CEO of Galaxy Corporation

Sir Michael Smith

Travis Greene,

President - Analytics

1st October 2020

Implementing Distributed Analytical Systems on a Cloud Architecture

Project Charter

Revision 1.0

Table of Contents

1. General Provisions

This section of the project charter outlines the comprehensive elements contained in this document folder, as well as the set of documents it mentions.

1.1. This document refers to the project "Implementing Distributed Analytical Systems on a Cloud Architecture", which is implemented by the organization "Galaxy Corporation".

1.2. The Project Charter formulates the practical requirements and documents the high-level description of the product, service or other deliverable that should meet these requirements and will be result of the project.

1.3. The Project Charter is issued and signed by the Initiator or Sponsor of the project and formally legitimized the existence of the project. It provides the Project Manager with the authority to use the organizational assets in the project operations.

1.4. The approved Project Charter formally initiates the project.

1.5. If the project consists of several phases, the Charter can be updated to verify or improve the decisions taken during the previous iteration of the Project Charter development. In such a case the Charter is re-issued and signed as amended.

2. Normative References

2.1. This document refers to the following documents:

Number	Name	Publishing Date	Author
1	Organizational Detail	2020	PM
2	Business Case Document	2020	PM
3	SWOT Analysis	2020	PM
4	Stakeholder Power/Interest Grid	2020	PM
PMBOK4.1.1.1	Business Documents	2020	PM
PMBOK5.2.3.2	Requirements Traceability Matrix	2020	PM
PMBOK5.3.3.1	Project Scope Statement	2020	PM
PMBOK6.2.3.3	Milestone List	2020	PM
PMBOK7.3.3.1	Cost Baseline	2020	PM
PMBOK10.1.3.1	Communications Management Plan	2020	PM
PMBOK11.2.3.1	Risk Register	2020	PM
PMBOK12.2.3.2	Agreement	2020	PM

3. Terms, Notations, Abbreviations

This defines all the terms, notations and abbreviations used or mentioned that is essential to comprehend this document.

3.1. Terms used in this document:

Term	Definition
Organizational Process Assets	Plans, processes, policies, procedures and knowledge bases that are specific to and used by the performing organization.
Deliverable	Any unique and verifiable product, result or capability to perform a service that is required to be produced to complete a process, phase or project.
Requirement	A condition or capability that is necessary to be present in a product, service or result to satisfy a business need.
Stakeholder	An individual, group or organization that may affect, be affected by or perceive itself to be affected by a decision, activity or outcome of a project, program or portfolio.
Enterprise Environmental Factors	Conditions, not under the immediate control of the team, that influence, constrain or direct the project, program or portfolio.
Objective	*Something toward which work is to be directed, a strategic position to be attained, a purpose to be achieved, a result to be*

Term	Definition
	obtained, a product to be produced or a service to be performed.
Distributed System	*Refers to a system with multiple components residing on different machines (virtual or server- based) that communicate and coordinate actions to achieve a common goal.*
AWS Elastic Computing Instance	*This is an Amazon product that bestows a scalability capacity in the cloud (AWS).*
Amazon Web Services (AWS)	*This is a cloud platform from Amazon that is made up of infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS).*
Hadoop Distributed File System (HDFS)	*This is a distributed file system that handles big data on common commodity.*
SWOT Analysis	*A strategic planning technique used by organizations to identify strengths, weaknesses, opportunities and threat linked to business competition or project planning.*

3.2. Notations used in this document:

Notation	Complete Form
Organization	*Galaxy Corporation*

3.3. Abbreviations used in this document:

Abbreviation	Complete Form
PMBOK	A Guide to the Project Management Body of Knowledge
AWS	Amazon Web Services
HDFS	Hadoop Distributed File System
SWOT	Strengths, Weaknesses, Opportunities, and Threats

4. Project Purpose

The IT department of Galaxy Corporation has witnessed a spike in the volume and variety of data generated from our business areas in recent time, which are used for analytics and making decisions. As the data grows, the current traditional data warehousing techniques cannot efficiently handle the data requirements from our functional areas, which is a threat to our strategic goals. These current challenges are:

I. Scaling the systems linearly incurs huge software licenses and maintenance costs.

II. Analytical processing ability of the systems becomes increasingly inefficient as data computation is performed centrally.

III. Maintenance and Operations are difficult as users of the system have limited control over the infrastructure.

The purpose of the project is to develop and implement a cloud-based distributed analytical system for data processing.

5. Measurable Project Objectives and Related Success Criteria

- The objective of the project is to extract timely value from big data; reduce processing and data storage cost, improve scalability and data query optimization; and build a high performance and resilient system.
- *Success Criteria:*
    - *The project must be completed and fully functional by the end of the 8th month of its life cycle.*
    - *80% reduction in data storage cost.*
    - *Data query optimization's throughput improved by 50 seconds from 60 seconds. Thus, reducing the wait time of complex query results to 10secs.*

6. High-Level Requirements

The requirement of this project is to develop and implement a high-level cloud-based distributed analytical system that will increase data analytical efficiency.

7. High-level Project Description, Boundaries and Key Deliverables

The goal of the project is to implement a secured cloud-based distributed analytical system to reduce data storage cost, faster data processing, and analysis. Relevant elements of the final product of this project has been outlined in Appendix 2 – Business Case. The initial strategic plan of this project is in five (5) phases. These are Requirements Gathering, System Plan & Design, Project Execution, Implementation and Testing, and Performance Monitoring in an 8-month life cycle spanning from November 2020 to June 2021.

8. Overall Project Risk

The high-level risk of this project is when the firm loses key stakeholders affecting the finances that in turn leads to a cut or elimination of the project budget. This draws down to the organization losing out on the current Senior Manager who serves as the project sponsor by accepting a new role in a different organization or branch of Galaxy Corp. This poses as a risk if the new senior manager does not agree with the project purpose and decides to either cut funds or terminate the project. Another high-level risk prone to the project is the non-confidentiality level of implementing a system to reside on a third party's cloud platform. How secure/safe is this platform? What will be the impact on the project? How can these be avoided or mitigated? Having security issues with the cloud will be a big blow to this project since the system would have to be taken down to avoid further damage. This is of grave concern to the project sponsor who will not take chances of any data leaks or breach leading to sanctions from the government. In order to avoid this, the project team must ensure that a contingency plan to safeguard the system and data is implemented.

Unlike the traditional analytical system, designing a distributed system is complex and involves deep understanding of functional areas, network protocols and the client's organizational structure. Also, the fact that we sometimes must replicate data in other nodes, asides from its originating source, adds to its complexity. These can lead to some mishaps during and after the implementation of the system that will lead to more time and funds being spent on debugging and fixing. This will be a negative decline per the project purpose.

Another high-level risk has to do with system compatibility and bug fixing. Galaxy corporation is a very large organization and our proposed system will reflect the organizational structure. Our organization does not employ same type of systems throughout its operations, so connecting systems with different makers and capabilities will pose a problem of incompatibility in a distributed ecosystem. Furthermore, debugging a distributed system as large like ours relatively takes more time. Will the budget be able to absorb this cost resulting in a termination delay of the project?

A qualitative risk analysis will be conducted, probability and impact assessment, by the chief security analyst to secure the project from data leaks or hacks.

Further documentation of project risk will be mapped and prioritized in the risk register based on the level of the stated risks. There will also be another documentation, risk management plan, that will keep track of the risk control strategy and an in-depth action plan on how we can manage each assessed risk.

9. Pre-Approved Financial Resources

The project is estimated to cost $1.5m with an 8-month life cycle spanning from the kick-off meeting to deliverable. However, $50,000 has been set aside to pay for hiccups that may

occur after the implementation of the system.

10. Key Stakeholder List

Name	Position/Organization	Role in the Project
Galaxy Corp.	Client	IT Department
Employees	Users/Galaxy Corporation	End users of the system
Suppliers	Vendors/Various Organizations	Proprietary software owners
Competitors	Various Organizations	Industry's rivals
Project Sponsor	Senior Manager/Galaxy Corp	Monitors the project progress
Govt Agency	Data Regulatory Agency	Regulates data privacy and usage
Customers	Various Entities	Entities from which data are collected

11. Project Approval Requirements

The deliverable of the project must meet the goals/purpose of the project. The functional system must achieve the goal of reducing data storage cost, faster processing and high efficiency in analytics.

12. Project Exit Criteria

These criteria must be satisfied in order for the project to be closed successfully. In the event that the project does not meet one or all the stated criteria, the PM must swiftly assemble the project team for further clarifications.

A. Completed Final Report

B. Final Product (Deliverable) that achieves the goal of the project. This must be fully functional.

C. All bugs in the system must be fully documented from the beginning of the project to its deployment.

However, the project will qualify for an early termination before its target termination date when there is a rising issue of data breach and insecurity during the execution phase. This is to avoid a long legal battle with stakeholders that may result in a huge loss to the company. Another criterion for an early termination of this project will be when the budget apportioned is withheld for the execution of this project. This will gravely impact the schedules and milestones, as well as the entire project as a whole.

13. Project Manager and Team Members

The project manager is the lead in this project and must be accorded the necessary respect and information. The project manager must ensure that the project team members are working to achieve the scope of this project. The project manager in consultation with functional managers must create a work breakdown structure to track tasks, milestones, and budgets. The team members must ensure that there is uniformity and communication within the team. Team members should communicate with their functional heads when a challenge arises during the execution of this project. The chart below illustrates the level of authority and position of the project manager and the team.

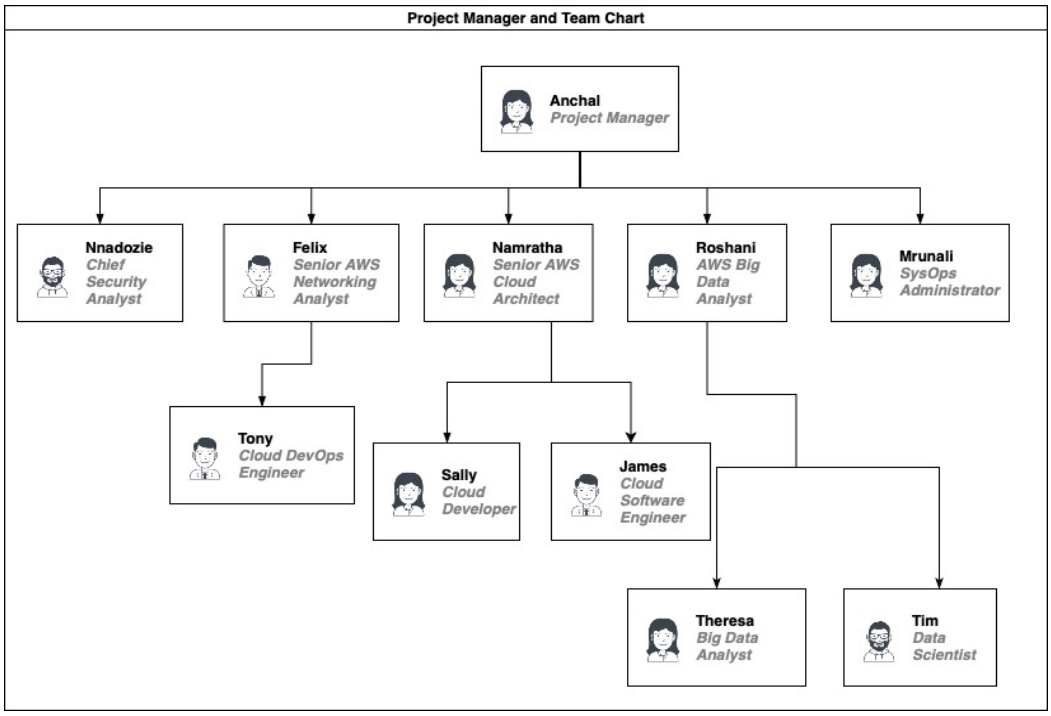


Figure 1. Project Manager & Project Team Members based on hierarchy. The project manager oversees the actions and inactions of the project team, touching base with the team through frequent weekly check-in meetings.

The RACI Matrix below identifies personnel working on this project. However, some of these personnel have subordinates they are working with that was not represented on here. These superiors must ensure that their subordinates are kept in the "loop" during the entire project life cycle. This matrix is to aid the project manager to easily track who is responsible for what, who must approve what, and who must be informed.

Legend:

R – Responsible C – Consult A – Accountable I – Inform

RACI Chart	**Person**					
Activity	***Anchal***	***Nnadozie***	***Felix***	***Namratha***	***Roshani***	***Mrunali***
Create Charter	A	I	I	I	I	I
Requirement Gathering	I	C	R	C	I	A
System Planning & Design	I	C	C	R	I	C
System Integration/Implementation	I	R	R	R	C	R
System Test Plan	A	I	I	I	R	C
Submit Change Request Due to Bugs or unforeseen Circumstances	A	R	R	C	I	C
Final Report	A	R	R	R	R	C

14. Appendixes

Appendix 1 – Organization

Industry: Information Technology (IT), Semiconductor, Research and Consulting (For Profit)

Key Data:

- Company Name: **Galaxy Corporation**
- Date Founded: **1986**
- No. of Employees: **7500**
- Headquarters: **Cleveland-Ohio, US**
- Locations: Mumbai-India; Accra-Ghana; California-US; Wuhan-China; Berlin-Germany

Product and Services:

- Producing- computer chips and circuit, wafers, data processing storage hardware.
- Services: R&D in chips designs, cloud computing, SAS consulting.

Competitors:

- Intel Corp., Samsung, TSMC, Nvidia, Broadcom, Qualcomm, SK Hynix, Texas Instruments

Collaborators:

- Vision Technology, a fabless manufacturing company based in Lagos, Nigeria; Swift Inc., a supply-chain consulting firm in Chicago, US.

Organization Strategy:

To leverage on cutting-edge technologies in delivering the best IT products and solutions to our global customers, conducted in an eco-friendly and professional manner, supported by an intelligent workforce.

Mission Statement:

To continuously improve on ways things are done through advancement in information technology.

Vision:

Powering communities, people, and organizations with reliable and cost-effective technology.

Core Values:

- Innovation
- Continuous Learning
- Customer centric
- Mutual Respect
- Best Practices

Strategic Goals:

- Promote the adoption of environmental-friendly production processes
- Deliver best IT products and services at competitive price
- Continuously grow our market share and increase shareholders value
- Encourage personal learning and career growth of our work force
- Supporting our corporate activities with investment in R&D

Appendix 2 - Business Case

Galaxy Corporation has witnessed a spike in the volume and variety of data generated from our business areas in recent time, that is utilized by the IT department for analytics and making decisions. As the data grows, the current traditional data warehousing techniques cannot efficiently handle the data requirements from our functional areas, which is a threat to our strategic goals. Thus, the existing traditional data warehouse is gradually failing in handling big data within the organization. These current challenges are:

- Scaling the systems linearly incurs huge software licenses and maintenance costs.
- Analytical processing ability of the systems becomes increasingly inefficient as data computation is performed centrally.
- Maintenance and Operations are difficult as users of the system have limited control over the infrastructure.

Proposed Solution: We proposed an implementation of a distributed analytical system on a cloud-based architecture for data processing. This solution is in line with best practices to manage efficiently the processing, storage, and analysis of big data at a lower cost

Project Scope, Budget, and Time: Design and implement a distributed analytical system using AWS Elastic Computing instances. The project is estimated to cost $1.5m with a timeline of 8mths.

Objectives: extracting timely value from big data; reduce processing and data storage cost, improve scalability and data query optimization; build a high performance and resilient system.

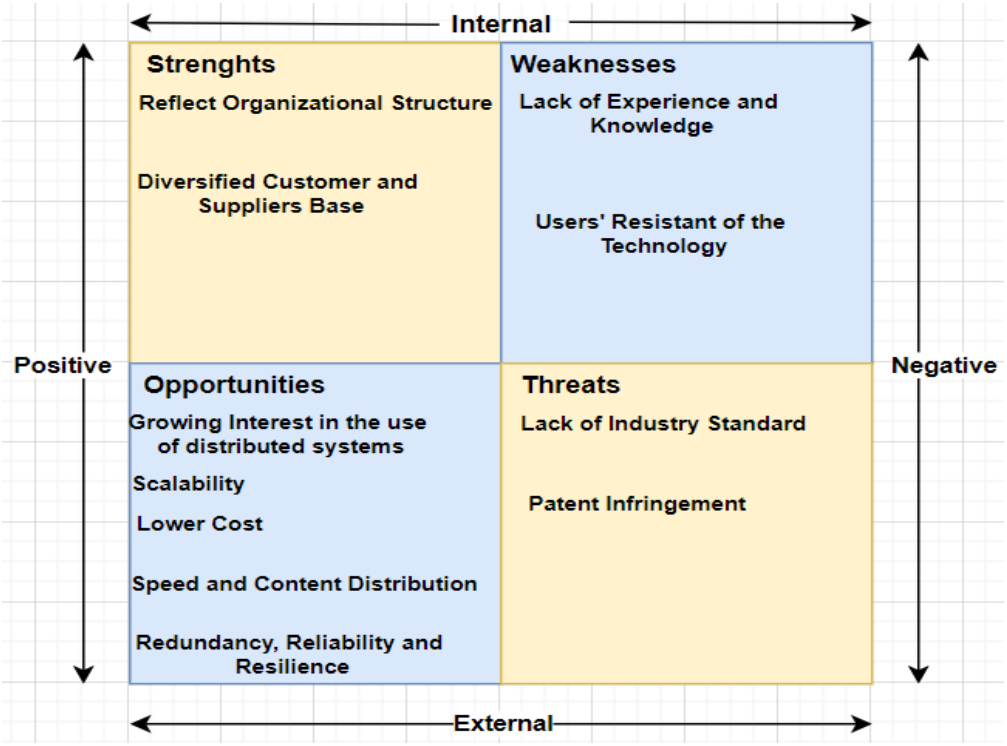
Deliverables:

- Customize and implement an Apache Spark analytical processing engine for big data.

- Implement a Hadoop distributed file system (HDFS), which makes it easy to plug in distributed analytics applications.
- Construct cloud tools for implementing solutions on AWS Elastic Computing instances.

Appendix 3 - SWOT Analysis

This report examines the strength, weakness, opportunities, and threat of our proposed *distributed analytical system* we intend to deploy to replace our tradition, linearly analytical system. Below is a four-quadrant depiction of the SWOT analysis, which is elaborated thereafter.



Strengths

- Reflect Organizational Structure: One remarkable strength of the proposed system is its reflection of the internal structure of our organization. Galaxy Corporation is structured along various functional areas, therefore, it's efficient for its database design to be naturally distributed over these areas. For instance, a location should keep a database containing transactions details of the location, thus a worker at the location can make local inquiries from the database without recourse to a central repository.
- Diversified Customer and Supplier Base. Our diversified customer and supplier network is a good blend to this technology, as various forms of data will be captured accurately and stored at endpoints closed to sources where they are captured and required for analysis. Also, we enjoy unwavering respect from vendors in the industry, which is a plus to our organization in implementing this project successfully.

Opportunities

- Growing Interest in the use of distributed systems: In this era of big data, distributed analytical system is a disruptive technology that Galaxy corporation needs to adopt to revolutionized its operation and leverage on the system's scaling advantage in making timely business decision in order to achieve its strategic goals.
- Scalability: Distributed computing is built upon the premise of seamless scalability. This is the ability of a system to adjust systematically to its performance by effectively distributing computing resources or data among its connected nodes.

- Lower Cost: It is cost-effective to have smaller connected systems with the same computing power serving as data storages than having a central, large repository (mainframe), which is one of the opportunities of a distributed system. Also, it's relatively less expensive adding this smaller workstation to the distributed network than periodically maintaining mainframe server.
- Redundancy, Reliability and Resilience: This is one of the greatest strengths of this proposed systems because we are moving away from an era of frequent server downtime and data redundancy to a period of maximum system uptime. If this project is adequately implemented, a downtime or collapse of any of its workstations does not necessarily mean that the organization activities will be put on hold. Since data is easily replicated across various nodes, a downtime in any workstations simply means users can work with same dataset as contained in the affected workstation from a different node because its hold a copy of same dataset.
- Speed and Content Distribution: Distributed systems ensures timely data processing and efficient content distribution than the conventional system. This leads to quick business decision being made from the vast amount data coming into our system.

Weaknesses

- Lack of experience and knowledge: This involves training end-users on gaining the required knowledge and experience on how the distributed analytical systems work.
- End Users' resistant of the technology: Since inception, most data facing employees of our organization use the legacy system. With the deployment of this solution at completion,

there could be internal bickering among employees. We intend to resolve this foreseeable weakness via active stakeholder engagement and effective communication/diplomacy.

Threats

- Lack of industry standards: Distribution systems architecture is a burgeoning area in the IT industry, so network/system protocols and standards are still at its developing stages. This lack of a consensus in standards limit the potentials of the system.
- Patent Infringement: We consider this a threat because some of the IaaS and SaaS products that will be implemented in this project are from third party/vendors.

Despite the weakness and threat identified, we are confident that if the proposed system is implemented efficiently and with all necessary controls and training put in place, significant portion of the threat and weakness will be addressed. This project is in line with the business case of significantly reducing data processing and storage cost and improve upon timely decision making.

Appendix 4 - Stakeholder Power/Interest Grid

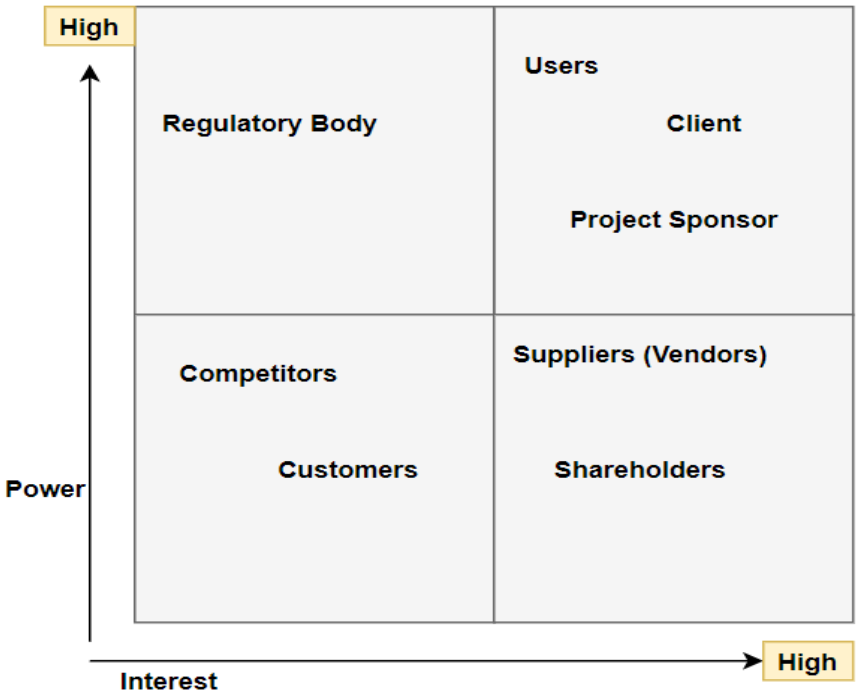
Identifying stakeholders is a continuous activity that takes place throughout the project life cycle, so as we move from one phase of the project to another, new stakeholders are identified and priorities are reevaluated to determine which stakeholder's interest and concern need utmost attention and urgent consideration. At this stage of our project, we have identified some key stakeholders which interest and concerns could influence the scope, budget, and schedule of the project. These stakeholders are categorized as follows:

High Interest and Power Stakeholders:

- Client (Galaxy Corp.)
- Project sponsor
- Users (Employees)
- Regulatory Bodies

Low Interest and Power Stakeholders:

- Competitors
- Shareholders
- Suppliers (Vendors)
- Customers



High power, high interest

The reason why users from relevant functional areas, client and project sponsor are placed in this quadrant is because these stakeholders determine the scope, budgets and deliverables of the projects, and the final product must meet their requirements. It is crucial to invite them to product strategy and road mapping workshops, program iteration events and sprint review meetings as they are empowered to make decisions that shapes the scope and schedule of the project.

High power, low interest

Relevant regulatory bodies that set standards on IT protocols and data privacy fit in here. Though they have limited or no interest in the implementation of this IT solution in our organization, we are guided by these regulations, which in turn influence our scope and deliverables. Therefore, it highly important for the project manager and the team to be conversant with IT protocols and standards and seek clarifications from these agencies on grey areas.

High Interest, Low power

The suppliers of proprietary software (and shareholders) fall in this quadrant. The owners of cloud-based software and distributed file systems have lot of influence on the tools deployed, because a change in some modifications could alter some functionalities of the project scope. Thus, it is necessary to keep them engaged or reach some negotiations through service level agreements.

Low power, Low Interest

This quadrant houses stakeholders whose influence on the project is expected to be low; some of these stakeholders are unaware that such project is being undertaken by the client. However, these stakeholders should not be ignored; they should be monitored to ensure that interest levels haven't changed.