

## **REQUEST FOR QUOTES (RFQ)**

## **Best Value Analysis Process for:**

## OLDC's initiative to implement an updated data architecture in Azure

OregonBuys: S-52500-00011234 HECC# 24-024

Issue Date: 8/8/2024

Issuing Office: Higher Education Coordinating Commission (HECC) on

behalf of the office of Oregon Longitudinal Data

Collaborative ("OLDC")

Single Point of Contact ("SPC"): Shannon Ziglinski

503-979-5850

hecc.procurement@hecc.oregon.gov

Offer Due Date and Time: August 22, 2024 1:00 PM Pacific Time

Service Category: 6.3 Extract Transform Load (ETL) Development

Issued to: MPSA Holders name and MPSA #

Deloitte	8482
Plante	8407
Microsoft Corporation	8487
The North Highland Company, LLC	8477
Resource Data, Inc.	8468
Windsor Solutions, Inc.	8466

## **EVENTS SCHEDULE**

EVENT	DATE	TIME (PT)
Issue RFQ	August 8, 2024	4:30 PM PT
Due date for questions/clarification ("Q&A")	August 16, 2024	10:00 AM PT
Answers to questions/clarification (approx.)	August 20, 2024	5:00 PM PT

Quote due date and time	August 22, 2024 by 5:00 PM PT
Notice of Intent to Award (approx.)	To be determined
Protest test period ends	7 calendar days after notice of intent to award announcement

#### 1. Introduction

The Higher Education Coordinating Commission (HECC) is looking to contract with a person or firm to provide Extract Transform Load (ETL) development services for OLDC's project to implement an updated data architecture in Azure.

The length of the contract is estimated to be 120 days and commence before September 18, 2024.

HECC reserves the right to extend the work order contract up to eighteen months and to award more than one Contract if it is HECC's best interest.

#### 2. Definitions

- 1. **Authorized Purchaser** means HECC or OLDC.
- 2. **Consultant** means the Offeror selected for Contract award.
- 3. **Offeror** means the organization submitting a quote in response to the RFQ.
- 4. **Project** means the OLDC's initiative to improve data architecture.
- 5. **Services** means the totality of ETL development identified in the RFQ.
- 6. **APM** means Authorized Purchaser's Project Manager(s).
- 7. **ETL Developer** means Consultant's ETL Developer(s).
- 8. **OLDC** means the Oregon Longitudinal Data Collaborative.
- 9. **SLDS** means the Statewide Longitudinal Data System.
- 10. **MDM** means Master Data Management.
- 11. **EIS** means Enterprise Information Services, the organization led by the State Chief Information Officer.
- 12. RFQ means Request for Quote.

## 3. Background Information

### **About the OLDC**

Established in 2019 and situated within the HECC, the Oregon Longitudinal Data Collaborative (OLDC) is an inter-agency research office that brings together data from various agency data partners and links individual records for analysis and reporting. The OLDC research team uses this data to provide recommendations aimed at enhancing education and workforce policies. The data is maintained in the SLDS and is managed by the OLDC.

The OLDC has a well-established data sharing agreement with its data partners, including the Oregon Department of Education (ODE), HECC, Oregon Employment Department, and the Teacher Standards and Practices Commission. This agreement specifies what data elements will be provided, its refresh schedule (either annually or quarterly), and the method of transfer (via direct connections or flat file).

### About the SLDS

The SLDS currently uses Informatica PowerCenter and MDM products for data processing and identity matching, enabling longitudinal analysis. The system is maintained by OLDC's data team, which includes a Systems Architect/MDM Developer responsible for the design and architecture, as well as maintaining the identity matching process. The data team also includes a Solution Architect/Data Engineer, though the position is currently vacant. The team is led by the program's Data Strategy and Operations Lead, who ensures data governance compliance and provides project management and business/system analysis. The OLDC team also includes a Business Data Analyst who maintains the data catalog and participates in data governance development, and a Researcher who conducts research and reporting based on the Research Agenda set by the Governance Committee.

In 2022, the SLDS data and infrastructure transitioned from the State's data center to Azure in Microsoft's Commercial Government cloud, following the State's and HECC's IT Strategic Plan. Last year, the OLDC assessed options to enhance data processing and chose to transition ETL processes from Informatica PowerCenter to Azure Synapse. From late 2023 to early 2024, the OLDC worked with Microsoft to develop a more efficient data architecture to boost performance, cut costs, and expand data utilization. The outcome is a Data Lakehouse and Metadata Driven architecture built on Azure Synapse using PySpark Notebooks for ETL and that is integrated with Informatica's MDM product for identity matching.

### About the data and data flow

There are approximately 200 tables across the four data partners with an estimated total size of 156 gb.

The updated architecture includes the following data flow:

- Landing Zone. Ingest data as-is and ensure compliance according to the Data Sharing Agreement.
- Cleansed Zone. Move data from Landing Zone and apply agency data quality rules.
- Matching Zone. Move data from Cleansed Zone that will be used for identity matching and send to Informatica MDM to begin matching process. Retrieve matched results from Informatica MDM.
- Curated Zone. Move data from Cleansed Zone (i.e., data not used for Identity Matching) and Matched Results and apply global OLDC-maintained transformations and standardizations. Create a presentation layer for researchers that complies with data minimization rules (e.g., hide metadata columns, personally identifiable information, etc.).

## 4. Project Description, Overview of Services, Project Oversight and Project Structure

## 4.1 Project Description

This Request for Quotes (RFQ) seeks a consultant firm to provide ETL development services (Services) for the OLDC Project. HECC expects the successful Offeror to begin work within approximately 10 business days after execution of a Work Order Contract (Contract).

The objective is to enhance the efficiency, performance, and use of the data of the SLDS by transitioning ETL processes to Azure Synapse. The project, with a completion target of December 13, 2024, is currently in the phase of setting up access, configuring security and auditing controls, implementing a metadata database, and establishing data connections to our data partners' systems.

There are three key milestones for this project:

• October 4, 2024: complete phase 1: Implement ETL through the Cleansed Zone

- October 25, 2024: complete phase 2: Implement ETL through Curated and Implement an Existing Report
- January 3, 2024: complete phase 3: Create Research Presentation Layer and Project Wrap up

#### 4.2 Overview of Services

Specific tasks expected of the ETL Developer are detailed in section 5.1.1. Tasks will include, but are not limited to:

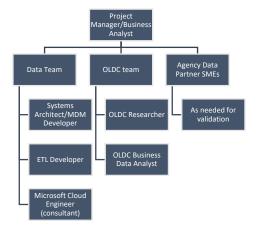
- 1. Develop Azure Synapse notebooks and pipelines according to data architecture
- 2. Document code, processes, and components
- 3. Refine curated zone according to research and reporting needs
- 4. Implement an existing annual report that includes applying archival rules
- 5. Make recommendations to data governance policies and processes

## 4.3 Project Oversight

The OLDC Data Strategy and Operations Lead will serve as the project manager, with the System Architect reviewing tasks and deliverables. The OLDC data governance committee will approve data governance policies, ensuring management of privacy, security, data quality, retention, and destruction.

## **4.4 Project Structure**

The project is led by the OLDC Data Strategy & Operations Lead, with development performed by the Systems Architect/MDM Developer and the ETL Developer. OLDC has an agreement for consultation with a Microsoft Cloud Engineer though the end of the project. Other OLDC team members will engage with the project as needed related to data governance or researcher requirements. Agency data partner subject matter experts (SMEs) will be involved as necessary for technical data connection issues or data quality questions.



## 5. Scope of Services:

## 5.1 General Requirements

**5.1.1** Consultant shall provide ETL services necessary to implementing the SLDS architecture in Azure.

## In addition to completing the deliverables described below, ETL development tasks include:

- 1. Participating in ongoing OLDC Project status meetings and provide task progress and status.
- 2. Providing demonstrations to assure mutual understanding, and that resulting project work adheres to the SLDS Data Architecture.
- 3. Obtaining OLDC approval and sign-off for Project deliverables.
- 4. Engaging with agency data partners regarding data quality questions.
- 5. Providing data governance policy recommendations.
- 6. Documenting operational processes throughout the ETL development.

## 5.1.2 Phase 1: Implement ETL through Cleansed Zone

The Consultant will evaluate and update existing PySpark Notebooks for each data partner to validate the data received in the Landing Zone, ensuring compliance with the Data Sharing Agreement. Feedback should enhance performance, error handling, or other improvements.

The Consultant will be required to write PySpark Notebooks for each data partner that moves data from Landing Zone to the Cleansed Zone and applies approved agency-specific data quality rules. The consultant will be provided with the rules but can propose data quality improvements. For example, a data partner asked us to remove records that have an empty string in a license field which indicates an error during the collection process.

The Consultant will ensure that notebooks are cost-efficient, use git for version control, implement proper error handling and logging, and comments clearly describe the code and expected results.

## 5.1.3 Phase 2.1: Implement ETL through Curated Zone

The Consultant will write four PySpark Notebooks, one for each data partners to move data from the Cleansed Zone (excluding identity-matching data) and the Matching Zone (identity-matched results) to the Curated Zone, incorporating the golden record ID and applying OLDC transformation rules. Rules provided by the governance committee should be followed, with opportunities to propose standardizations or other transformations to enhance data usability. For example, standardize gender data as 'male, female, nonbinary' across all sources.

The Consultant will ensure notebooks are cost-efficient, use git for version control, implement robust error handling and logging, and include clear comments describing code and expected results.

## 5.1.4 Phase 2.2: Implement an existing report

The Consultant will create a Notebook and save results for an existing data file that OLDC provides annually in October to ODE for federal reporting requirements related to Career Technical Education (CTE) students. Results will be saved for the current year and the results from prior years will be moved into long-term archive. Existing logic can be reused and adapted. The report flags CTE students (approximately 18,000 rows) who are working and/or attending a college or university after high school graduation.

### **5.1.5** Phase **3.1**: Create presentation layer for researchers

The Consultant will develop a presentation layer for researchers, considering options including but not limited to: Views, Materialized Views, 'Select as Table'. The presentation layer will apply data minimization rules such as hiding personal identifiable information and metadata fields.

## 5.1.6 Phase 3.2: Deploy to Production and project wrap up

The Consultant will schedule pipelines, deploy them to the production environment, and ensure all ETL-related processes are documented in the Wiki pages in Azure DevOps.

## **5.1.7 Required Regularly Scheduled Meetings**

Consultant will be required to prepare for and participate in the following meetings:

- 1. Bi-weekly Project Sprint Planning meeting (60 minutes)
- 2. Weekly Project Status Check-in (45 minutes per week)
- 3. Weekly guided work sessions with Microsoft Cloud Engineer (up to 120 minutes)

### 5.2 Proposed Tasks and Deliverables

Offeror must address each of the Proposed Tasks and Deliverables, detailed in RFQ Attachment A, in its Offer in the context of:

- 1. Approach, methodology and availability to deliver Services;
- 2. Organizational qualifications and experience in comparable projects involving similar Services;
- 3. Deployed personnel qualifications and experience in comparable projects involving similar Services; and
- 4. Cost for delivery of such Services.

## 5.3 Deliverable Submissions, Walkthroughs, and Acceptance Criteria

### 5.3.1 Submissions and Acceptance

- 1) Consultant shall submit each Deliverable to Authorized Purchasers in written form. Authorized Purchasers will review the Deliverable to identify any major flaw, error, or omission in the Deliverable.
- 2) Authorized Purchasers will have up to 5 business days to provide feedback to Consultant.
- 3) Consultant will address feedback as necessary and re-submit the deliverable for final review.
  - a. If a deliverable is <u>not accepted</u>, Authorized Purchasers shall describe the reason(s) the deliverable is not being accepted, and the Consultant will have five business days to correct, update and resubmit the Deliverable for Approval.
  - b. When a deliverable is <u>accepted</u>, the Consultant will obtain formal signatures for approval using either a formal Deliverable Approval sheet or a signature section within the Deliverable.
- 4) Submission of Code and Documentation: Consultant shall submit all code, scripts, and associated artifacts (e.g., documentation, configuration files) to Authorized Purchasers via the

designated version control system or code repository. Additionally, Consultant shall document all processes and relevant information in Azure DevOps Wiki pages.

## 5) Code and Documentation Review:

- c. Authorized Purchasers will review the submitted code and Wiki documentation to identify any major flaws, errors, omissions, or gaps in information. Feedback will be provided within 5 business days.
- d. For Wiki documentation, Authorized Purchasers will review the clarity, completeness, and accuracy of the content.

## 6) Revisions and Resubmission:

- e. Consultant will address the feedback and make necessary revisions to both the code and Wiki documentation. The updated code and Wiki pages must be resubmitted for a final review.
- f. If a submission is not accepted, Authorized Purchasers will provide specific reasons for rejection. Consultant will have 5 business days to correct, update, and resubmit both the code and Wiki documentation for re-evaluation.

## 7) Acceptance and Approval:

- g. Upon acceptance of the deliverables, including code and Wiki documentation, Consultant shall obtain formal approval from Authorized Purchasers. Approval may be documented through an electronic approval process, such as comments or approval workflows within Azure DevOps.
- h. Accepted deliverables will be documented as approved in the version control system and Azure DevOps Wiki, with relevant metadata or notes indicating formal acceptance.
- 8) **Finalization and Integration**: Once approved, the code will be integrated into the main codebase or system, and the Wiki documentation will be finalized and published in Azure DevOps. The Consultant will assist with the integration and finalization processes if necessary.

## 6. Questions and Requests for Clarification:

All questions and requests for clarification regarding this RFQ must be submitted in writing by e-mail to the Authorized Representative and must be received no later than the date and time identified in the Events Schedule.

When appropriate, as determined by Authorized Purchaser in its sole discretion, revisions, substitutions or clarifications of the RFQ will be sent electronically.

### 7. Quotation Submittal:

Quotes must be received on or before the Offer Due Date and Time at the location identified above to be considered. Quotes will not be accepted after the Offer Due Date and Time. Quotes will be accepted by e-mail Authorized Representative identified above. Authorized Purchaser may extend the Offer Due Date when it is in the best interest of Authorized Purchaser.

Authorized Purchaser may reject all Quotes or to cancel this RFQ if in its sole determination, it is in the best interest of Authorized Purchaser.

## 8. Submittal Requirements:

- a. Quotes must include a description of Offeror's methodologies for providing the Services and a draft Statement of Work based upon the Scope of Services above including suggested schedule and Key person responsible for each task and deliverable.
- b. Quotes must include a description of Offeror's Key Persons, other staff, their experience and availability.
- c. Quotes must include a description of past projects Offeror has completed that are similar in scope to what is being requested in this RFQ.
- d. Quotes must include 3 references from customers for whom Offeror has performed similar projects within the past five years.
- e. Cost Offer. Offeror must complete and return Attachment B as Offeror's cost offer, including estimated hours per Task and Deliverable. The cost offer must be provided as a separate file from the remainder of the Offer.
- f. Security Certification. Offer must contain a statement demonstrating Offeror's agreement that if awarded a Contract:
  - Offer and Offeror's staff with access to State systems, facilities, data, and confidential
    information will submit to all security checks requested by HECC, which may include
    any combination of fingerprinting, Oregon Law Enforcement Data Systems (LEDS) and
    Federal Bureau of Investigation Criminal Justice Information Services (FBI CJIS)
    background checks; and
  - Upon request, Offeror and Offeror's staff will sign a non-disclosure agreement for any and all data or information received or processed on its equipment from the State of Oregon; and
  - Offeror will protect at all times State of Oregon sensitive material; and
  - Offeror will meet or exceed the State of Oregon's security standards as set forth in the following:
    - Privileged Access Monitoring and Reporting viewable at: https://www.oregon.gov/das/Policies/107-004-140.pdf
    - Statewide Information and Cyber Security Standards viewable at: <a href="https://www.oregon.gov/das/OSCIO/Documents/2019StatewideInformationAndCy">https://www.oregon.gov/das/OSCIO/Documents/2019StatewideInformationAndCy</a> berSecurityStandardsV1.0.pdf
    - Statewide Cloud Computing policy: http://www.oregon.gov/das/policies/107-004-150.pdf
- g. Submitted Quotes Subject to Disclosure as Public Records. If an Offeror/Consultant includes information and data with its submitted Quote that Offeror/Consultant regards as proprietary, privileged, or otherwise confidential; Offeror/Consultant must identify such information in a separate document submitted with its Quote and provide a redacted submission along with the original submission. Otherwise, State will assume that Offeror/Consultant consents to public disclosure of the original submission.

## 9. Offeror/Consultant Screening and Selection Process:

Quote submissions will be reviewed to determine if all Submittal Requirements have been met. Those meeting the Submittal Requirements will be evaluated to determine the "Best Value" for the State. "Best Value" is based solely on the evaluator's determination of what best meets the needs of Authorized Purchaser taking into account the considerations set out in this section of the RFQ.

Offers will be evaluated on a two-tier evaluation system.

- 1. The first tier is an initial evaluation of written Offers by an evaluation team.
- 2. The second tier, if conducted, will consist of interviews with Offeror's Key Persons and additional presentations that OLDC may require.

## 9.1 First Tier – Evaluation of Written Offers

The first tier is an evaluation of Offers by an evaluation team assembled by OLDC, which may include representatives from other agencies or entities, to determine the apparent successful Offeror/Consultant under this RFQ.

### **Required Offeror/Consultant Qualifications**

- 1. Five or more years of experience in the Data Engineering field or in a related area
- 2. Make recommendations surrounding improving processes, efficiency and best practices based on lessons learned from past engagements
- 3. Proficiency in using one or more programming languages for ETL within Azure
- 4. Experience with distributed computing frameworks (e.g., Apache Spark, Microsoft Azure Databricks)
- 5. Experience with data pipelines and workflow management tools (e.g., Azure Synapse, Azure DevOps)
- 6. Advanced SQL skills and data modeling
- 7. Excellent problem-solving, communication, and organizational skills

## **Desired Offeror/Consultant Qualifications**

- 1. Microsoft Azure certifications
- 2. Data migration experience
- 3. Experience with Informatica PowerCenter or MDM products
- 4. Experience with regulated and personally identifiable data (e.g., FERPA)
- 5. Experience working within a governmental (federal, state, or municipality) environment/culture

## Scoring Criteria (BVA) and Weights

Best Value Analysis	Maximum Points Available	
Qualifications and Experience of the Offeror/Consultant Firm	100	
Qualifications and Experience of Key Persons	350	
Availability and Resource capacity	300	
Cost	250	
TOTAL	1000	

Each non-cost category will be scored based on the information provided in the Quotes in the context of the Rating Scale set out below.

SCORE	EXPLANATION
90 – 100%	<b>OUTSTANDING</b> - Response meets all the requirements and has demonstrated in a clear and concise manner a thorough knowledge and understanding of the subject matter and opportunity. The Offeror provides deep and meaningful insight into its expertise, knowledge, and understanding of the subject matter.
61 – 89%	GOOD TO VERY GOOD – Response provides useful information, while showing experience and knowledge within the agency program/project area. Response demonstrates above average knowledge and ability with no apparent deficiencies noted.
50 - 60%	SATISFACTORY – Response meets all requirements in an acceptable manner. Response demonstrates an ability to comply with guidelines, parameters, and requirements, but with little or no additional value provided.
1 – 49%	POOR TO FAIR – Offeror meets minimum requirements but does not demonstrate "Satisfactory" knowledge of the subject matter.

Based on scores from the first-tier evaluation, Authorized Purchaser may determine to award a contract to the top-ranking Offeror, select one or more of the top-scoring Consultant(s) as finalists for the second-tier evaluation, or determine to cancel the solicitation.

Authorized Purchaser's determination is final.

### 9.2 Second Tier - Evaluation Interviews and Additional Presentations

If a second-tier evaluation is conducted, an interview panel will interview one or more of the top scoring Offerors selected during the first-tier evaluation. The second-tier in-person interview evaluation may involve the submission of sample work products for evaluator review, the interview of the participating Consultants, and revised cost quotes.

If conducted, the second-tier evaluation will determine the Consultant that will be accepted for work under a Work Order Contract resulting from this RFQ. Commitments made by the Consultant at the oral interview, if any, will be considered binding.

#### 10. Award:

The Offeror with the highest-scored Quote will be awarded a Contract in a form substantially similar to the Work Order Contract form attached to the MPSA. SPC may negotiate Contract terms and conditions with the successful Offeror. Authorized Purchaser may award less than the full scope described in this RFQ. The Offeror with the most advantageous quote will be awarded a contract in a form substantially similar to the Work Order Contract attached as Attachment C to the MPSA. Authorized Purchaser may negotiate contract terms and conditions with the successful Offeror.

CONTRACT AWARD DETERMINATION IS FINAL BASED ON THE SCORES PROVIDED BY THE EVALUATION COMMITTEE, SUBJECT TO A TIMELY, COMPLIANT AWARD PROTEST FROM AN ADVERSELY AFFECTED OFFEROR.

AWARD PROTEST OPPORTUNITY CONSISTENT WITH ORS 279B.410. Authorized Purchaser will provide a seven (7) calendar day period following release of any Notice of Intent to Award ("Notice") during which an adversely affected Offeror may protest such award. An Offeror who claims to have been adversely affected by the selection of a competing Offeror shall have seven (7) calendar days from the date of such Notice in which to deliver a signed protest to the Authorized Purchaser. The Authorized Purchaser will deliver the Notice to all Offeror via email transmission to the Offerors' named representatives. Adversely affected Proposers will have until 5:00 p.m. Pacific Time, on the protest deadline date to submit a protest in accordance with the requirements in ORS 279B.410.

The submitted protest must bear the signature of a company official authorized to represent Offeror's interests. Offerors may use an electronic signature copy via email to satisfy the protest submission deadline. To be adversely affected, an offeror must demonstrate that all higher ranked quotes were ineligible for selection. Authorized Purchaser will not consider a protest submitted after the deadline. Protests will be resolved according to Oregon Administrative Rules. Authorized Purchaser will not consider protests that do not include the information required under ORS 279B.410.

RFQ Attachments A and B begin on the next page.

## Attachment A

## Proposed Tasks and Deliverables

## **Task 1: ETL Development**

Authorized Purchaser will require Consultant to perform the following Tasks:

## **Task 1 Deliverables**

Item#	Description	Task Due
1.1	Onboarding (e.g., equipment setup and access)	One (1) week from start date
1.2	Perform code review of existing PySpark Notebooks and make agreed upon enhancements	Two (2) weeks from start date
1.3	Write and test PySpark Notebooks to move data from Landing to Cleansed	Four (4) weeks from start date
1.4	Test results of identity matching output	Seven (7) weeks from start date
1.5	Write and test PySpark Notebooks to move data from Cleansed Zone and Matching Zone into Curated Zone	Seven (7) weeks from start date
1.6	Create a Notebook and save results for an existing report that includes applying archival rules (ODE CTE Annual Report)	October 25 <sup>th</sup> 2024
1.7	Create presentation layer from curated zone for researcher use	Ten (10) weeks from start date
1.8	Schedule pipelines to run Notebooks	Ten (10) weeks from start date
1.9	Create and build process to deploy Prod environment	Fourteen (14) weeks from start date
1.10	Document processes in Azure DevOps Wiki pages	Sixteen (16) weeks from start date
1.11	Bi-weekly Sprint Planning Meetings (60 min)	Every other Monday
1.12	Data team status check-ins (30 min)	Every other Monday; Weekly on Friday
1.13	Guided work sessions with Microsoft Cloud Engineer (up to 120 min)	Weekly on Thursday

Task due dates are subject to change according to change management structure/process identified by OLDC and terms and conditions of the Department of Administrative Services, Procurement Services Master Price and Services Agreement for Information Technology Professional Business Services.

-End of Attachment A to the RFQ-

# Attachment B

## Cost Offer

ITEM	# DESCRIPTION	ESTIMATED HOURS	COST
TASK	<b>X1</b>		
1.1	Onboarding (e.g., equipment setup and access)		
1.2	Perform code review of existing PySpark Notebooks and make agreed upon enhancements		
1.3	Write and test PySpark Notebooks to move data from Landing to Cleansed		
1.4	Test results of identity matching output		
1.5	Write and test PySpark Notebooks to move data from Cleansed Zone and Matching Zone into Curated Zone		
1.6	Create a Notebook and save results for an existing report that includes applying archival rules (ODE CTE Annual Report)		
1.7	Create presentation layer from curated zone for researcher use		
1.8	Schedule pipelines to run Notebooks		
1.9	Create and build process to deploy Prod environment		
1.10	Document processes in Azure DevOps Wiki pages		
1.11	Bi-weekly Sprint Planning Meetings (60 min)		
1.12	Data team status check-in's (30 min)		
1.13	Guided work sessions with Microsoft Cloud Engineer (up to 120 min)		
	Task 1 Total Maximum Cost:		
	MAXIMUM TOTAL COST:		
Consult	ant Firm Name:		
Authori	zed Signature:	Date:	_//
Printed Name of Authorized Representative:			
Title:			