

Project Investment Justification

Integrated Data System Development

E021001

Office of Economic Opportunity

Contents

1. General Information	2
2. Meeting Pre-Work	2
3. Pre-PIJ/Assessment	3
4. Project	4
5. Schedule	4
6. Impact	5
7. Budget	6
8. Technology	6
9. Security	10
10. Areas of Impact	11
11. Financials	13
12. Project Success	15
13. Conditions	16
14. Oversight Summary	16
15. PIJ Review Checklist	17

1. GENERAL INFORMATION

PIJ ID: EO21001

PIJ Name: Integrated Data System Development

Account: Office of Economic Opportunity

Business Unit Requesting: Office of Economic Opportunity

Sponsor: Ken Burns

Sponsor Title: COO, Arizona Commerce Authority

Sponsor Email: kenb@azcommerce.com

Sponsor Phone: (602) 845-1229

2. MEETING PRE-WORK

2.1 What is the operational issue or business need that the Agency is trying to solve? (i.e....current process is manual, which increases resource time/costs to the State/Agency, and leads to errors...):

To better understand and improve services, the state needs an integrated data system (IDS) to combine administrative data from various state agencies and educational institutions with the purpose of assessing the impact of those organizations on the lives of Arizonans over time.

2.2 How will solving this issue or addressing this need benefit the State or the Agency?

An IDS will allow the state to better measure the value of various government programs and interventions; to understand how individuals may be affected by more than one agency; and to refine government efforts to generate better outcomes.

The initial focus is to assess the efficacy of the education and workforce systems using the IDS. For example, an IDS with integrated education-workforce data can be used by state policymakers to design and fund programs that help people gain skills effectively and move into sustainable jobs; by workforce program managers to measure the effectiveness of their services based on employment outcomes; by community colleges and universities to benchmark and improve outcomes for their students; by parents and students to support informed decisions on education choices; and by researchers who want to know what type of services are most effective.

Future studies may entail including data about health, health care, incarceration and other topics, allowing the state to have a broader view of how state agencies intersect with and affect the lives of Arizonans.

2.3 Describe the proposed solution to this business need.

In this project, OEO will establish the IDS using the following data: (1) Education records from the Maricopa County Community College District; (2) Education records from the Pima County Community College District; (3) Participant records from the Workforce Innovation & Opportunity Act programs in the Arizona Department of Economic Security; (4) Participant records from the Adult Education program in the Arizona Department of Education; and (5) Wage records from the Unemployment Insurance data repository in OEO.

The IDS will be designed to be expandable so that the integration of data from additional sources can be done over time.

2.4 Has the existing technology environment, into which the proposed solution will be implemented, been documented?

Yes

2.4a Please describe the existing technology environment into which the proposed solution will be implemented.

2.5 Have the business requirements been gathered, along with any technology requirements that have been identified?

Yes

2.5a Please explain below why the requirements are not available.

3. PRE-PIJ/ASSESSMENT

3.1 Are you submitting this as a Pre-PIJ in order to issue a Request for Proposal (RFP) to evaluate options and select a solution that meets the project requirements?

No

3.1a Is the final Statement of Work (SOW) for the RFP available for review?

3.2 Will you be completing an assessment/Pilot/RFP phase, i.e. an evaluation by a vendor, 3rd party or your agency, of the current state, needs, & desired future state, in order to determine the cost, effort, approach and/or feasibility of a project?

No

3.2a Describe the reason for completing the assessment/pilot/RFP and the expected deliverables.

3.2b Provide the estimated cost, if any, to conduct the assessment phase and/or Pilot and/or RFP/solicitation process.

3.2e Based on research to date, provide a high-level cost estimate to implement the final solution.

4. PROJECT

4.1 Does your agency have a formal project methodology in place?

Yes

4.2 Describe the high level makeup and roles/responsibilities of the Agency, Vendor(s) and other third parties (i.e. agency will do...vendor will do...third party will do).

OEO will provide stewardship and funds for system development; and also a project manager to coordinate and oversee the project.

SpringML will develop the system on Google Cloud; and provide a technical project manager to guide the development life cycle.

Maricopa county community college district, Pima county community college district, ADE, DES and OEO will designate data stewards to work collaboratively with SpringML to design data repositories and data quality processes; designate information security liaisons to work with the security company to implement the risk management framework; and designate IT personnel to extract, transform and transmit data to the IDS.

4.3 Will a PM be assigned to manage the project, regardless of whether internal or vendor provided?

Yes

4.3a If the PM is credentialed, e.g., PMP, CPM, State certification etc., please provide certification information.

4.4 Is the proposed procurement the result of an RFP solicitation process?

No

4.5 Is this project referenced in your agency's Strategic IT Plan?

Yes

5. SCHEDULE

5.1 Is a project plan available that reflects the estimated Start Date and End Date of the project, and the supporting Milestones of the project?

No

5.2 Provide an estimated start and finish date for implementing the proposed solution.

Est. Implementation Start Date

Est. Implementation End Date

1/18/2021 12:00:00 AM

6/7/2021 12:00:00 AM

5.3 How were the start and end dates determined?

Other

5.3a List the expected high level project tasks/milestones of the project, e.g., acquire new web server, develop software interfaces, deploy new application, production go live, and estimate start/finish dates for each, if known.

Milestone / Task	Estimated Start Date	Estimated Finish Date
------------------	----------------------	-----------------------

Project Charter - develop scope, objectives, roles and responsibilities	12/14/20	12/31/20
Solution Design - develop final architecture and design from notional architecture in technical design document	01/02/21	01/18/21
Solution Development	01/18/21	05/03/21
Hardening Sprint - implement security controls, control enhancements, modifications based on security control assessment and continuous monitoring.	04/12/21	05/03/21
User Acceptance Testing	05/03/21	05/21/21
Production Rollout	05/21/21	06/07/21

5.4 Have steps needed to roll-out to all impacted parties been incorporated, e.g. communications, planned outages, deployment plan?

Yes

5.5 Will any physical infrastructure improvements be required prior to the implementation of the proposed solution. e.g., building reconstruction, cabling, etc.?

No

5.5a Does the PIJ include the facilities costs associated with construction?

5.5b Does the project plan reflect the timeline associated with completing the construction?

6. IMPACT

6.1 Are there any known resource availability conflicts that could impact the project?

No

6.1a Have the identified conflicts been taken into account in the project plan?

6.2 Does your schedule have dependencies on any other projects or procurements?

Yes

6.2a Please identify the projects or procurements.

Procuring services for implementing a risk management framework for the IDS using the statewide Knowledge Services contract. Based on the timeline from Knowledge Services, OEO will complete selection of a company by November 20.

6.3 Will the implementation involve major end user view or functionality changes?

No

6.4 Will the proposed solution result in a change to a public-facing application or system?

No

7. BUDGET

7.1 Is a detailed project budget reflecting all of the up-front/startup costs to implement the project available, e.g., hardware, initial software licenses, training, taxes, P&OS, etc.?

Yes

7.2 Have the ongoing support costs for sustaining the proposed solution over a 5-year lifecycle, once the project is complete, been determined, e.g., ongoing vendor hosting costs, annual maintenance and support not acquired upfront, etc.?

Yes

7.3 Have all required funding sources for the project and ongoing support costs been identified?

Yes

7.4 Will the funding for this project expire on a specific date, regardless of project timelines?

No

7.5 Will the funding allocated for this project include any contingency, in the event of cost over-runs or potential changes in scope?

Yes

8. TECHNOLOGY

8.1 Please indicate whether a statewide enterprise solution will be used or select the primary reason for not choosing an enterprise solution.

Other (please specify)

8.2 Will the technology and all required services be acquired off existing State contract(s)?

Yes

8.3 Will any software be acquired through the current State value-added reseller contract?

Yes

8.3a Describe how the software was selected below:

SPSS and SAS are statistical software used widely by social scientists and analysts in state government for data analysis. OEO consulted with its staff experienced in using the software and potential users of the IDS.

OEO will acquire software using CDW-G and the IBM passport advantage contract.

8.4 Does the project involve technology that is new and/or unfamiliar to your agency, e.g., software tool never used before, virtualized server environment?

Yes

8.5 Does your agency have experience with the vendor (if known)?

No

8.6 Does the vendor (if known) have professional experience with similar projects?

Yes

8.7 Does the project involve any coordination across multiple vendors?

Yes

8.8 Does this project require multiple system interfaces, e.g., APIs, data exchange with other external application systems/agencies or other internal systems/divisions?

No

8.9 Have any compatibility issues been identified between the proposed solution and the existing environment, e.g., upgrade to server needed before new COTS solution can be installed?

No

8.9a Describe below the issues that were identified and how they have been/will be resolved, or whether an ADOA-ASET representative should contact you.

8.10 Will a migration/conversion step be required, i.e., data extract, transformation and load?

Yes

8.11 Is this replacing an existing solution?

No

8.11a Indicate below when the solution being replaced was originally acquired.

8.11b Describe the planned disposition of the existing technology below, e.g., surplus, retired, used as backup, used for another purpose:

8.12 Describe how the agency determined the quantities reflected in the PIJ, e.g., number of hours of P&OS, disk capacity required, number of licenses, etc. for the proposed solution?

Agencies and colleges participating in the development of the IDS provided information on the data volume (size and rows) of both historical and ongoing data extracts that they intend on sending to the IDS. Vendors used this information to provide estimates of cloud service costs.

To correct assumptions made by vendors on frequency of operations, OEO and ASET provided clarification and sought revised estimates from all 3 vendors.

OEO's statisticians and demographers provided guidance on which SAS and IBM SPSS software modules are most likely useful for longitudinal analysis of IDS data. OEO determined the best licensing option by holding discussions with both companies. Based on the ongoing dialogue with potential users in the participating organizations, OEO may choose to acquire fewer licenses than shown in the financials section.

SpringML, the chosen vendor, provided development costs based on its experience with a similar project in the State of Utah.

The statewide information security and privacy office provided best estimates for implementing the risk management framework using a consulting company. This procurement process is underway.

ADOA's risk manager provided an estimate for a ten million dollar cyber insurance policy based on PII volume and informal quotes from insurance companies.

8.13 Does the proposed solution and associated costs reflect any assumptions regarding projected growth, e.g., more users over time, increases in the amount of data to be stored over 5 years?

No

8.14 Does the proposed solution and associated costs include failover and disaster recovery contingencies?

Yes

8.14a Please select why failover and disaster recovery is not included in the proposed solution.

8.15 Will the vendor need to configure the proposed solution for use by your agency?

No

8.15a Are the costs associated with that configuration included in the PIJ financials?

8.16 Will any app dev or customization of the proposed solution be required for the agency to use the project in the current/planned tech environment, e.g. a COTS app that will req custom programming, an agency app that will be entirely custom developed?

Yes

8.16a Will the customizations inhibit the ability to implement regular product updates, or to move to future versions?

No

8.16b Describe who will be customizing the solution below:

Upon evaluation of Google services for data matching, SpringML may determine standard open source libraries could provide better outcomes than using native Google services. If necessary, SpringML data scientists will develop a custom solution for data matching.

8.16c Do the resources that will be customizing the application have experience with the technology platform being used, e.g., .NET, Java, Drupal?

Yes

8.16d Please select the application development methodology that will be used:

Agile/Scrum

8.16e Provide an estimate of the amount of customized development required, e.g., 25% for a COTS application, 100% for pure custom development, and describe how that estimate was determined below:

Upon evaluation of Google services for data matching and examination of the data, SpringML may determine standard open source libraries could provide better outcomes. An estimate of the amount of customization is difficult. It could be 0% if Google services like Data Fusion are adequate, or it could be 25% of the entire system if an open source library is used for data matching.

8.16f Are any/all Professional & Outside Services costs associated with the customized development included in the PIJ financials?

Yes

8.17 Have you determined that this project is in compliance with all applicable statutes, regulations, policies, standards & procedures, incl. those for network, security, platform, software/application &/or data/info found at aset.az.gov/resources/psp?

Yes

8.17a Describe below the compliance issues that were identified and how they have been/will be resolved, or whether an ADOA-ASET representative should contact you:

8.18 Are there other high risk project issues that have not been identified as part of this PIJ?

No

8.18a Please explain all unidentified high risk project issues below:

9. SECURITY

9.1 Will the proposed solution be vendor-hosted?

No

9.1a Please select from the following vendor-hosted options:

Commercial data center environment, e.g AWS, Azure

9.1b Describe the rationale for selecting the vendor-hosted option below:

9.1c Has the agency been able to confirm the long-term viability of the vendor hosted environment?

Yes

9.1d Has the agency addressed contract termination contingencies, e.g., solution ownership, data ownership, application portability, migration plans upon contract/support termination?

Yes

9.1e Has a Conceptual Design/Network Diagram been provided and reviewed by ASET-SPR?

No

9.1f Has the spreadsheet located at <https://aset.az.gov/arizona-baseline-security-controls-excel> already been completed by the vendor and approved by ASET-SPR?

No

9.2 Will the proposed solution be hosted on-premise in a state agency?

No

9.2a Where will the on-premise solution be located:

9.2b Were vendor-hosted options available and reviewed?

9.2c Describe the rationale for selecting an on-premise option below:

9.2d Will any data be transmitted into or out of the agency's on-premise environment or the State Data Center?

9.3 Will any PII, PHI, CGIS, or other Protected Information as defined in the 8110 Statewide Data Classification Policy be transmitted, stored, or processed with this project?

Yes

9.3a Describe below what security infrastructure/controls are/will be put in place to safeguard this data:

A security company that is separate from SpringML will: (1) conduct a privacy impact assessment; (2) develop and document system security architecture and design; (3) conduct (and document) initial and periodic security risk assessments; (4) select the baseline security controls for moderate-impact systems pursuant to state requirements (AZRAMP), add additional security controls as necessary, tailor the selected controls, document the security controls in a system security plan, and periodically update the system security plan; (5) develop, document and assist OEO in implementing a continuous monitoring strategy for the IDS; (6) conduct security control assessments; (7) conduct penetration testing; (8) prepare and maintain a plan of action and milestones document based on assessments and activities like continuous monitoring; and (9) develop an incident response plan.

10. AREAS OF IMPACT

Application Systems

New Application Development

Database Systems

Other

BigQuery

Software

Other

Google cloud services

Hardware

Other

Google cloud services

Hosted Solution (Cloud Implementation)

Other

Google Cloud

Security

Encryption;Firewall;Intrusion Detection System (IDS);Intrusion Prevention System (IPS);Security Controls/Systems - Other

AZRAMP baseline security controls for moderate impact systems

Telecommunications

Enterprise Solutions

OEO and ASET jointly explored the option of creating a statewide enterprise solution using the IDS project, but found the solution could not be supported with the available OEO operational budget.

11. FINANCIALS

Description	PIJ Category	Cost Type	Fiscal Year Spend	Quantity	Unit Cost	Extended Cost	Tax Rate	Tax	Total Cost
Project Charter	Professional & Outside Services	Development	1	1	\$60,000	\$60,000	0.00 %	\$0	\$60,000
Project Manager (\$85/hr; 1/10/2020 to 6/1/2021)	Professional & Outside Services	Development	1	1	\$99,280	\$99,280	0.00 %	\$0	\$99,280
Solution Design	Professional & Outside Services	Development	1	1	\$180,000	\$180,000	0.00 %	\$0	\$180,000
Hardening Sprint	Professional & Outside Services	Development	1	1	\$90,000	\$90,000	0.00 %	\$0	\$90,000
Solution Development	Professional & Outside Services	Development	1	1	\$180,000	\$180,000	0.00 %	\$0	\$180,000
User Acceptance Testing	Professional & Outside Services	Development	1	1	\$150,000	\$150,000	0.00 %	\$0	\$150,000
Production Migration	Professional & Outside Services	Development	1	1	\$78,000	\$78,000	0.00 %	\$0	\$78,000
Risk Management Framework Implementation (estimate)	Professional & Outside Services	Development	1	1	\$100,000	\$100,000	0.00 %	\$0	\$100,000
Annual Cloud Service Cost	Other	Development	1	1	\$51,758	\$51,758	860.00 %	\$4,451	\$56,209
SPSS software-25 licenses	License & Maintenance Fees	Development	1	1	\$205,828	\$205,828	860.00 %	\$17,701	\$223,529
SAS software-10 licenses	License & Maintenance Fees	Development	1	1	\$83,480	\$83,480	860.00 %	\$7,179	\$90,659
Annual Cloud Service Cost	Other	Operational	2	1	\$51,758	\$51,758	860.00 %	\$4,451	\$56,209
Managed Services	Professional & Outside Services	Operational	2	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
SAS software maintenance	License & Maintenance Fees	Operational	2	1	\$22,090	\$22,090	860.00 %	\$1,900	\$23,990

Change Management & User Training	Professional & Outside Services	Operational	2	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
SPSS software maintenance (20% price + 10% year over year)	License & Maintenance Fees	Operational	2	1	\$45,282	\$45,282	860.00 %	\$3,894	\$49,176
Annual Security Control Assessment (estimate)	Professional & Outside Services	Operational	2	1	\$30,000	\$30,000	0.00 %	\$0	\$30,000
Annual Security Control Assessment (estimate)	Professional & Outside Services	Operational	3	1	\$30,000	\$30,000	0.00 %	\$0	\$30,000
SPSS software maintenance (20% price + 10% year over year)	License & Maintenance Fees	Operational	3	1	\$49,810	\$49,810	860.00 %	\$4,284	\$54,094
SAS software maintenance	License & Maintenance Fees	Operational	3	1	\$22,090	\$22,090	860.00 %	\$1,900	\$23,990
Change Management & User Training	Professional & Outside Services	Operational	3	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
Managed Services	Professional & Outside Services	Operational	3	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
Annual Cloud Service Cost	Other	Operational	3	1	\$51,758	\$51,758	860.00 %	\$4,451	\$56,209
Annual Security Control Assessment (estimate)	Professional & Outside Services	Operational	4	1	\$30,000	\$30,000	0.00 %	\$0	\$30,000
SPSS software maintenance (20% price + 10% year over year)	License & Maintenance Fees	Operational	4	1	\$54,791	\$54,791	860.00 %	\$4,712	\$59,503
SAS software maintenance	License & Maintenance Fees	Operational	4	1	\$22,090	\$22,090	860.00 %	\$1,900	\$23,990
Change Management & User Training	Professional & Outside Services	Operational	4	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
Managed Services	Professional & Outside Services	Operational	4	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600

Annual Cloud Service Cost	Other	Operational	4	1	\$51,758	\$51,758	860.00 %	\$4,451	\$56,209
Annual Security Control Assessment (estimate)	Professional & Outside Services	Operational	5	1	\$30,000	\$30,000	0.00 %	\$0	\$30,000
Annual Cloud Service Cost	Other	Operational	5	1	\$51,758	\$51,758	860.00 %	\$4,451	\$56,209
Managed Services	Professional & Outside Services	Operational	5	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
Change Management & User Training	Professional & Outside Services	Operational	5	1	\$39,600	\$39,600	0.00 %	\$0	\$39,600
SAS software maintenance	License & Maintenance Fees	Operational	5	1	\$22,090	\$22,090	860.00 %	\$1,900	\$23,990
SPSS software maintenance (20% price + 10% year over year)	License & Maintenance Fees	Operational	5	1	\$60,270	\$60,270	860.00 %	\$5,183	\$65,454

Base Budget (Available)	Base Budget (To Be Req)	Base Budget % of Project
\$0	\$0	0%
APF (Available)	APF (To Be Req)	APF % of Project
\$0	\$0	0%
Other Appropriated (Available)	Other Appropriated (To Be Req)	Other Appropriated % of Project
\$0	\$0	0%
Federal (Available)	Federal (To Be Req)	Federal % of Project
\$0	\$0	0%
Other Non-Appropriated (Available)	Other Non-Appropriated (To Be Req)	Other Non-Appropriated % of Project
\$2,293,500	\$0	100%

Total Budget Available	Total Development Cost
\$2,293,500	\$1,307,677
Total Budget To Be Req	Total Operational Cost
\$0	\$985,823
Total Budget	Total Cost
\$2,293,500	\$2,293,500

12. PROJECT SUCCESS

Please specify what performance indicator(s) will be referenced in determining the success of the proposed project (e.g. increased productivity, improved customer service, etc.)? (A minimum of one performance indicator must be specified)

Please provide the performance objective as a quantifiable metric for each performance indicator specified.

Note: The performance objective should provide the current performance level, the performance goal, and the time period within which that performance goal is intended to be achieved. You should have an auditable means to measure and take corrective action to address any deviations.

Example: Within 6 months of project completion, the agency would hope to increase "Neighborhood Beautification" program registration by 20% (3,986 registrants) from the current registration count of 19,930 active participants.

Performance Indicators

Within 6 months of project completion, the IDS will be used to produce 3 research products approved by the Workforce Data Task Force.

Note: The Workforce Data Task Force is a statutory governance body (A.R.S. 41-5404) that oversees development and operations of the IDS.

13. CONDITIONS

Conditions for Approval

Should development costs exceed the approved estimates by 10% or more, or should there be significant changes to the proposed technology scope of work or implementation schedule, the Agency must amend the PIJ to reflect the changes and submit it to ADOA-ASET, and ITAC if required, for review and approval prior to further expenditure of funds.

14. OVERSIGHT SUMMARY

Project Background

The Arizona Office of Economic Opportunity (OEO) exists to consolidate and improve analytic capacity for the state with the goal of tracking Arizona's key economic indicators. Its mission is to generate, collect, store, analyze and share data efficiently, safely and reliably using modern techniques and tools. In order to better understand and improve services, the state seeks to implement an integrated data system (IDS) to combine administrative data from various state agencies and educational institutions, with the purpose of assessing the impact of those organizations on the lives of Arizonans over time. With this project, the Arizona Office of Economic Opportunity will establish the IDS using data from community college education records, participant records from certain Arizona Department of Economic Security and the Arizona Department of Education programs, and wage records from the Unemployment Insurance data repository. The IDS will be designed to be expandable so that the integration of data from additional sources can be completed over time.

Business Justification

The IDS will enable the state to better measure the value of various government programs and interventions; to understand how individuals may be affected by more than one agency; and to refine government efforts to generate better outcomes. The initial focus of the implementation will be to assess the efficacy of the education and workforce systems. Integrating education and workforce data with the IDS will help enable state policymakers to design and fund programs that help people gain skills effectively and move into sustainable jobs; workforce program managers to measure the effectiveness of their services based on employment outcomes; community colleges and universities to benchmark and improve outcomes for their students; parents and students to support informed decisions on education choices; and researchers to better understand which services are the most effective. Within six months of completing the project, the IDS will be used to produce three new research products approved by the Workforce Data Task Force, a statutory governance body (A.R.S. 41-5404) that oversees development and operation of the IDS.

Implementation Plan

The solution will be hosted in the OEO Google Cloud Platform and OEO will be responsible for the security of the environment. The Arizona Office of Economic Opportunity will provide stewardship, funds for system development, and a project manager to coordinate and oversee the project. The vendor, SpringML, will be

responsible for system development and will provide a technical project manager to guide the development life cycle. Maricopa county community college district, Pima county community college district, ADE, DES and OEO will designate data stewards to work collaboratively with the vendor to design data repositories and data quality processes; designate information security liaisons to work with the security company to implement the risk management framework; and designate IT personnel to extract, transform and transmit data to the IDS.

Vendor Selection

Quotes and proposals were received from preferred vendors selected by the cloud service providers GCP (Google Cloud Platform), AWS (Amazon Web Services) and Microsoft Azure. Vendor proposals evaluated during the selection process included submissions from the Nerderly (AWS), Tallan (Microsoft) and SpringML (GCP). SpringML was ultimately selected as the chosen vendor for the project.

Budget or Funding Considerations

The project will be funded using available Other Non-Appropriated funds.

15. PIJ REVIEW CHECKLIST

Agency Project Sponsor
Ken Burns

Agency CIO (or Designee)
Chris Kontz

Agency ISO (or designee)
Chris Kontz

OSP Representative

ASET Engagement Manager

ASET SPR Representative
Thomas Considine

Agency SPO Representative
Eric Bell

Agency CFO
Ken Burns