* 11(1), 513-513. doi:10.1186/1471-2105-11-513

Gartner Research. (2008, July 15). Gartner Says Security Delivered as a Cloud-Based Service Will More Than Triple in Many Segments by 2013. Retrieved from http://gtnr.it/XTPupf

Gorban, A. (2012, June 28). Cloud Computing in Regulated Environments. Retrieved from http://bit.ly/Zj36zj

Heritage, T. (2012, March). Hosted Informatics: Bringing Cloud Computing Down to Earth with Bottom-line Benefits for Pharma. Retrieved from http://bit.ly/ZquDvC

Juniper Network**s**. (2012, February). 2011 Mobile Threats Report. Retrieved from www.juniper.net/us/en/local/pdf/additional-resources/jnpr-2011-mobile-threats-report.pdf

Knorr, E. and Gruman, G. (2009, December 23). What cloud computing really means. Retrieved from http://bit.ly/c2dL6P

Kubick, W. R. (2011). Are we ready to fly into the cloud? Cloud computing is now becoming possible for pharmaceutical companies, but still faces challenges. *Applied Clinical Trials,* 20(2), 28.

May, M.. (2010). Forecast calls for clouds over biological computing. Nature Medicine, 16(1), 6. Retrieved June 5, 2012, from Research Library. (Document ID: 1935297111).

McCarthy, A., (2012). Drug Discovery in the Clouds, Chemistry & *Biology*, 19(1), 27 January 2012, Pages 1-2, ISSN 1074-5521, 10.1016/j (http://www.sciencedirect.com/science/article/pii/S1074552112000233)

McCracken, G. (1998), *The Long Interview*, Newbury Park, CA: Sage Publications

Mullin, R. (2009, May 25). Chemical Engineering News - The New Computing Pioneers. Retrieved from http://bit.ly/WpJ1O

Nguyen, A. (2012, April 25). BT launches life sciences R&D cloud. Retrieved from http://bit.ly/Ib6sP4

Ponemon Institute. (2013, June). The Risk of Regulated Data on Mobile Devices. Retrieved from http://info.watchdox.com/rs/watchdox/images/WatchDoxWhite%20PaperFINAL2.pdf

Proffitt, A.. (2009, November). Pharma's Early Cloud Adopters. *Bio - IT World*, 8(6), 31-32. Retrieved June 6, 2012, from ProQuest Computing. (Document ID: 1955425641).

Qiu, J., Ekanayake, J., Gunarathne, T., Choi, J. Y., Bae, S., Li, H., ...(2010). Hybrid cloud and cluster computing paradigms for life science applications. *BMC Bioinformatics,* 11 Suppl 12(Suppl 12), S3-S3. doi:10.1186/1471-2105-11-S12-S3

Reisinger, D. (2012, April 4). Mobile and Wireless: Android Security Is a Major Threat: 10 Reasons Why. Retrieved from www.eweek.com/c/a/Mobile-and-Wireless/Android-Security-Is-a-Major-Threat-10-Reasons-Why-148798/

Sansom, C. (2010). Up in a cloud? *Nature Biotechnology*, 28(1), 13-15. doi:10.1038/nbt0110-13

Savitz, E. (2011, December 1). The Road Ahead: Gartner's Outlook for 2012 and Beyond. Retrieved from http://onforb.es/tD5cuG

  ISSN: 1543-5962-Printed Copy ISSN: 1941-6679-On-line Copy

Shurell, A.. (2010, November). LIFE SCIENCES JOINS THE CLOUD. Pharma, 6(6), 54,56,58. Retrieved June 5, 2012, from *ProQuest Health and Medical Complete*. (Document ID: 2249071871).

Sommer, T. (2013). Cloud Computing In Emerging Biotech and Pharmaceutical Companies. *Communications of the IIMA*, 13(3), 37-54.

Taylor, R. C. (2010). An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics. *BMC Bioinformatics*, 11 Suppl 12(Suppl 12), S1-S1. doi:10.1186/1471-2105-11-S12-S1

Vandeweyer, G., Reyniers, E., Wuyts, W., Rooms, L., & Kooy, R. F. (2011). CNV-WebStore: Online CNV analysis, storage and interpretation. *BMC Bioinformatics*, 12(1), 4-4. doi:10.1186/1471-2105-12-4

APPENDIX

Figure A-1: Survey Questions and Interview Protocol.

1)	What is your strategic role within your organization?
2)	What are the reasons behind your organization's use of Cloud Computing?
3)	Which service model does your organization currently utilize, based on this Cloud Computing taxonomy
4)	Which deployment model does your organization currently utilize, based on this Cloud Computing taxonomy?
5)	Who currently hosts and manages your cloud computing environment?
6)	Which IT services or applications, that support your business processes, have/would you migrate to Cloud Computing?
7)	Are you or would you be willing to outsource to multiple cloud computing providers?
8)	In your assessment of the feasibility and profitability of your cloud computing environment, what are the biggest advantages?
9)	In your assessment of the feasibility and profitability of your cloud computing environment, what are the biggest disadvantages?
10)	What are your main concerns regarding your organization's approach to Cloud Computing?

Figure A-2: OpenID Connect Protocol Suite.