Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Greco M. Johnson Date

Genesis: Laboratory Data Processing and Analytics System

By

Greco M. Johnson
Master of Public Health

Applied Public Health Informatics

_____

Laura M. Gaydos, PhD
Committee Chair

_____

Melissa Alperin, EdD, MPH, MCHES
Committee Member

Genesis: Laboratory Data Processing and Analytics System

By

Greco M. Johnson

B.S., Kennesaw State University
Emory University
2018

Thesis Committee Chair: Laura M. Gaydos, PhD.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Applied Public Health Informatics
2018

Abstract

Genesis: Laboratory Data Processing and Analytics System
By Greco Johnson

The number of emerging infectious disease threats continues to increase as society, technological, and environmental factors change over time. The introduction of unknown infectious diseases could negatively impact the health of communities around the world. As a result, a significant amount of attention should be placed on the ability of public health laboratories to improve efforts of processing, analyzing, and reporting infectious disease findings. This thesis document presents the recommendation and proposed solution for improving current state laboratory data processing and reporting within public health laboratories.

The Centers for Disease Control and Prevention (CDC) provides oversight for a number of domestic and global laboratory studies and investigations of infectious disease of unknown etiologies. The CDC "strives to improve public health and safety nationally and globally through investigating, monitoring, and controlling sickness, disability, and death caused by highly lethal viral, bacterial, prion, and related infections and diseases of unknown origin".[1] A key component to satisfying this organizational mission is to ensure a systematic approach exist when performing test on tissue specimens from humans and animals that are sick or deceased from unknown causes. The actions taken from these procedures allow laboratory test results to be analyzed and interpreted, which promote the monitoring of lethal diseases, improve disease diagnosis, prevention, and treatment and allow for emergency response preparedness.

The Genesis Laboratory Data Processing and Analytics System is an automated solution designed to enhance current state laboratory data processing procedures. In addition, Genesis will ensure the rapid return of laboratory test results and associated data to ensure any unknown infectious disease threats are limited in their impact. This cloud-based computing system will provide public health laboratories with an innovative, interoperable, and secure approach to processing laboratory test data, while satisfying any requirements to fully automate current state manual processes.

Genesis: Laboratory Data Processing and Analytics System

By

Greco Johnson

B.S., Kennesaw State University
Emory University
2018

Thesis Committee Chair: Laura M. Gaydos, PhD.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Applied Public Health Informatics
2018

Acknowledgements:

I would not be at this point in my education or life if it were not for the grace of GOD.

A very special thanks to my wife Janine. You have always been my biggest supporter and my loving inspiration. Your words have inspired me to push forward with the mission of helping others less fortunate than ourselves. I would also like to thank my children, Evan and Nicole for withstanding this process. Especially, during all of the times I was not totally focused on your own personal endeavors. Lastly, I want to thank the rest of my family and friends for their prayers and loving support throughout this process.

I would like to give a special thanks to my work family at the Public Health Informatics Institute and the Task Force for Global Health. You have given me the opportunity to truly understand what it takes to make difference and change the lives of people across the world.

Finally, thank you to the faculty and staff of the Executive MPH program for all of your support over the last 3 years.

Table of Contents

Background

Infectious disease outbreak has been characterized as the single most important public threat to the human population. The growing number of emerging infectious diseases are directly related to the increased number of deaths in humans. Infectious disease can have a social and economic impact on our global community. As disease surveillance activities take place and subsequent data has to be properly collected and identified. In addition, systems have to be in place to ensure the data collected during those surveillance activities is secure, faultless and available for transmission as requested across public health agencies.[1,2,3]

The CDC conducts laboratory studies of pathology laboratory test obtained from internal and external laboratories in order to properly identify and confirm cases of infectious disease in human specimens. In 2007, reporting laboratories estimated 6.3 million laboratory test[2] was performed for two disease characteristics (i.e. chlamydia and gonorrhea). DHCCP conducts test on a wide array of infectious disease categories. Current CDC policies require sponsored projects participants (e.g. international, federal, state and private laboratories) to share laboratory test findings so that additional internal analysis can be performed.

To this end, technology can be a driving force to improve the efforts of public health agencies to limit the impact of this public health threat. Enhancing the informatics capabilities of public health laboratories to identify known and unknown threats is a critical aspect in preventing human and non-human causalities of infectious disease. It is imperative that public health laboratories take advantage of every opportunity to improve data processing and statistical analysis procedures when identifying infectious disease. In addition, the methods for presenting the overall results of laboratory data can lead to improved delivery times. Improving the capabilities of public health laboratories can also lead to better policy making as the results of data is presented to policy makers. Corrective action can be taken to ensure work procedures are more than adequate. Lastly, global standards of operation can be established across all laboratories.[3,4]

Problem Statement

A major goal of the Trump Administration is to cut cost within the Public Health Sector of the Federal Government. “The fiscal year (FY) 2018 President’s Budget request for Centers for Disease Control and Prevention (CDC) and Agency for Toxic Substance and Disease Registries (ATSDR) includes a total funding level of $6,037,243,000 in discretionary budget authority and the Prevention and Public Health Fund (PPHF). This is an overall decrease of $1,222,431,000 below the FY 2017 Annualized Continuing Resolution (CR) level, which is a 17% reduction.”[13] As a result, the CDC is conducting a review of current laboratory testing cost and procedures due to pending budget cuts coming from the Executive Branch of Government. The goal of this review is to identify areas of improvement, which can assist in the agency in meeting the requirements defined by the budgetary constraints.

To address potential funding restrictions, the agency will review all operational activities including, current laboratory testing processing and procedures due to pending budget cuts defined by the Executive Branch of Government. A review of current laboratory procedures has determined a new more efficient process is needed. A majority of laboratory analysis and review procedures involve exhaustive manual review processes. Creating efficiencies in the review process includes, developing a public health informatics solution that will automate current state manual processes through the introduction of a laboratory information management system. As a result laboratory test submitted by CDC internal and external laboratories can be collected, analyzed, and interpreted in a timely, accurate and concise manner. This has the potential to lead to a reduction in overhead cost associated with current laboratory review and analysis procedures.

The current state process requires external laboratories to upload their findings to an electronic file storage application. Internal laboratory resources retrieve the file submissions and analyze the results using legacy analysis applications. Once lab results are obtained, it takes approximately 45 minutes to 1 hour from the initial data retrieval to make a final determination for a single interpreted result (e.g. Haemophilus influenzae). In addition, it was determined that all data reviews are manually performed by lab scientist. The timetable required to deliver on interpreted results is also impacted by other

laboratory work factors. This can extend the time required to perform disease interpretation analysis and interpretation. On average, it takes 7 – 21 days[3] for an internal CDC laboratory to perform its analysis of specimens, review external laboratory test and define an interpreted result for a particular lab test. Introducing an automated system will cut the final lab analysis and interpretation time to 15 minutes in total. Additionally, the automated system will reduce labor cost and streamline work processes.[6,7]

Public Health and Business Impact

A major goal of the current Executive Office Administration is to cut cost within the Public Health Sector of the Federal Government. Introducing the Genesis system as the primary tool for processing and analyzing laboratory test data will reduce agency expenses significantly by cutting personnel cost and eliminate undesired work processes. Additionally, the system will improve the quality, timeliness, and the delivery of laboratory result data. The Genesis system will also support efforts to better deliver lab reports, which will allow agency leadership to better understand where areas of improvement can be achieved in regards to policy making, training and workforce development. The system will allow the agency to retire legacy systems and establish a single authoritative source for laboratory test result analysis and data storage. Lastly, the system is designed to be interoperable with leading laboratory analysis systems as to ensure low cost of implementation.

Methods

The Genesis Laboratory Data Processing and Analytics system will support internal and external operational procedures associated with laboratory specimen processing, analysis, and interpretation. The system will enhance the process by which current internal CDC and external laboratory sites obtain, share, and store laboratory test results. Genesis has the capability to incorporate and integrate current laboratory processing technology standards, which ensure technology interoperability across laboratory systems. Developing a fully integrated laboratory information system will expedite the time needed to gather critical data need to combat public health infectious disease threats.

Introducing data visualizations as a means to communicate test results ensures information can be retrieved real time. As lab results are analyzed in the Genesis, raw statistical data can be transformed into visual diagrams. In many instances, key stakeholders such as laboratory managers and public health leadership are informed of critical data points and decisions can be made to improve laboratory testing efforts.

The ability to gather accurate and timely laboratory test results is a critical component associated with the combating public health threats. Currently, the time taken to gather, analyze, and interpret the results of laboratory test submitted by CDC and external laboratories is unsatisfactory. The increase in effort to make lab result determinations available in a timely manner potentially prohibits the agency from meeting mandates handed down by the Executive Office to decrease cost for all federal agencies.

Funding:

The timeline to develop and implement the Genesis Laboratory Data Processing and Analytics tool as the primary system for laboratory data collection, processing, analysis, and data dissemination is 9 – 12 months. A second, phase of development or system maintenance will take an additional 9 – 12 months. The project budget is estimated to be $1.5 million. Included in this budget are the necessary technological systems, software, and development services, which are required to successfully implement the system.

Scope and Objectives:

Improve the current process by which the CDC collects, analyzes and sets determination for laboratory specimen test results.

The objectives of the Genesis System are as follows:

- Provide external laboratories with an automated system to submit electronic lab results directly to the CDC.
- Implement a systematic algorithm to interpret the results of analyzed laboratory test results submitted by external and internal laboratories.

- Store analyzed and interpreted lab test result data for the life expectancy of the program.
- Develop a system component that is cost effective and efficiently illustrates and reports the results of system laboratory test data.

The inclusive boundaries of the project include:

- Fulfillment of CDC sponsored laboratory project objectives
- A single authoritative source of laboratory information
- Improved processes and procedures for interpreting laboratory test results.

The exclusive boundaries of the project include:

- The system will not facilitate non-CDC specimen analysis
- The system will not integrate with external laboratory systems

Lifecycle Component and Resource:

The following information technology lifecycle components are needed to complete the project. In addition, these components are supported by the applicable project resources.

Lifecycle Component	**Resource**
Planning	Project Manager
Requirements Development	Business Analyst
Application/System Development	Application Developer, Database Administrator, Network Administrator
Testing and Deploy	QA Analyst, Application Developer and Network Administrator
Maintenance	All

Major Deliverables:

The following table represents the major deliverables that are the result of the implementing the Genesis system.

Major Deliverable	Deliverable Description
Project Charter Approval	Stakeholder and Business Steward Approval of Project
Vendor Contracts	Contractual Agreements between CDC and External Vendors
Project Plan	Documented Project Timelines and Milestones
End User Web Portal	End User Web Interface for Submitting Lab Specimen Results
Laboratory Information System	Enterprise System for Lab Data Collection, Processing, and Storage
Data Visualizations	Data Visualizations and Reports for End-User Consumption

Project Timeline:

The following tables represent the high-level and descriptive/detailed requirements, which are necessary to ensure the Genesis system, services and results ensure the project objectives are satisfied. These requirements were collected and will provide guidance for the development of the Genesis Data Laboratory and Analytics System.

Action or Step to be Completed Method for Completion	Timeline	Milestones	Person(s) Responsible	Status/ Completion Date
Project Initiation (Stakeholder Identification, Business Case,	Weeks 1-3	1. Business Case Documented 2. Charter Documented	Project Manager,	February 2018

Charter, Communications Plan)		3. Communications Plan Completed	Business Analyst(s)	
Activity/Task Evaluation and Estimation	Week 4-5	1. Project Plan Developed	Project Manager	February – March 2018
Requirements Elicitation	Weeks 6-10	1. Stakeholder/End User Requirement Meetings 2. Requirements Documented 3. Requirements Signoff	Business Analyst(s), Project Manager	March – April 2018
Application Development	Weeks 11-28	1. Development Environments Created 2. Sprints 1- 8 Completed	Application Developer(s), Database Administrator(s), Business Analyst(s)	April – July 2018
Unit Testing	Weeks 13-30	1. Test Cases and Scripts Documented 2. Sprint Testing Complete	Quality Assurance Analyst(s), Developer(s)	May – August 2018
Defect Development/Testing	Weeks 30-31	1. Defects fixed 2. Internal Testing Sign-off	Application Developer(s), Database Administrator(s), Quality Assurance Analyst(s)	August 2018
System Documentation	Weeks 32-35	1. System and User Documents Completed 2. Stakeholder Sign-off	Business Analyst(s)	September 2018
User Acceptance Testing	Weeks 35-38	1. End-user Module Testing Complete 2. Stakeholder Sign-off	Stakeholders	October 2018
Application Release/Deployment	Week 39	1. “Go Live”	Application Development Team	November 2018

Project Requirements:

The following tables represent the high-level and descriptive/detailed requirements, which are necessary to ensure the Genesis product, services and results ensure the project objectives are satisfied. These requirements were collected and will provide guidance for the development of the Genesis Data Laboratory and Analytics System.

High-Level Project Requirements:

Requirements #	Requirements Description
1.	Development of Cloud-based Enterprise System Architecture
2.	Laboratory Data Processing Algorithm
3.	Data Integration and Warehousing Architecture
4.	Laboratorian Web-portal User Interface
5.	Enterprise Information Security Policies
6.	Develop Data Visualizations for Laboratory Test Submissions
7.	Data Visualization Reports
8.	Data Access and Use Policies
9.	Systems Documentation

Descriptive/Detailed Project Requirements:

1. Enterprise System Architecture Requirements
1.1. Identification of Cloud-based Service Provider
1.2. Identification of Application Development Architecture
1.3. Identification of Database Platform
1.4. Network and Systems Communication Architecture Development
1.5. Data Visualization Environment Scan

2. Laboratory Data Processing Algorithm
2.1. Develop Java Based Computing Algorithm
2.2. Identify and Implement All Laboratory Pathogen Test Result Patterns
2.3. Develop Lab Data Export Format

3. Data Integration and Warehousing Architecture Requirements
3.1. Establish Data Store Procedures for Laboratory Test Results
3.2. Develop Data Sanitation Algorithm for Data Visualization Tool
3.3. Define Data Retention and Long-term Storage Plans and Procedures

4. Laboratorian Web-portal User Interface Requirements
4.1. Develop Laboratorian Web Portal Interface
4.2. Create Lab/Site Specific System Access Protocols and Properties

5. Enterprise **Information Security Policy Requirements**
5.1. Identification of Cloud-based Service Provider
5.2. Identification of Application Development Architecture
5.3. Identification of Database Platform
5.4. Network and Systems Communication Architecture Development

6. Develop Data Visualizations for Laboratory Test Submissions
6.1. Create/Design a data visualization application that would allow CDC internal and external laboratory partners to analyze and review overall laboratory test findings.
6.2. Grant system access to designated system users (faculty) across the CDC and external laboratory partners
6.3. Ensure data visualization tool provides the capability to generate multiple data representations for queried system data.

7. Data Visualization Reports
7.1. Provide system users with the capability to create/view multiple system reports within the data visualization tool.
7.2. Create reports to identify lab test performed by sites.
7.3. Create reports to identify the total number of test performed by type.
7.4. Create reports to identify lab extraction fails

8. Data Access Use and Request Requirements
8.1. Identification of Cloud-based Service Provider
8.2. Identification of Application Development Architecture
8.3. Identification of Database Platform
8.4. Network and Systems Communication Architecture Development

9. Systems Documentation Requirements
9.1. Development of System Architecture Documentation
9.2. Development of System Requirements
9.3. Development of System Data Specifications
9.4. Development of Quality Assurance Test Cases and Scenarios
9.5. Development of End-User Data Use Agreements
9.6. Development of End-User Training Guides

Business Process 1:

Tier 1: Business Area – Tier 2: Business Category	
Tier 3: Laboratory Test Result Analysis and Interpretation	
Description	Lab Test Data Interpretation
Trigger Event	Lab test results submitted by the CDC and external laboratories
Result	The analysis algorithm processes the data analyzed results and makes a final determination for specimen test.
Business Process Step	External lab data submission to the web portal. Internal CDC lab data submission to the web portal
Shared Data	Data stored and retrieved from a centralized database.
Predecessor	Lab specimens analyzed by laboratories using analysis software.
Successor	Laboratory Data Processing and Analytics System writes lab results to the central database.
Constraints	Data submission is in incorrect format.
Failures	Data submitters alerted for failed processing request.
Performance Measure	Time to analyze and return interpreted results.

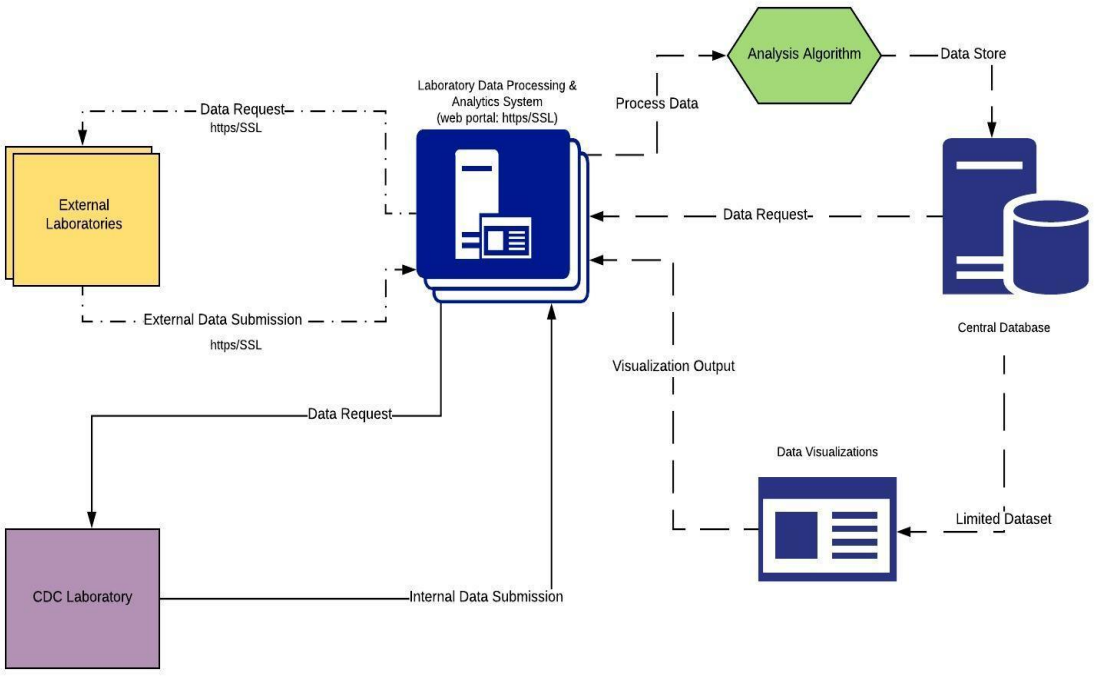
Business Process 2:

Tier 1: Business Area – Tier 2: Business Category	
Tier 3: Health Information Exchange	
Description	Lab Test Data Request
Trigger Event	Internal CDC or external laboratories request final laboratory datasets.
Result	The laboratory information system provides an electronic report of the final lab test results.
Business Process Step	External lab data request lab results from the web portal. Internal CDC lab data request lab results from the web portal.
Shared Data	Data retrieved from the centralized database.
Predecessor	Lab test results have been interpreted in the laboratory information system.
Successor	Final laboratory test results presented in the web portal as a pdf.
Constraints	System error interpreting lab data submissions.

Failures	Data requestor alerted for failed data request.
Performance Measure	Time to report interpreted results decreased.

Information Flow Diagram:

The current information flow and data distribution for the Genesis Laboratory Data Processing and Analytics System. The proposed system must process external and internal data submissions from laboratories using a secure web portal. The laboratory information system will also analyze and interpret data submissions. The analyzed data will be stored in the systems central database. Any request from external and internal laboratories will be processed through the web portal. Analyzed data records captured in the system will be conjoined into a limited dataset and exported into the data visualization tool for consumption by system end-users.

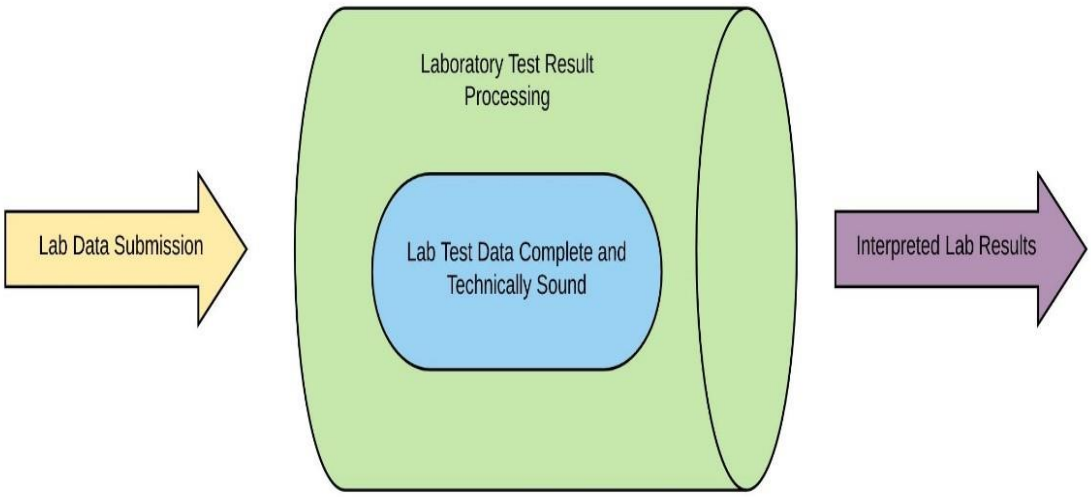


MITA Process Model:

The Genesis Laboratory Data Processing and Analytics System MITA Diagram is initiated by an event or trigger. The trigger is a submission of lab results from internal CDC or external partner laboratory.

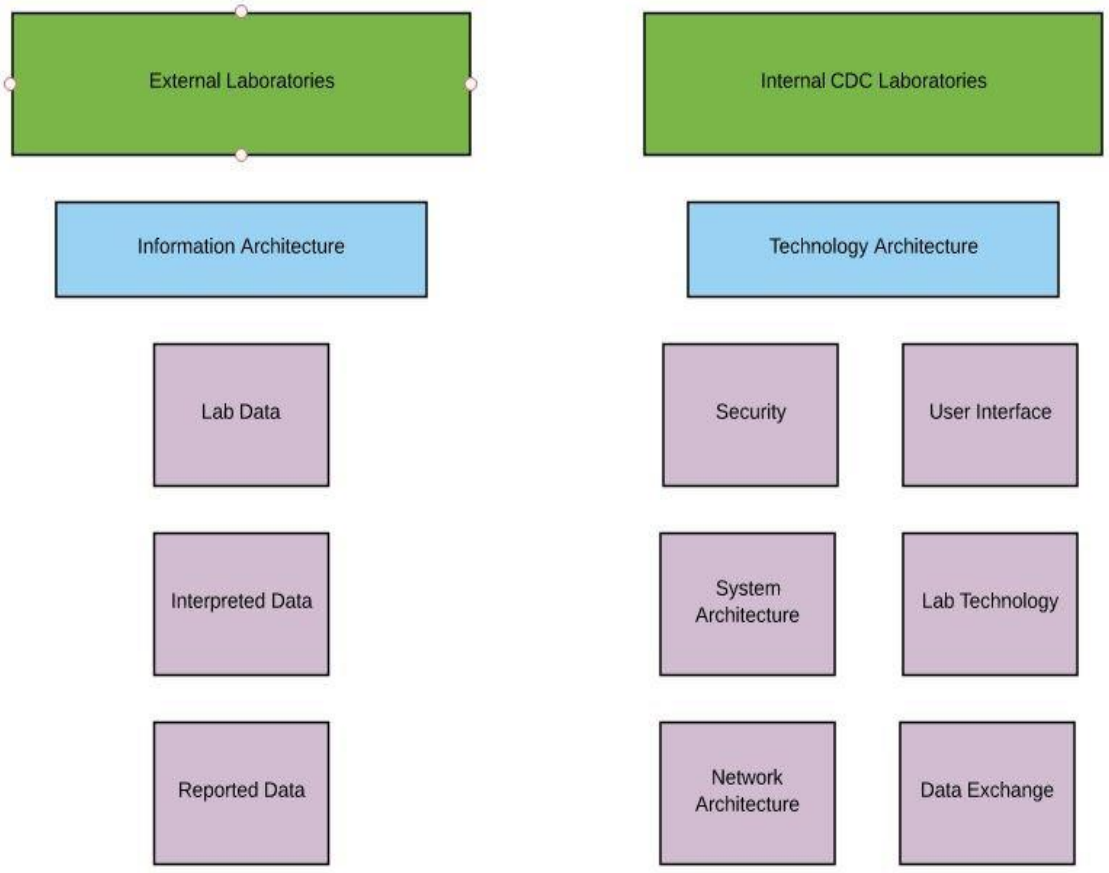
The system business rules requires lab data to be complete for all associated specimen test required or identified in the data submission. The submitted data should include an export of the analyzed and unanalyzed versions of the laboratory analysis software files and an exported excel file. These file formats are needed for the laboratory information systems' algorithm to return interpreted results for the combined data set.

Once the data is received and the interpretation process runs on the associated specimen results; the final determination is captured in the system. The results of the final test determination can be queried from all associated laboratories. The results will be made available in a web and pdf formatted report.



Business Architecture:

Internal CDC or external partner laboratories submit lab test result data through the secure laboratory information system web portal using https and SSL. The submitted data is de-identified. It is analyzed, interpreted, and stored by the Genesis Laboratory Data Processing and Analytics System into the central database. The Internal CDC and external partner laboratories request the final analyzed test results. Final laboratory test results can only be requested from persons with accounts stored on the laboratory information system. Only laboratory data from approved CDC entities and partners are allowed on the system. All data transactions must take place through the secure web portal.



Technology Architecture:

The Genesis Laboratory and Data Processing and Analytics system must support interoperability with only approved 3rd party laboratory analysis technology solutions. In addition, the system must support the exchange of data electronically. A user-friendly interface is required for end user interaction with the system. The system must support the latest advancements in system and network architecture. The system must provide confidentiality, integrity, and availability.

Service Access and Delivery			
Access Channels	**Delivery Channels**	**Service Requirements**	**Service Transport**
Web Browser	Internet	Legislative / Compliance	Supporting Network Services
Collaboration Communications		Authentication/Single Sign-on	Service Transport
	VPN	Hosting	

Service Platform and Infrastructure				
Support Platform	**Delivery Servers**	**Hardware / Infrastructure**	**Software Engineering**	**Database / Storage**
Independent Platform	Web Servers	Servers / Computers	Integrated Development Environment (IDE)	Database
	Application Servers	Embedded Technology Devices	Software Configuration Management	Storage
	Portal Servers	Peripherals	Test Management	

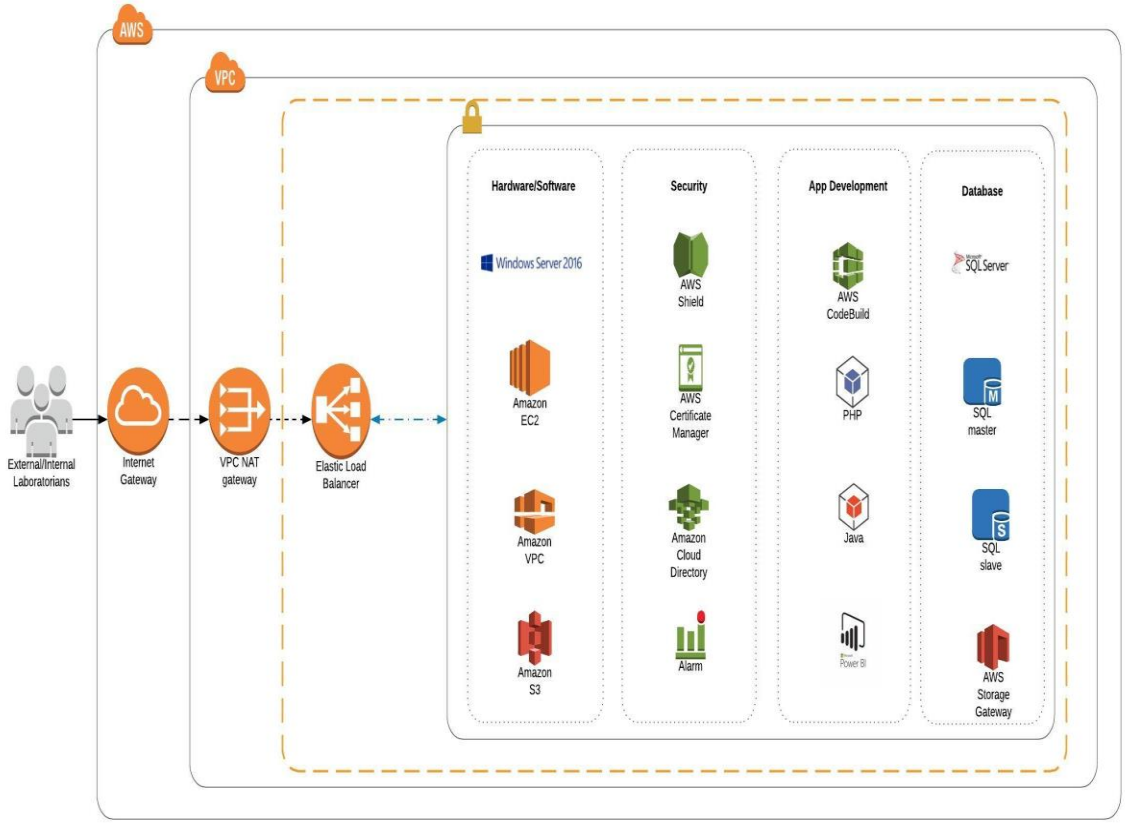
		Wide Area Network (WAN)	Modeling	
		Local Area Network (LAN)		
		Network Devices / Standards		

Component Framework				
Security	**User Presentation / Interface**	**Business Logic**	**Data Management**	**Data Interchange**
Certificates / Digital Signatures	Dynamic Server-Side Display	Platform Independent Technologies	Database Connectivity	Data Exchange
Supporting Security Services	Content Rendering		Reporting and Analysis	

Service Interface and Integration		
Integration	**Interoperability**	**Interface**
Middleware	Data Format / Classification	Service Discovery
Enterprise Application Integration	Data Types / Validation	
	Data Transformation	

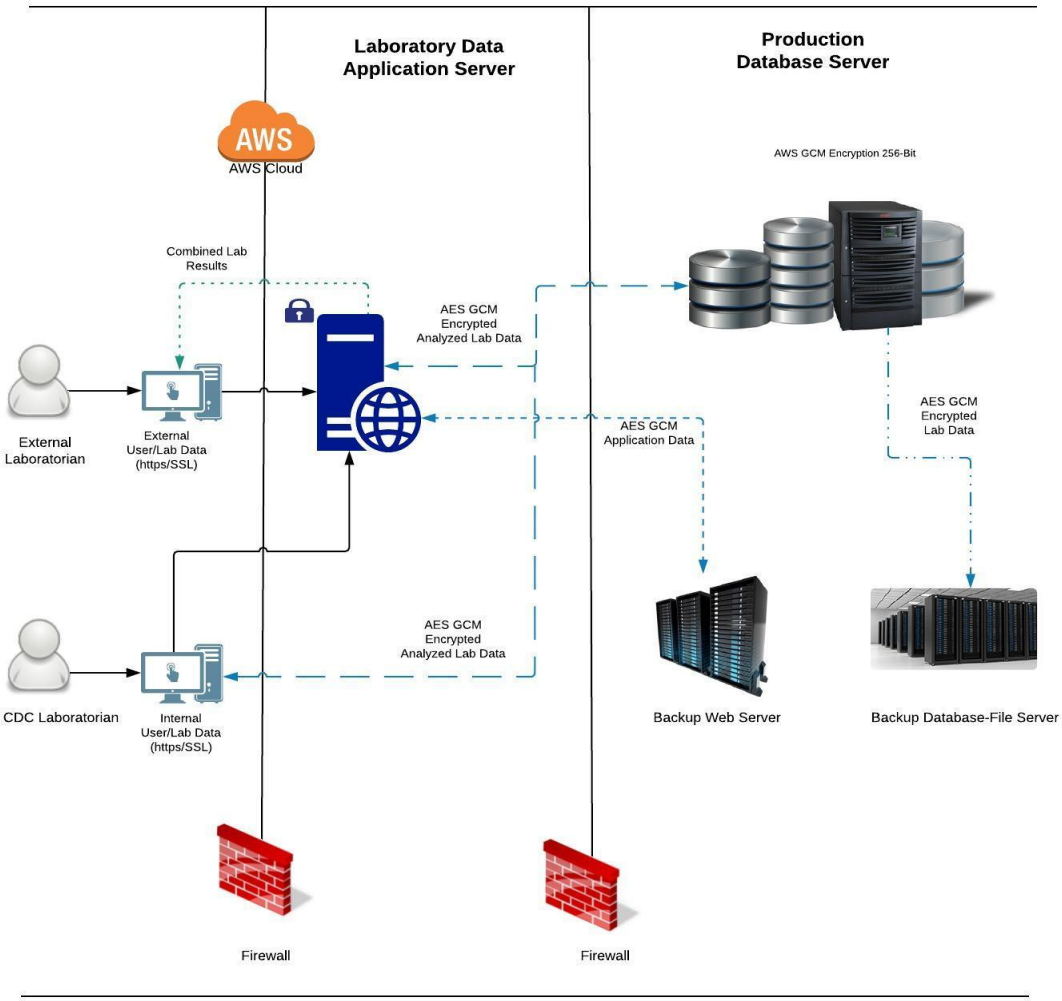
Enterprise Architecture Diagram:

The Genesis enterprise architecture is an AWS cloud-based environment. This diagram highlights key system components including the Hardware/Software, Security, Application Development and Database architectures. The hardware/software layer consists of the Amazon EC2, VPC, S3 platforms. Routing for this architecture component is maintained by a router. The platform runs on a Windows Server 2016 environment. The security layer consists of AWS shield and certificate manager tools. Amazon cloud directory is used to manage system personas. The application development layer runs AWS code build. Application development languages are PHP and Java. Power BI is used to develop analytic dashboards. The database platform is SQL Server. The data redundancy is managed by SQL slave and AWS storage gateways.



Network Security Diagram:

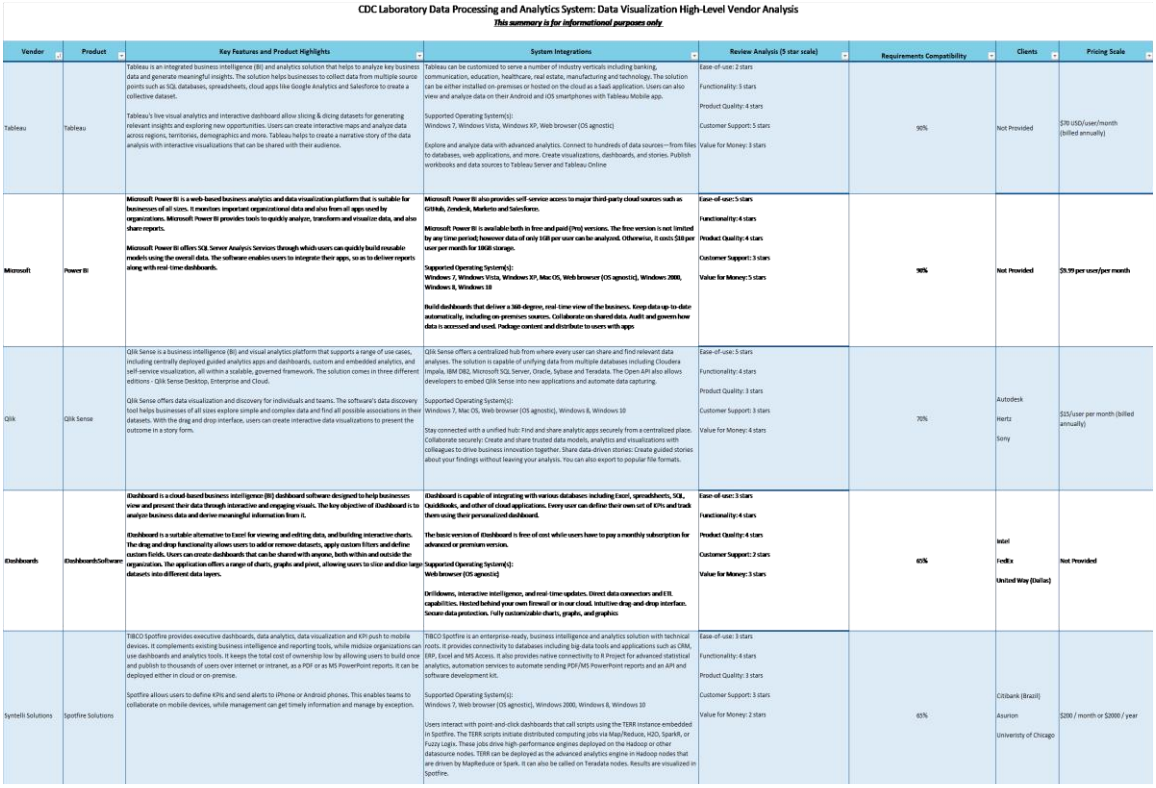
Data security protocols required by the National Health Institute, HIPAA, and HI-TECH are supported by the Genesis system. Regulatory guidelines are set by the National Institute of Technology Standards (NIST). The security protocols meet the minimum security requirements for Federal Information and Information Systems (e.g. Federal Information Processing Standards; FIPS-200)7. Laboratory test data is transmitted using Secure Socket Layer (SSL). The instances of those transmissions are via the internet https protocol. Data inside of the registry web services infrastructure are encrypted using AWS AES GCM. The encryption standard for all lab data within the cloud infrastructure is 256 AES GCM.



Environmental Scan Matrices:

The environmental scan below depicts the results returned from a review of the available data visualization tools, the service offerings of the vendors, and how closely the tools met the required system requirements.[17, 18, 19, 20, 21]

Note: Please click on the image to display the full environmental scan matrices.



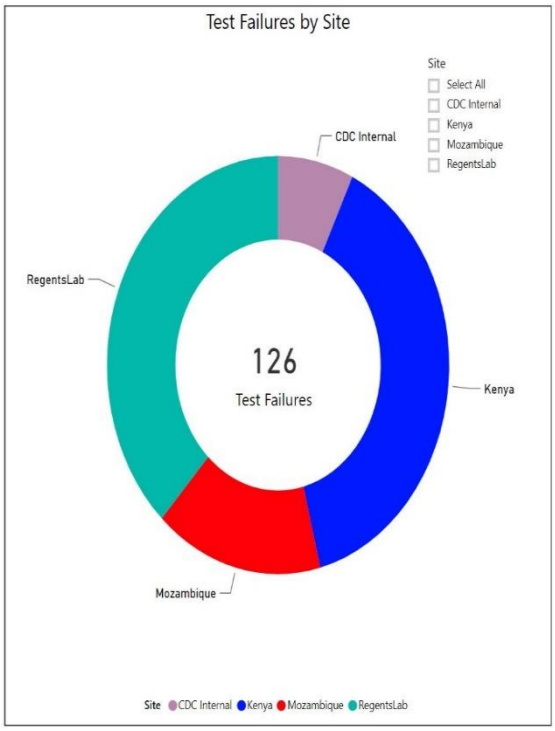
CDC Laboratory Data Processing and Analytics System: Data Visualization High-Level Vendor Analysis

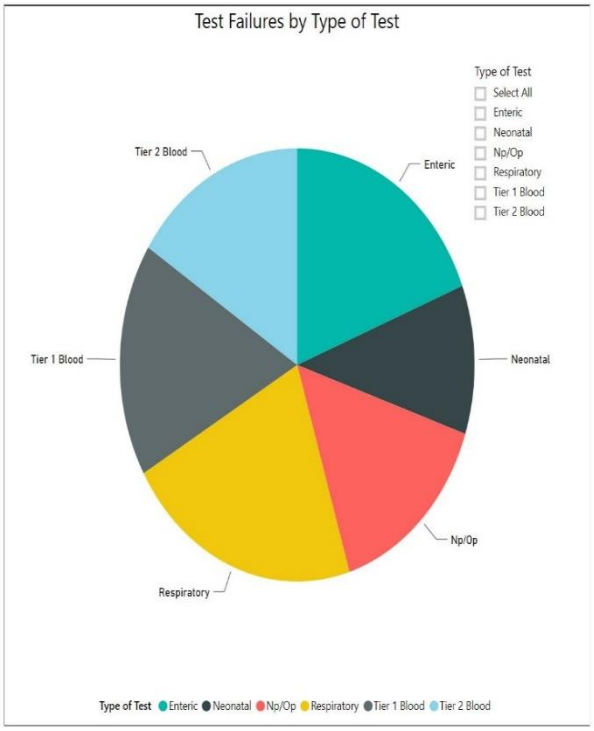
This summary is for informational purposes only

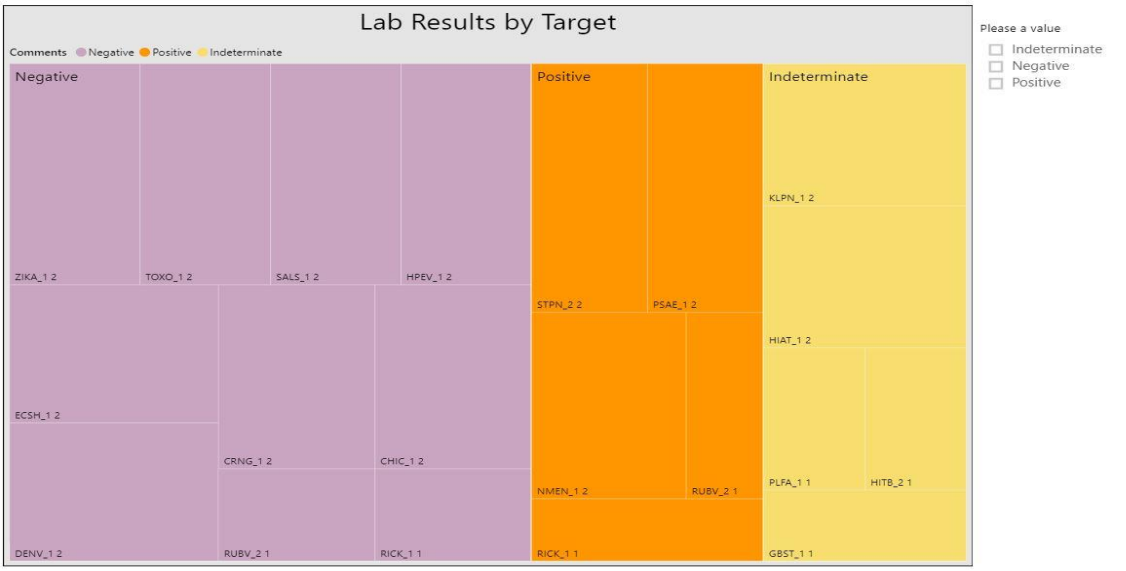
Vendor	Product	Key Features and Product Highlights	System Integrations	Review Analysis (5 star scale)	Requirements Compatibility	Clients	Pricing Scale
Tableau	Tableau	Tableau is an integrated business intelligence (BI) and analytics solution that helps to analyze key business data and generate meaningful insights. The solution helps businesses to collect data from multiple source points such as SQL databases, spreadsheets, cloud apps like Google Analytics and Salesforce to create a collective dataset. Tableau's live visual analytics and interactive dashboard allow slicing & dicing datasets for generating relevant insights and exploring new opportunities. Users can create interactive maps and analyze data across regions, territories, demographics and more. Tableau helps to create a narrative story of the data analysis with interactive visualizations that can be shared with their audience.	Tableau can be customized to serve a number of industry verticals including banking, communication, education, healthcare, real estate, manufacturing and technology. The solution can be either installed on-premises or hosted on the cloud as a SaaS application. Users can also view and analyze data on their Android and iOS smartphones with Tableau Mobile app. Supported Operating System(s): Windows 7, Windows Vista, Windows XP, Web browser (OS agnostic) Explore and analyze data with advanced analytics. Connect to hundreds of data sources—from files to databases, web applications, and more. Create visualizations, dashboards, and stories. Publish workbooks and data sources to Tableau Server and Tableau Online	Ease-of-use: 2 stars Functionality: 5 stars Product Quality: 4 stars Customer Support: 5 stars Value for Money: 3 stars	90%	Not Provided	$70 USD/user/month (billed annually)
Microsoft	**Power BI**	**Microsoft Power BI is a web-based business analytics and data visualization platform that is suitable for businesses of all sizes. It monitors important organizational data and also from all apps used by organizations. Microsoft Power BI provides tools to quickly analyze, transform and visualize data, and also share reports.** **Microsoft Power BI offers SQL Server Analysis Services through which users can quickly build reusable models using the overall data. The software enables users to integrate their apps, so as to deliver reports along with real-time dashboards.**	**Microsoft Power BI also provides self-service access to major third-party cloud sources such as GitHub, Zendesk, Marketo and Salesforce.** **Microsoft Power BI is available both in free and paid (Pro) versions. The free version is not limited by any time period; however data of only 1GB per user can be analyzed. Otherwise, it costs $10 per user per month for 10GB storage.** **Supported Operating System(s):** **Windows 7, Windows Vista, Windows XP, Mac OS, Web browser (OS agnostic), Windows 2000, Windows 8, Windows 10** **Build dashboards that deliver a 360-degree, real-time view of the business. Keep data up-to-date automatically, including on-premises sources. Collaborate on shared data. Audit and govern how data is accessed and used. Package content and distribute to users with apps**	**Ease-of-use: 5 stars** **Functionality: 4 stars** **Product Quality: 4 stars** **Customer Support: 3 stars** **Value for Money: 5 stars**	**98%**	**Not Provided**	**$9.99 per user/per month**
Qlik	Qlik Sense	Qlik Sense is a business intelligence (BI) and visual analytics platform that supports a range of use cases, including centrally deployed guided analytics apps and dashboards, custom and embedded analytics, and self-service visualization, all within a scalable, governed framework. The solution comes in three different editions - Qlik Sense Desktop, Enterprise and Cloud. Qlik Sense offers data visualization and discovery for individuals and teams. The software's data discovery tool helps businesses of all sizes explore simple and complex data and find all possible associations in their datasets. With the drag and drop interface, users can create interactive data visualizations to present the outcome in a story form.	Qlik Sense offers a centralized hub from where every user can share and find relevant data analyses. The solution is capable of unifying data from multiple databases including Cloudera Impala, IBM DB2, Microsoft SQL Server, Oracle, Sybase and Teradata. The Open API also allows developers to embed Qlik Sense into new applications and automate data capturing. Supported Operating System(s): Windows 7, Mac OS, Web browser (OS agnostic), Windows 8, Windows 10 Stay connected with a unified hub: Find and share analytic apps securely from a centralized place. Collaborate securely: Create and share trusted data models, analytics and visualizations with colleagues to drive business innovation together. Share data-driven stories: Create guided stories about your findings without leaving your analysis. You can also export to popular file formats.	Ease-of-use: 5 stars Functionality: 4 stars Product Quality: 3 stars Customer Support: 3 stars Value for Money: 4 stars	70%	Autodesk Hertz Sony	$15/user per month (billed annually)
iDashboards	**iDashboardsSoftware**	**iDashboard is a cloud-based business intelligence (BI) dashboard software designed to help businesses view and present their data through interactive and engaging visuals. The key objective of iDashboard is to analyze business data and derive meaningful information from it.** **iDashboard is a suitable alternative to Excel for viewing and editing data, and building interactive charts. The drag and drop functionality allows users to add or remove datasets, apply custom filters and define custom fields. Users can create dashboards that can be shared with anyone, both within and outside the organization. The application offers a range of charts, graphs and pivot, allowing users to slice and dice large datasets into different data layers.**	**iDashboard is capable of integrating with various databases including Excel, spreadsheets, SQL, QuickBooks, and other of cloud applications. Every user can define their own set of KPIs and track them using their personalized dashboard.** **The basic version of iDashboard is free of cost while users have to pay a monthly subscription for advanced or premium version.** **Supported Operating System(s):** **Web browser (OS agnostic)** **Drilldowns, interactive intelligence, and real-time updates. Direct data connectors and ETL capabilities. Hosted behind your own firewall or in our cloud. Intuitive drag-and-drop interface. Secure data protection. Fully customizable charts, graphs, and graphics**	**Ease-of-use: 3 stars** **Functionality: 4 stars** **Product Quality: 4 stars** **Customer Support: 2 stars** **Value for Money: 3 stars**	**65%**	**Intel** **FedEx** **United Way (Dallas)**	**Not Provided**
Syntelli Solutions	Spotfire Solutions	TIBCO Spotfire provides executive dashboards, data analytics, data visualization and KPI push to mobile devices. It complements existing business intelligence and reporting tools, while midsize organizations can use dashboards and analytics tools. It keeps the total cost of ownership low by allowing users to build once and publish to thousands of users over internet or intranet, as a PDF or as MS PowerPoint reports. It can be deployed either in cloud or on-premise. Spotfire allows users to define KPIs and send alerts to iPhone or Android phones. This enables teams to collaborate on mobile devices, while management can get timely information and manage by exception.	TIBCO Spotfire is an enterprise-ready, business intelligence and analytics solution with technical roots. It provides connectivity to databases including big-data tools and applications such as CRM, ERP, Excel and MS Access. It also provides native connectivity to R Project for advanced statistical analytics, automation services to automate sending PDF/MS PowerPoint reports and an API and software development kit. Supported Operating System(s): Windows 7, Web browser (OS agnostic), Windows 2000, Windows 8, Windows 10 Users interact with point-and-click dashboards that call scripts using the TERR instance embedded in Spotfire. The TERR scripts initiate distributed computing jobs via Map/Reduce, H2O, SparkR, or Fuzzy Logix. These jobs drive high-performance engines deployed on the Hadoop or other datasource nodes. TERR can be deployed as the advanced analytics engine in Hadoop nodes that are driven by MapReduce or Spark. It can also be called on Teradata nodes. Results are visualized in Spotfire.	Ease-of-use: 3 stars Functionality: 4 stars Product Quality: 3 stars Customer Support: 3 stars Value for Money: 2 stars	65%	Citibank (Brazil) Asurion University of Chicago	$200 / month or $2000 / year

Data Visualization Output(s):

The data visualizations below depict the results or output of analyzed data from the Genesis Laboratory Data Processing and Analytics system. As laboratory results are uploaded into the system, the system algorithm runs and performs data interpretations. The data visualization tool, in this instance Power BI, queries the data from the collection of test results in the system database and outputs a visualization that shows test results by type of test and the site conducting the test.







Conclusion:

The recommendation for addressing the issues related to the current state of the laboratory test procedures, processes, and looming budgetary cuts is for the CDC/DCHPP to implement the customized laboratory data processing and analytics tool. Implementing Genesis will improve current state processes and procedures associated with the collection, analysis, and interpretation of laboratory specimen test for CDC internal and external laboratory partners. The cost of implementing this system will be minimal, as the ROI for the project outweigh the long-term cost associated with man hours required to conduct current laboratory test procedures. In addition, the investment will ensure budgetary goals defined by the Executive Office will be satisfied.

In addition, implementing the Genesis system as described simplifies the process of analyzing and disseminating current laboratory test result data in the laboratory information system and outputting that data in a visual format. The proposed tool will use "Business Intelligence to transform raw data into smart information." Implementing a data visualization tool will further enhance current efforts of the CDC to strategically analyze, report, and manage internal and external laboratory workforce activities. The introduction of the Genesis system supports ongoing efforts to make better decisions concerning public health initiatives when combating known or unknown infectious disease threats. [6, 7, 8, 9, 10, 13]

References:

[1] Georgsson, Mattias & Staggers, Nancy. (2016). *Quantifying usability: An evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics*. Journal of the American Medical Informatics Association : JAMIA. 23. 5-11. 10.1093/jamia/ocv099.

[2] Theparee, T., et al. (2018). *Total Laboratory Automation and Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry Improve Turnaround Times in the Clinical Microbiology Laboratory: a Retrospective Analysis*. Journal of Clinical Microbiology 56(1).

[3] Hawkins, R. (2007). Laboratory turnaround time. *The Clinical biochemist*. Reviews, 28(4), 179-94.

[4] B. Reeder, et al. (2012). *Public health practice within a health information exchange: information needs and barriers to disease surveillance.* Online J Pub Health Inform, 4 (3).

[5] D.L. Heymann, L. Brilliant (2011). *Surveillance in eradication and elimination of infectious diseases: a progression through the years.* Vaccine, 29 (Suppl. 4) D141–4.

[6] H. Chen, D. Zeng, P. Yan (2010). *Data visualization, information dissemination, and alerting Infectious disease informatics,* Springer, Boston, MA pp. 73-87

[7] D. Karlsson, et al. (2013). *Visualization of infectious disease outbreaks in routine practice*
Stud Health Technol Inform, 192, pp. 697-701. doi 10.3233/978-1-61499-289-9-697

[8] Hersh, W.R. (2002). *Medical Informatics Improving Health Care Through Information,* JAMA, 288(16):1955-1958. doi:10.1001/jama.288.16.1955

[9] Shirts, B. H., Jackson, B. R., Baird, G. S., Baron, J. M., Clements, B., Grisson, R., Hauser, R. G., Taylor, J. R., Terrazas, E., … Brimhall, B. (2015). *Clinical laboratory analytics: Challenges and promise for an emerging discipline*. Journal of pathology informatics, 6, 9. doi:10.4103/2153-3539.151919

[10] Lindahl, J. F., & Grace, D. (2015). The consequences of human actions on risks for infectious diseases: a review. *Infection ecology & epidemiology*, *5*, 30048. doi:10.3402/iee.v5.30048

Appendix A

[11] Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Diseases, Division of High-Consequence Pathogens and Pathology: About Page (updated 2018). https://www.cdc.gov/ncezid/dhcpp/index.html; and https://www.cdc.gov/ncezid/dhcpp/about.html

[12] Centers for Disease Control and Prevention, Overview of the CDC FY2018 Budget Request (2017). https://www.cdc.gov/budget/documents/fy2018/fy-2018-cdc-budget-overview.pdf';

[13] Lebied, M., (2017). *Your Data Won't Speak Unless You Ask It The Right Data Analysis Questions*, The Datapine Blog; https://www.datapine.com/blog/data-analysis-questions/

[14] Proof of Concept – Scope Document Chattisgarh – State Resident Data Hub, Chhattisgarh Infotech Promotion Society (2016). https://CHIPS.gov.in

[15] *20 Questions You Should Ask While Considering a Visualization Tool*, (2014). http://bigdata-madesimple.com/20-questions-you-should-ask-while-considering-visualization-tools/

[16] Aggarwal, S., McCabe, L.; *Twelve Questions to Ask Your Data Visualization Software Vendor*, SMB Group White Paper, https://www.sas.com/en_us/insights/articles/big-data/data-visualization-vendor-questions.html

[17] *Tableau Technical Specifications and Features*, Tableau Product Page, (2018). https://www.tableau.com/

[18] *Power BI Technical Specifications and Features*, Microsoft BI Product Page, (2018). https://powerbi.microsoft.com/en-us/

[19] *QLIK Technical Specifications and Features*, QLIK Product Page, (2018). https://www.qlik.com/us

[20] *iDashboards Technical Specifications and Features*, iDashboards Product Page, (2018). https://www.idashboards.com/

[21] *Spotfire Technical Specifications and Features*, Spotfire Product Page, (2018). https://spotfire.tibco.com/products/big-data-analytics