



Leidos Biomedical Research, Inc.

Frederick National Laboratory for Cancer Research

April 28, 2019

Mr. Scott Keasey, Contracting Officer

Dr. Toby Hecht, COR  
9609 Medical Center Drive  
Bethesda, MD 20892

**Reference:** Contract HHSN261201500003I

**Subject:** Task Order HHSN26100076

**NCI Action:** Review and Acceptance of Task Order Deliverable

Dear Dr. Hecht:

In accordance with the above referenced contract and task order, the deliverable summarized below is provided for your review and acceptance.

**Table 1: Deliverable Summary**

<b>Task Order Number:</b>	HHSN26100076	<b>Project Title:</b>	Development of an Integrated Canine Data Commons (ICDC)
<b>Deliverable Item Number:</b>	2	<b>Deliverable Description:</b>	Quarterly CSP Report
<b>Reporting Period:</b>	1/21/2019 - 4/20/2019	<b>Quantity:</b>	1
<b>Primary Program Manager (PPM):</b>	John Otridge	<b>Contracting Officer's Representative (COR):</b>	Toby Hecht
<b>PPM Email:</b>	<a href="mailto:John.Otridge@nih.gov">John.Otridge@nih.gov</a>	<b>COR Email:</b>	<a href="mailto:Toby.Hecht@nih.gov">Toby.Hecht@nih.gov</a>
<b>PPM Phone:</b>	240.276.5653	<b>COR Phone:</b>	301.435.9162

Respectfully,

Connie Suders  
Contract Administrator

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

Project Information									
<b>Project Title</b>	HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)					<b>Project Overall Status: RYG</b>	<b>G</b>		
<b>Project Description and Deliverables</b>	The objective of this project is to leverage the Center for Biomedical Informatics and Information Technology's (CBIIT) NCI Cancer Research Data Commons (CRDC) experience and knowledge, and its development of Data Commons Framework Services (DCFS), to create a new, dynamic data commons for canine cancer data, including not only clinical outcomes and genomics findings from canine clinical trials being conducted by the Comparative Oncology Program (COP) in collaboration with NCI's Division of Cancer Treatment and Diagnosis (DCTD), but also the trials' molecular, pharmacological, microenvironment, medical imaging and other study data. Reporting deliverables include quarterly CSP reports and monthly meeting minutes.								
<b>LBR PM</b>	Matthew Beyers		<b>LBR Directorate</b>	BIDS/ADRD	<b>LBR Change Control Rep</b>	Eric Stahlberg			
<b>Total Funded Amount</b>	\$1,959,337		<b>Project Type</b>	Applied/Clinical	<b>Tier</b>	3	<b>Period of Performance</b>	2018-09-24 to 2020-09-23	
<b>PID</b>	<b>Milestone Planned Amount</b>		<b>LBR Project Expenses to Date</b>		<b>LBR Open Obligations</b>		<b>LBR Project Costs Invoiced to Government</b>		
400.041.0076.0001.001	\$1,959,336.71		\$140,819.21		\$60,582.45		\$112,424.31		
<b>Total as of:</b>	3/29/2019		140,819.21		\$60,582.45		\$112,424.31		
<b>Percent Spent:</b>	7%				<b>Percent Committed:</b>	10%			
<b>Milestone No. and Name</b>	<b>Description</b>						<b>POP</b>		
							<b>Start Date</b>	<b>End Date</b>	
1 – Base: Complete Prototype	Initial and incremental development of a prototype ICDC using existing data and implement						9/24/2018	9/23/2020	
<b>LBR Subcontracts Administrator</b>	<b>Name</b>			<b>Email</b>			<b>Phone</b>		
	Nick D'Abbraccio			<a href="mailto:dabbraccionn@nih.gov">dabbraccionn@nih.gov</a>			301-228-4323		
<b>Subcontractor or Supplier</b>					<b>Subcontract Amount</b>				
Essential Software, Inc.					\$112,608				

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

Project Status		
Assessment Type	Current Status	Future Plans
<p>Technical Scope and Status</p>	<p>System Infrastructure:</p> <ul style="list-style-type: none"> <li>-Completed install and testing of full-Gen3 system and demoed to NCI CBIIT staff. Determined that the stack (including the database) is insufficient for ICDC purposes and explained why.</li> <li>-Developed budget and staffing plan for software development to deliver minimum viable product of ICDC by Oct. 2019 with continued development effort until May 2020.</li> <li>-Submitted contract modification to acquire staff for development.</li> <li>-Began exploring mapping of nodes/edges to BRIDG model.</li> <li>-Began exploring NiFi as alternative method for ETL.</li> <li>-Continued meeting with Data Commons Framework staff to understand Fence/IndexD and how ICDC will integrate with them.</li> </ul> <p>Data:</p> <ul style="list-style-type: none"> <li>-Completed ETL of COTOC007b trial data, not including laboratory values, using Pentaho transformations.</li> <li>-Modified data model as needed using Model-Tool, developed by FNL staff.</li> <li>-Met with NCATS team and acquired their data structure. Began mapping this into the graph data model and scoping out changes needed. While the actual data for this study will not be available until NCATS has finalized patent and publishing issues, the system will be ready to accept the data when available.</li> <li>-Began developing use cases and user stories based on Steering Committee feedback.</li> </ul>	<p>System Infrastructure:</p> <ul style="list-style-type: none"> <li>-Will develop a true graph database structure and associated software interfaces for querying and loading of the data.</li> <li>-Design principles include: being database agnostic, being data model agnostic, being able to submit data, being able to query all data in the system, having a userfriendly and informative interface and having each of the models being independent of one another by use of standard APIs so as to avoid the tight integration seen with Gen3.</li> </ul> <p>Data:</p> <ul style="list-style-type: none"> <li>-Will continue to implement data structure for NCATS and build out the rest of the data structure for COTOC007b.</li> <li>-Will continue to be available to take on new data sets as they become available as decided by the Steering Committee.</li> </ul>

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

Project Status		
Assessment Type	Current Status	Future Plans
	<p>-Met with Seven Bridges Genomics to begin to understand how to connect their system to ours and how to help our Steering Committee develop and use tools in their system. They are currently the only one of the Cloud Resources on the Amazon platform (as is ICDC) and so it makes sense to integrate with them first.</p> <p>Steering Committee:</p> <p>-First Steering Committee meeting was held Feb. 20, 2019 and was considered successful.</p> <p>-A Data Governance Board is being developed to make decisions about which data sets with which priority will be entered into ICDC during the prototype phase.</p> <p>-A chair of the Steering Committee was identified and has been asked to serve.</p> <p>-The final composition of the Steering Committee is 11 external members and 9 NCI members with 3 FNL staff serving in a support role.</p> <p>-Second meeting was held on April 10, 2019 and discussed SC needs and wants from both computational and information points of view.</p> <p>Other:</p> <p>-Project Management Plan was completed, submitted for approval and initially reviewed.</p>	<p>Steering Committee</p> <p>-Establish Data Governance Board and process and then prioritize data sets to be loaded to ICDC.</p> <p>-Provide use cases and user stories to ICDC staff to be implemented into the system.</p>

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

Project Status												
Assessment Type	Current Status	Future Plans										
	ADRD/HR: <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td colspan="2" style="text-align: center;">Task Order 76 Staffing Summary</td> </tr> <tr> <td style="width: 80%;">Filled</td> <td style="text-align: right;">0</td> </tr> <tr> <td>Recruiting</td> <td style="text-align: right;">0</td> </tr> <tr> <td>Open</td> <td style="text-align: right;">0</td> </tr> <tr> <td style="text-align: right;">Total</td> <td style="text-align: right;">0</td> </tr> </table>	Task Order 76 Staffing Summary		Filled	0	Recruiting	0	Open	0	Total	0	
Task Order 76 Staffing Summary												
Filled	0											
Recruiting	0											
Open	0											
Total	0											
Schedule Milestones and Status	The project is still on track with regards to period of performance and schedule.	The project will pivot to mostly software development for the next several months and is expected to produce a minimum viable product by end of Sept. 2019.										
Cost Status	The project is currently underspending.	Because the project has underspent up to this point, there are sufficient funds to undertake the software development phase as outlined above.										
Terms and Conditions	No change.	No change expected.										
Assumptions	No change.	No change expected.										
Subcontractor Status	The subcontractor has performed at or above expectations and is within budget and schedule.	Because of excellent past performance, we expect this subcontractor to continue to provide the same level of effort as we move forward and have initiated a modification to the subcontract to add more funds for more staff to enable us to pivot to software development.										
Risk Status	The Gen3 tech stack maturity risk is now realized, and we have now implemented the mitigation strategy by increasing software development. There is a risk that, because we must now pivot to intense software development, that we will consume project reserves towards the end of the	See risk assessment below.										

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
Cost/Schedule/Performance Quarterly Report  
April 2019

<b>Project Status</b>		
<b>Assessment Type</b>	<b>Current Status</b>	<b>Future Plans</b>
	period of performance. We believe there are enough reserves to mitigate this risk. Should the project not need the reserves it has set aside, additional functionality will be added so that the project finishes within budget.	

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

**Cost Status Overview**

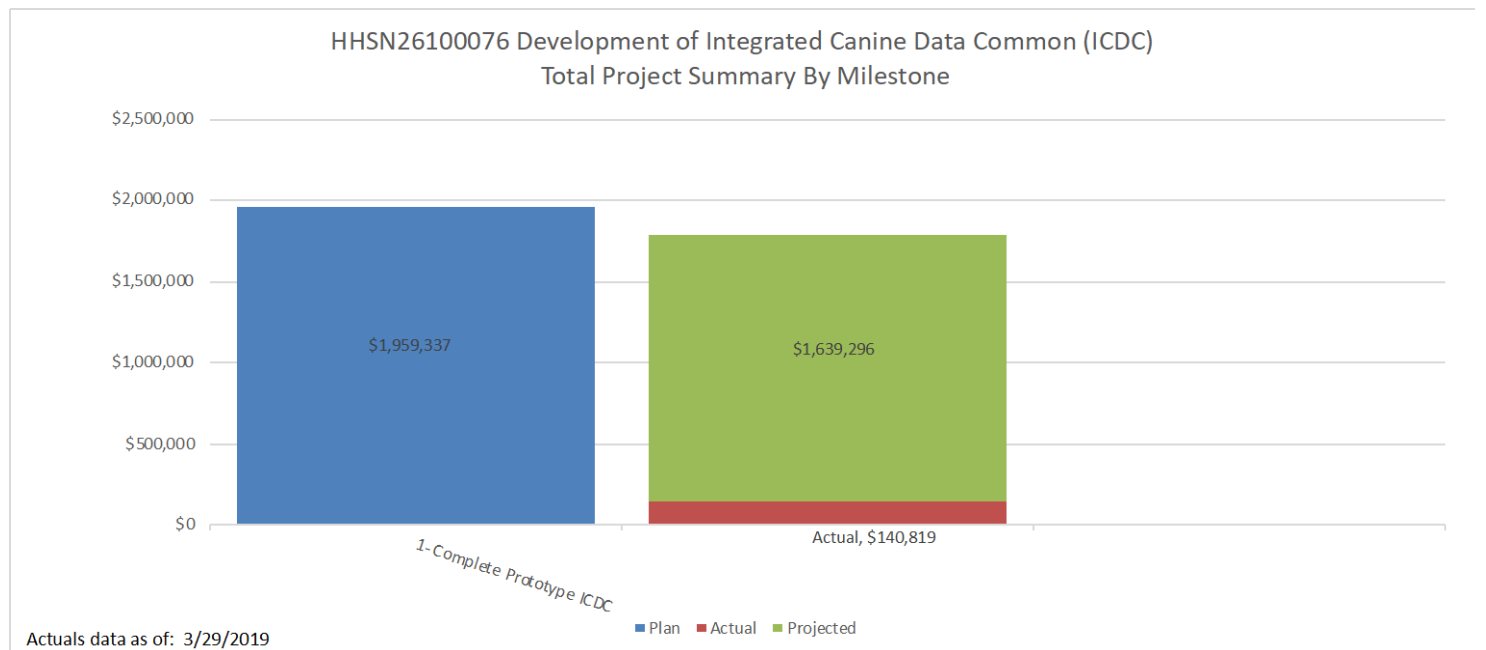


Figure 1. Bar graph shows Planned Spend compared to Actual and Projected expenses by milestone as of April 2019.

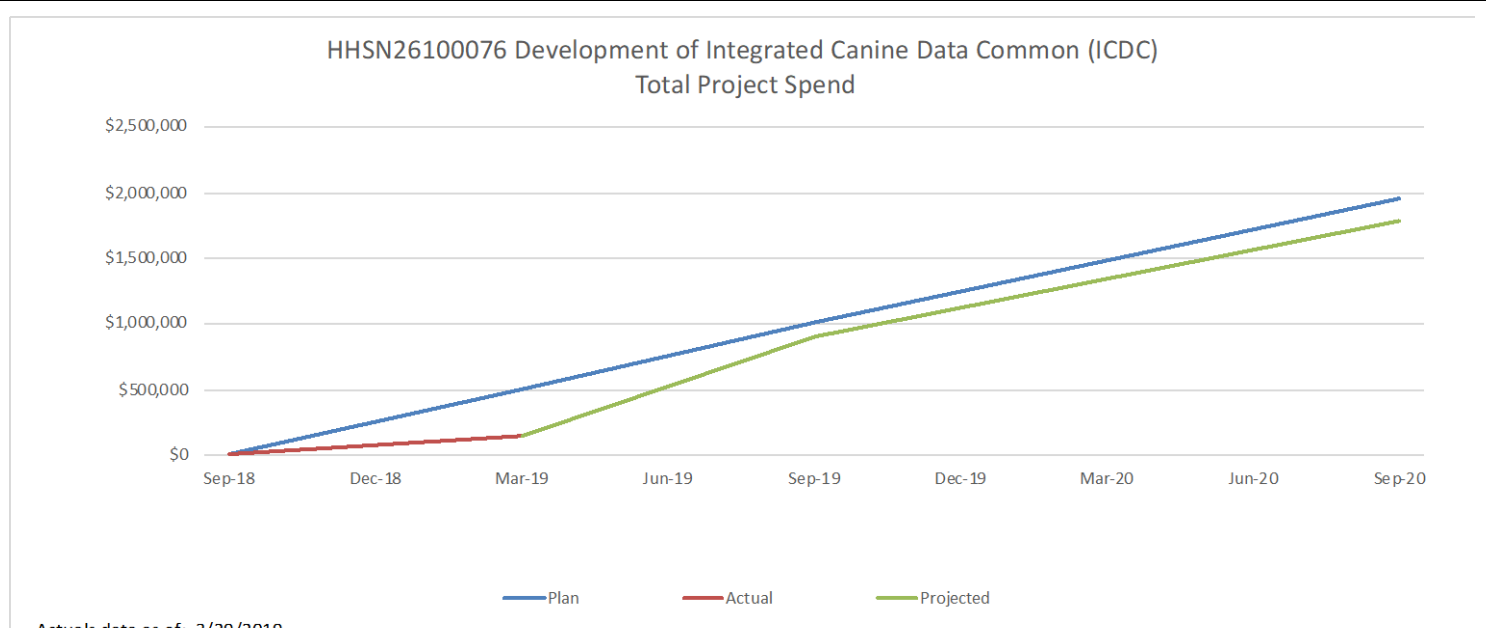


Figure 2. Line Graph showing the Actual and Projected costs compared to the Planned Spend by month and year as of April 2019.

**Project Performance Status**

Assessment Area	Past	Present	Future	Comments
Overall Assessment	G	G	G	

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

Project Performance Status				
Assessment Area	Past	Present	Future	Comments
Technical/Scientific	G	G	G	Demonstrated or projected ability to meet all technical metrics and no open unresolved technical issues. Project has finished initial evaluation and is moving into development phase to provide final product.
Schedule	G	G	G	Ability (actual and projected) to meet all schedule milestones.
Cost	G	G	G	Costs are being tracked and projected to show actuals versus plan/forecast. Project has significantly underspent to this point in anticipation of needing to do software development and is in a good position to afford the work.
Contract	G	G	G	Change Control Board running well and managing technical direction changes. And no significant contractual issues.
Subcontractors & Suppliers	G	G	G	Demonstrated or projected ability for supplier to meet all technical metrics.
Customer Environment	G	G	G	Customer perceptions aligned with PM perceptions. Continue to hold regular meetings with CBIT and DCTD staff to assure alignment.
Team Compliance & Fraud Concerns	G	G	G	No unusual circumstances that would give rise to fraud/corruption concerns.
Staffing	G	G	G	All key positions filled; no significant staffing shortfalls. Project team working effectively together. Good line management and functional support.
Infrastructure & Facilities	G	G	G	No Infrastructure needs.
Data Security	G	G	G	Required security and privacy plans current, self-assessment has been completed, employees have completed required training.
Risk				
Accepted or Realized Risks & Impact				



HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

<b>Risk</b>	
	<ul style="list-style-type: none"> <li>• Gen3 architecture took longer than expected to instantiate the shell, however, this was still within project schedule boundaries and effected a negligible impact. We worked through the issues encountered with UChicago. In particular, the code was less mature than desired, however, this was anticipated and the time required fell within our scheduling parameters. Probability: High; Impact: Minor; Mitigation: Software development to customize as needed using existing budget which was anticipated to be needed for this purpose.</li> <li>• The Sheepdog code in Gen3 is very specific to the GDC data model. We will have to iteratively modify our data model YAML to get Sheepdog to accept it. This may cause a delay in loading the model and therefore loading the data to the system. Probability: High; Impact: Minor; Mitigation: Software development to customize as needed using existing budget anticipated for this purpose.</li> <li>• The Gen3 architecture is still new and not extensively documented or field tested. It is possible there are missing elements needed to fully support the ICDC or there are performance gaps in functionality or stability. Probability: Low; Impact: Moderate; Mitigation: Gen3 was assembled by UChicago so by using UChicago as SME(s) during development we can be guided by them with respect to what features are incomplete and if there are roadmaps to complete those features. This will enable us to determine prioritization of ICDC system development to avoid any known issues and to plan around Gen3 development releases.</li> </ul>
<b>Open Red Risks &amp; Mitigation Plans</b>	
	<ul style="list-style-type: none"> <li>• The known Use Cases may only be a small fraction of the Use Cases the community requires. As such, our level of efforts estimates may not be enough to cover the effort required to meet the new use cases. Probability: High; Impact: Minor; Mitigation: Frequent communication with the NCI program leadership to prioritize Use Cases to use in the Prototyping and Production stages.</li> </ul>
<b>Open Yellow Risks &amp; Mitigation Plans</b>	
	<ul style="list-style-type: none"> <li>• The level of detail in the SOW is low and the Data Commons concept is new. So, there are a lot of unknowns that will only be encountered during implementation. So, this adds a lot of uncertainty to the timelines and the effort estimates. Probability: Medium; Impact: Moderate; Mitigation: Focus on uncovering those unknowns during the Prototyping stage so they do not arise late in the project at Production. At the completion of the Prototyping phase we will conduct an assessment of costs and schedule for the development of the Production system.</li> </ul>
<b>Open Green Risks</b>	

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
Cost/Schedule/Performance Quarterly Report  
April 2019

<b>Risk</b>	
	<ul style="list-style-type: none"><li>• The Gen3 architecture is still new and not extensively documented or field tested. It is possible there are missing elements needed to fully support the ICDC or there are performance gaps in functionality or stability. Probability: Low; Impact: Moderate; Mitigation: Gen3 was assembled by UChicago so by using UChicago as SME(s) during development we can be guided by them with respect to what features are incomplete and if there are roadmaps to complete those features. This will enable us to determine prioritization of ICDC system development to avoid any known issues and to plan around Gen3 development releases.</li><li>• Amount of data to be stored is larger than the free-storage can handle, so could exceed our estimated costs. Probability: Low; Impact: Moderate; Mitigation: Work with the NCI programs to identify this issue if it arises and evaluate options before implementing a solution.</li><li>• Unable to staff the project in a timely fashion with either/or FNL or subcontractor staff. This could delay progress towards meeting milestones. Probability: Low; Impact: Moderate; Mitigation: The initial phases will focus on activities such as data inventory, harmonization and use case definition that utilize existing or soon to be hired staff (anticipated to be onboard before project starts). This will allow time to find any additional staff or subcontractors to staff up.</li></ul>
<b>Open Issues, Action Items and Resolution Plans</b>	
	No Open Issues

HHSN26100076: Development of an Integrated Canine Data Commons (ICDC)  
 Cost/Schedule/Performance Quarterly Report  
 April 2019

