

**Deliverable Acceptance Form**

**Date:** 10/08/2023

**Project Name:** Data Alignment Project

**Deliverable Name:** Data Alignment Project Initiating

**Project Manager:** Yoga Sundaram

**Project Sponsor:** Chief Information Officer

**Acknowledgement and Acceptance:**

We, the undersigned, acknowledge and accept the delivery of the work completed for this deliverable on behalf of our organization. My signature attests to my agreement that this deliverable has been completed. No further work should be done on this deliverable.

**1. Was this deliverable completed to your satisfaction?**

**[ ] Yes**

**[ ] No**

**2. Please provide the detailed reasons for your satisfaction or dissatisfaction for this deliverable:**

**3. If the deliverable is not acceptable, describe in detail what additional work must be done to complete it:**

**Contact’s signature for resubmission of deliverable if found unacceptable:**

**[Contact's Signature]**

**PROJECT SPONSOR:** Chief Information Officer

**Business Case for Data Alignment Project**

**Date:**10/08/2023

**Prepared by:** Yoga Sundaram

|  |
| --- |
| 1. **Introduction/ Background:**   The "Data Alignment Project" is initiated within the Tools Team of Amazon Web Services (AWS), an entity that is primely responsible for developing and maintaining an array of tools and services that empower AWS customers worldwide. In this growing field of Computer vision of both video and image processing, one unavoidable critical challenge it has is effective data alignment in semantic segmentation on the backend records and the Metadata.  **Company Overview:**  Amazon Web Services (AWS) stands as a global leader in cloud computing, delivering a comprehensive suite of services that cater to organizations of all sizes. With a steadfast commitment to excellence, AWS continues to evolve and refine its offerings, shaping the future of technology. The company has a total employee of 70 annotators as FTC’s, 10 FTES’s and 2 Project Managers and a Program Manager.  The product that the AWS provides are High Accuracy data for the Science team that uses the same Data to actively train the Machine learning algorithm for its Alexa devices.  **Background Information:**  Semantic segmentation, a fundamental technique in computer vision, holds immense promise for various applications within AWS and beyond. It basically involves the pixel-level classification of objects within an image, providing invaluable insights for a myriad of tasks, from content moderation to autonomous navigation.  **Issue Description:**  Yet, amid the vast potential of semantic segmentation, a crucial issue has always persisted to the Image and Video annotation (IVA) team. The data generated through semantic segmentation at IVA is currently experiencing a Dissociation with the post-processed data—its non-synchronization with metadata and the backend system is not accurate and inconsistent, which affects the team’s productivity tracking where the entire ML model data flow is relied upon. This dissociation leads to *data inconsistencies*, *complicating data utilization*, *duplicity in the generated data*, also unclean data and hindering the full realization of the technology's potential.  **Importance of Data Alignment:**  Data Alignment is the crucial element that ensures the cohesion of semantic segmentation data with the overarching AWS ecosystem. The effective resolution of this issue is paramount, as it directly impacts data quality, usability, and the efficiency of downstream applications in Object detection through computer vision, which Plays a huge role in the AI research.  **Project Objective:**  The primary objective of the Data Alignment Project is to implement a system or a tool that ensures accurate and timely Alignment of semantic segmentation results, i.e., accurate real-time data that reflects in the metadata and the backend system. By doing so, AWS aims to fortify its capabilities in the field of computer vision and provide customers with a seamless, high-quality experience and products. |
| **2.0 Business Objective:**  **Business Reasons for Initiating the Data Alignment Project/Business Objective**  **Business Process Improvement**  This category typically involves optimizing existing processes to enhance efficiency while reducing costs, and also improve the overall performance. It aims to streamline the data alignment process and make it more efficient.  **Current Business Processes**   1. Data Generation source point: Security cameras and other camera equipment are the sources of the essential data that is utilized for semantic segmentation. The scientific team collects video and images from Alexa devices and uses them to train the machine learning algorithm via a secure connection. 2. Data Utilization and Processing: The IVA team is granted exclusive access to categorized data, which is then distributed into various queues for annotation by the annotation team. This annotation process occurs within designated timeframes. Following annotation, the data is sent to the auditor's queue for verification, where auditors identify and annotate any deviations from standard operating procedures (SOP). The annotated videos are subsequently returned to the science team for final review.   Employee contributions and productivity are determined by the volume of data they handle daily in the system. The science team uses this data to allocate videos and images to other annotation teams within IVA's queue system. Due to the issue, employee productivity is also affected, and rewards are unequally distributed.  **Problems with the Current System:**   1. Data inconsistency: Employee contributions and productivity are determined by their daily data processing volume in the system, this leads to discouraged consistency in the workflow. Any misalignment due to the backend records leads to manual calculation of productivity which at times leads to inaccuracy and employee dissatisfaction. 2. Efficiency loss: This non-alignment leads to data inconsistencies, False identification of object by the ML algorithm that causes some usability concerns, and inefficient manual reconciliation processes. This raises strain on Human resource constraint. 3. Resource Constraint: Due to budget constraints and data confidentiality, human resources cannot be increased as well.   **Business Opportunities:**   1. Successful data alignment will lead to improved data quality, benefiting data analysts, Scientists, Annotators and end-users. 2. Improved Data alignment can provide AWS with a competitive advantage by offering more reliable and efficient services to customers. |
| **3.0 Current Situation and Problem/Opportunity Statement**   1. Currently, the data generated through semantic segmentation is not effectively synchronized with metadata or the backend system. 2. This misalignment leads to data inconsistencies, False identification of object by the ML algorithm that causes some usability concerns, and inefficient manual reconciliation processes. This raises strain on Human resource constraint. 3. The issue adversely affects data quality and operational effectiveness. Employee effectiveness score depends on the attributes that are reflected in the Annotation process. If the attributes marked on the data is inconsistent, employee credit faces a challenge. 4. Uneven productivity calculation on the throughput also affects the operation cost. |
| **4.0 Critical Assumption and Constraints**  **Assumptions:**  1)The project team will have access to necessary critical and confidential data and organizational resources.  2) Stakeholder support and cross-team cooperation will be obtained.  3) The project will adhere to the estimated budget and schedule constraints.  Support from the end-user i.e., Science team and Annotation team.  **Constraints:**  1)Limited project timeline due to time dependent process-flow.  2) Resource constraints.  3) Compatibility with existing systems and infrastructure. |
| **5.0 Analysis of Option and Recommendation**  The primary available option is to develop and implement a data alignment solution that correctly and seamlessly integrates into the semantic segmentation results with the metadata and backend records. This option is recommended due to its potential to not only eliminate data inconsistencies but enhance usability, and improve the overall operational efficiency as discussed in the Critical Assumptions and Constraints. |
| **6.0 Preliminary Project Requirements**  **Preliminary Project Requirements**  **Technical Requirements**:  1) Development of a data Alignment algorithm or a Tool that tracks and records the Data generated during the Annotation process in IVA.  2) Integration with existing semantic segmentation and backend systems.  3) Testing and validation procedures.  4) Migration of few public data and test them on the new tool  **Stakeholder Requirements**:  1) User-friendly interface for data Alignment.  2) Training and documentation of the data for end-users. |
| **7.0 Budget Estimate and Financial Analysis**  The project budget estimate includes expenses related to development, testing, and deployment of the data alignment solution.  The initial estimation of the budget for the project is expected to be $ 400,000(four hundred and thousand dollars. The initiating estimate for the analysis is concluded by planning the working hours for four months, The project manager works 40 hours per week, earning 70$ /hr. and Employees work for 50 hours per week for 80$ per hour. Additional costs will include the software, hardware and services. And an estimated additional cost for the four-month timeframe is $50,000.  With a budget of $400,000, the initial budget for the project is $175000. remaining from the total budget, which can be allocated to other project-related expenses or contingencies as needed. |
| **8.0 Schedule Estimate**  The Project sponsor expects the project to be delivered with the deliverables in an estimate time of 3.5 to 4 months |
| **9.0 Potential Risks**  Identified Risks:   * 1. Technical challenges in developing the Data Alignment algorithm.   2. Annotators and testers aligning with new practice might require more process time for at least initial phase of the project of 2-3 weeks.   3. Additional training is required for the Annotation team, which might cost data confidentiality with the tools team accessing the critical data.   4. Integration complexities with existing systems might occur.   5. Stakeholder resistance to change might also be there.   Risk Mitigation:   * 1. Engagement with the technical experts for algorithm development.   2. Collaborate closely with stakeholders during integration.   Implement a change management plan to address resistance |
| **10.0 Exhibits** |

**Preliminary Scope Statement**

**Project Title:** Data Alignment Project

**Date:** 10/08/2023

**Prepared by**: Akhil Dasari

**Project Justification:**

The project, called “Data Alignment Project", aims to solve an important IT problem with two of the main points.

Firstly, it was determined that the semantic segmentation process continued to record incorrect uptime, especially when using Amazon products such as AWS. This problem has a huge impact because it hinders the organization's ability to effectively track and account for significant work time and the activities that is performed by the employees of Annotation team. Based on this problem, the project recognizes the urgent need for the accurate Metadata technology. To solve the problem of working without information, metadata technology needs to be integrated.

By seamlessly integrating metadata technology into semantic segmentation systems, we hope to bridge the gap between employee activities and the system's ability to record and correlate those activities. These strategies will not only ensure accuracy and accountability, but they will also smoothly increase operational efficiency and effectiveness.

The work of the project therefore involves many methods, including data fusion and optimization as well as the integration of metadata processes. The main goal is to solve the problems caused by working without knowledge in Amazon's semantic segmentation system. We hope to improve the performance of the system through our efforts to clean, improve, and add the metadata. The end result will be increased productivity and accountability in the company by keeping employees working hard and satisfied customers with clean data.

**Product Characteristics and Requirements:**

1. **Data Accuracy Enhancement:**

Characteristic: The system shall improve data accuracy by aligning data generated throughout the semantic segmentation.

Requirement: The alignment process should reduce data discrepancies by at least less than 2% for all the data categories.

1. **Data Integration Efficiency:**

Characteristic: The system shall streamline data integration processes within the AWS ecosystem.

Requirement: The time required for the data integration should be reduced by at least half compared to the current process.

1. **Scalability:**

Characteristic: The system should be scalable to accommodate an increasing volume of data generated through semantic segmentation.

Requirement: The system should handle a 20% increase in data volume without degradation in performance.

1. **Data Security:**

Characteristic: The system shall ensure the security and privacy of aligned data.

Requirement: Data must be encrypted during transmission and storage, adhering to AWS data security standards.

1. **User Access Control:**

Characteristic: The system shall provide role-based access control to aligned data.

Requirement: Administrators should be able to define and manage user roles, limiting access to specific data categories as needed.

1. **Scalable Infrastructure:**

Characteristic: The system infrastructure shall be designed for high scalability.

Requirement: The infrastructure should be capable of not only by accommodating increased workloads but also by provisioning additional AWS resources.

1. **Data Availability:**

Characteristic: The system shall ensure that there is a high availability of aligned data.

Requirement: Aligned data should be available 24/7 with a minimum uptime of 98%.

**Summary of Project Deliverables:**

**Project management-related deliverables:**

1. Business Case
2. Project Charter
3. Scope Statement
4. Work Breakdown Structure (WBS)
5. Schedule
6. Cost Baseline
7. Final Project Report
8. Lessons-Learned Report

**Product-related deliverables:**

1. The actual data alignment solution, including the software, algorithms, and configurations developed to align data from various sources within the AWS ecosystem.
2. Detailed design documents that with specificity of the architecture along with data flow diagrams, and technical specifications of the data alignment solution.
3. Reports on research findings related to the Data alignment methods, corrective technologies, and best practices, this serves as reference materials for the project.
4. Data Alignment Algorithm
5. Software Code and configuration for the Data Alignment
6. Hardware
7. Testing and Quality assurance report with the manual
8. Deployment Strategy
9. Licensing agreement

**Project Success Criteria:**

1. Successful alignment of data generated through semantic segmentation within the AWS ecosystem.

2. Improved data accuracy and consistency.

3. Streamlined data integration processes.

4. Enhanced usability of data for analysis and decision-making within AWS.

**Stakeholder Analysis**

**Prepared by:** Akhil Dasari, Yoga Sundaram.

**Date:**10/08/2023

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stakeholder** | **Influence on the project** | **Estimated Influence** | **Level of interest** | **Interest area** |
| Project sponsor-CIO | Supportive, with a keen interest in improving employee work tracking for enhanced efficiency in the Data Alignment, The end product. | High | 1 | Project progress, budget allocation. |
| Project Manager | Ensure compliance among key players to ensure product success. All stakeholders should collaborate to see the project through | High | 1 | Project planning,  and scheduling, Task Management. |
| Data Scientists | Supportive, throughout the process with a focus on not only utilizing improved work tracking systems but also to enhance data analysis capabilities. | Medium | 2 | Data availability. |
| Data Analyst | Supportive, as they rely on accurate data and efficient work tracking for their analytical tasks. | Medium | 2 | Efficiency in data retrieval and analysis. |
| Science Team | The completed project should be able to address the Data misalignment mentioned previously and also, they provide the source data to test for the project. Their requirement varies periodically with accuracy. | High | 1 | data collection and analysis, experimentation design. |
| Amazon IT Department | Supportive, as they recognize the need  for technological enhancements  alignment to improve work tracking. | Medium | 2 | Technological enhancements, data security, scalability. |
| Employees (Annotation Team) | Expectant, as they anticipate improved work tracking systems that  their efficiency and productivity are kept on track. | High | 1 | Work tracking improvements, ease of use, impact on daily tasks. |
| Amazon Customers | Indirect stakeholders who may benefit  from improved service efficiency due to  enhanced work tracking systems and the Aligned data help them choose the product in the market that is developed based on the deliverable. | High | 1 | Satisfaction with the end product with the provided data |
| Competitors | Likely to be monitoring Amazon's initiatives closely to assess their impact on the market and potentially respond with their own strategies. | Medium | 2 | market positioning, and any vulnerabilities they can exploit. |
| Program Manager | Likely supportive, with a focus on overall project success and alignment with organizational goals. | High | 1 | stakeholder communication. |

Project Charter

**Project Title**: Data Alignment Project

**Project Start Date:10/08/2023** **Projected Finish Date:** 12/12/2023

**Project Manager:** Yoga Sundaram

**Project Objectives:**

**The Data Alignment Enhancement Project** aims to:

1. Improve Data Accuracy: Enhance the accuracy and consistency of the real-time data generated through semantic segmentation within the AWS ecosystem and record it in the database.
2. Enhance Data Usability: Improve the usability of data for analysis and decision-making, benefiting both AWS teams and customers.
3. Streamline Data Integration: Simplify the data integration processes to reduce time and resource requirements through effective means.

These objectives collectively drive the project's mission to align data effectively, fostering data-driven decision-making and operational efficiency within AWS.

**Main Project Success Criteria:**

1. Accurate Data generated that rewards the employee working hours and smooth flow of processed Data to the Science team.
2. Successful alignment of data generated through semantic segmentation.
3. Improved data accuracy and consistency.
4. Streamlined data integration processes.
5. Enhanced usability of the data for analysis and decision-making within AWS.

**Benefits:**

**Tangible Benefits:**

1. **Cost Savings:** Streamlined data processes are expected to reduce operational costs.
2. **Increased Efficiency:** Faster data alignment leads to improved team productivity.

**Intangible Benefits:**

1. **Improved Data Quality:** Enhanced data alignment ensures higher data accuracy.
2. **Enhanced User Satisfaction:** Teams within AWS benefit from better data usability.

**Key Schedule Milestone:**

**Project Initiation:** This milestone marks the project's official launch. It involves completing the project charter, assembling the project team, and beginning the preliminary planning processes, such as identifying, analysis of the stakeholders and defining the project's scope**.**

**Data Alignment Enhancement Implementation:** This accomplishment signifies the stage in which the real data alignment improvement solution is put into place and put through its paces inside the AWS environment.

**User Training and Adoption:** once Data alignment implementation is successful post the installation, this milestone will consist of informing the AWS Annotation team of the new data alignment system and assuring their efficient adoption.

**Roles and Responsibilities**

|  |  |  |  |
| --- | --- | --- | --- |
| **Role** | **Name** | **Organization/Position** | **Contact Information** |
| **Project Sponsor** | Rick Dalzell | Amazon Web Services /CIO | xyz@amazon.com |
| **Project Manager** | Yoga Sundaram | Amazon Web Services | xyz@amazon.com |
| **Project Co-Ordinator** | Akhil Dasari | Amazon Web Services | xyz@amazon.com |
| **Project Champion** | Uttej Reddy Pateti | Amazon Web Services | xyz@amazon.com |
| **Project Team Member** | Phanendra Ghoud Palle | Amazon Web Services | xyz@amazon.com |

**Sign-off:**

**Comments:**

**Project** **Kick-off Meeting**

**Date**:02/10/2023

**Meeting Objective:** To Get the project off to start by introducing the key stakeholders, The reviewing of project goals, and discussing future plans

**Agenda:**

1. Introduction of the attendees
2. Background of Semantic segmentation the business process of the current organization, and the challenge with their workflow were discussed
3. Review of project-related documents such as Business case (Data misalignment and the employee dissatisfaction it causes with missed data is discussed deeply), project charter.
4. Discussion about the Project organizational structure with emphasis to data management with organization that will have access to the critical data and its scope for confidentiality and its maintenance with integrity throughout the process.
5. Discussion of project scope, time, and cost goals
6. Discussion of other important topics related to Semantic segmentation and various phases of the project
7. List of action items from meeting

|  |  |  |
| --- | --- | --- |
| **Action Item** | **Assigned To** | **Due Date** |
| **Naming the project** | **All team** | **09/05/2023** |
| IT | Akhil Dasari | 11/08/2023 |
| Testing of the Data Alignment tool | Phanendra ghoud palle | 12/08/2023 |
| Business case | Yoga Sundaram | 10/09/2023 |
| Project charter | Yoga Sundaram | 09/30/2023 |

**Date and time of next meeting:**

09/10/2023 08:00 PM EDT.

**Feasibility Analysis**

**Technical Feasibility:**

1. **Infrastructure Readiness:** Assessing the current state of AWS IT infrastructure. Ensures that it can provide hardware, servers, and network components that are capable of supporting the proposed IT project.
2. **Software and Tools**: Evaluate the availability and compatibility of necessary software and tools for the project. Check if your team has access to the required licenses and if the tools align with your IT ecosystem. Familiarity with these tools among your team members might **enhance** technical feasibility.
3. **Development Expertise**: The IT team possesses the required development skills and expertise for the Data Alignment project. The proficiency in programming languages, frameworks, and technologies are relevant to the project.
4. Greater expertise will positively affect the technical feasibility.

**Technology Familiarity for Data Alignment:**

1. **Technology Stack**: Assessing the team's familiarity with **cloud services, databases**, **cybersecurity protocols**, will have a strong understanding for the crucial effective implementation.
2. **Training and Skill Development:** Consider investing in training and skill development if there are gaps in your team's knowledge. Strengthening their expertise in the required technologies can improve the overall effectiveness and feasibility of the project.

**Project Size and Testing:**

1. **Scope Analysis**: The scope of the project, Project aligns with the available resources, including manpower, time, and budget for the organization’s capability.
2. **Resource Allocation:** The necessary resources, such as human, financial, and technical to undertake the project successfully, are feasible, and The availability of skilled personnel and their capacity to handle the project's size and complexity are feasible as well.
3. **Testing and Quality Assurance:** Plan for rigorous testing will assure quality procedures to ensure that the integrated systems work seamlessly and are free from critical issues. **Delays in integration testing** can impact project timelines.

In IT project initiation, technical feasibility ensures that the organization has the capability, technology familiarity, and resources to build and deploy the proposed IT solution. Assessing these factors comprehensively concludes that the project can be effectively initiated from a technical standpoint.

**Economic Feasibility**:

1. **Development Costs**: the upfront costs, including hardware and software, personnel, and any Miscellaneous expenses required to develop and implement the data alignment solution.
2. **Intangible Benefits and Costs**: Recognize intangible factors, such as improved data quality, enhanced decision-making, and customer satisfaction, which may impact economic feasibility.
3. **Annual Maintenance Costs:** the ongoing operating costs, which encompass maintenance, support, licensing fees, and any additional expenses incurred after implementation.
4. **Annual Benefits**: Analyze the expected financial benefits of the data alignment project, such as cost savings, efficiency gains, and potential revenue increases.

**ROI :**

1. a) **Cost Savings:** Reduced savings if it is provided by the organization with reduced expenses

b) **Operational Efficiency**: faster and better service provided by the organization may get better scalability as well.

**Organizational Feasibility:**

**Strategic Alignment**: Ensure that the data alignment project aligns with AWS's strategic goals and its commitment to providing high-quality services to customers.

**Project Champion**: key advocates or champions within the Tools Team at AWS who are dedicated to the project's success and can mobilize support.

**Senior Management Support**: Secure support from senior management within AWS to provide the necessary resources, direction, and leadership for the project.

Users and Stakeholder Involvement: Engagement of the data analysts, scientists, and other stakeholders within AWS ensured that the project addresses their needs and aligns with their objectives and doesn’t violate any regulations.

**Time Feasibility**:

**Project Schedule:** A detailed project schedule and timeline are provided and will be developed based on the feedback provided on each phase

**Availability:** The necessary resources, including the skilled personnel from the AWS and the equipment to test the tool, are available to meet the expectations of the project on the said deadlines.

**Risk Mitigation:** Potential risks that could impact project timelines and develop strategies are. Effective risk management is critical for time feasibility. The tools Team at AWS, who are dedicated to the project's success, can mobilize required support.

By considering these factors, including time feasibility, we can conclude that Data Alignment Enhancement Project is within the AWS satisfactory condition, also viable. The above completed analysis aids in determining whether the project is technically achievable, economically viable, and aligned with organizational goals, within the desired timeframe. And promises to achieve the desired set of goals as expected