

**I-Catalyst Program - Project 1: OPHSS/CSELS Data Hub**

GenIC Submission under OMB #0920-1158

Juliana K. Cyril, MPH, PhD  
Director, Office of Technology and Innovation  
Office of the Associate Director for Science  
Centers for Disease Control and Prevention  
Ph: 404-639-4639  
Fax: 404-639-4903

Team Lead – Veer Pawate, CDC/OPHSS/CSELS/DHIS

Submission Date:

4-24-2017

OMB Project Number: 0920-1158

## Contents

A. Justification.....	3
1. Circumstances Making the Collection of Information Necessary.....	3
2. Purpose and Use of Information Collection.....	3
3. Use of Improved Information Technology and Burden Reduction.....	4
4. Efforts to Identify Duplication and Use of Similar Information.....	4
5. Impact on Small Businesses or Other Small Entities.....	4
6. Consequences of Collecting the Information Less Frequently.....	4
7. Special Circumstances Relating to the Guidelines of 5 CFR 1320.5.....	4
8. Comments in Response to the Federal Register Notice and Efforts to Consult Outside the Agency.....	4
9. Explanation of Any Payment or Gift to Respondents.....	5
10. Assurance of Confidentiality Provided to Respondents.....	5
11. Justification for Sensitive Questions.....	5
12. Estimates of Annualized Burden Hours and Costs.....	5
13. Estimates of Other Total Annual Cost Burden to Respondents or Record Keepers.....	6
14. Annualized Cost to the Government.....	6
15. Explanation for Program Changes or Adjustments.....	6
16. Plans for Tabulation and Publication and Project Time Schedule.....	7
17. Reason(s) Display of OMB Expiration Date is Inappropriate.....	7
18. Exceptions to Certification for Paperwork Reduction Act Submissions.....	7

## GenIC Package & Attachments

1. NOA
2. Supporting Statement A
3. ICR PRA Part II Worksheet
4. Att. 1: I-Cat Interview Protocol Guide and Questions

- CDC's Division of Health Informatics and Surveillance (DHIS) in the Center for Surveillance, Epidemiology, and Laboratory Services maintains a Data Hub of externally-acquired data sources [American Hospital Association (AHA), Centers for Medicare and Medicaid Services (CMS), Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project (HCUP), and Truven Health Analytics MarketScan® data] for agency use. DHIS coordinates agency-wide purchases and acquisitions of external data sources, and provides services to facilitate the use of these data. Evolving technology creates opportunities for adding innovative services (e.g., visualization) to the existing portfolio to better meet data user needs. The team will explore whether data users and other stakeholders have a need for additional services, such as visualization, to enhance data access and availability.
- Teams will use convenience sampling methods to select subjects who are readily available and within close proximity.

and policies that support the CDC/ATSDR. Federal scientific agencies, like the CDC, rely on research and scientific findings to help them develop solutions to public health problems which ultimately are disseminated to customers and stakeholders for adoption and use. However, anecdotal and empirical data show that many well-meaning, robust solutions are never used or adopted by the intended customer. This has resulted in a path littered with failed and unused government sponsored outputs. The CDC I-Catalyst program guides participants through a “customer discovery” process aimed at helping teams with a new solution to identify their customers. This is done by taking a team’s main assumptions about who their customer is, the exact problem they are solving for the customer, and how the customer wants to receive or use the solution from the team—and turning those assumptions into hypotheses which the teams will then test (mainly through interviews with potential customers). Only conversations with potential customers (stakeholders) can provide the facts from which hypotheses are proven or disproven about whether a solution (product, process, etc.) creates value for the intended beneficiaries. It is expected that participants will leave the program with the ability to evaluate and translate their insights into solutions that have high levels of efficacy and user acceptability. The information collection is necessary to guide CDC project teams to create usable solutions that are customer centric and meaningful to users, whether it’s adhering to recommendations, policies, protocol or interventions.

This request seeks approval for a GenIC approval for subproject I-Catalyst – OPHSS/CSELS Data Hub. The ultimate goal of the I-Catalyst Project OPHSS/CSELS Data Hub is to explore opportunities for future epidemiology work (including data file content visualization), communication activities with data users and other stakeholders, evaluation efforts, and IT options for easy data access and availability. The results of the project will help CDC Data Hub team to make the case for support to advance important improvements and solutions of data tool acquisitions, customization and usages. The efforts of CDC activities is authorized under Section 301 of the Public Health Service Act 42 U.S.C.241.

## 2. Purpose and Use of Information Collection

Data CDC’s Division of Health Informatics and Surveillance (DHIS) in the Center for Surveillance, Epidemiology, and Laboratory Services (CSELS) maintains a Data Hub of externally-acquired data sources [American Hospital Association (AHA), Centers for Medicare and Medicaid Services (CMS), Agency for Healthcare Research and Quality Healthcare Cost and Utilization Project (HCUP), and Truven Health Analytics MarketScan® data] for agency use. DHIS coordinates agency-wide purchases and acquisitions of external data sources, and provides services to facilitate the use of these data. Evolving technology creates opportunities for adding innovative services (e.g., visualization) to the existing portfolio to better meet data user needs. Recently, Exploratory Data Analysis (EDA), an approach for summarizing and visualizing the characteristics of a data set, was conducted for one of the data sets managed through the Data Hub and an online interactive and dynamic query visualization

system was developed using MS Excel 2013 and Excel's add-on Power Pivot. The project will explore opportunities for future epidemiology work (including data file content visualization), communication activities with data users and other stakeholders, evaluation efforts, and IT options for easy data access and availability.

The project hopes to seek input on data tool usage through I-Catalyst and determine how best to support the needs of scientific researchers and data visualization. The information collected will be used for internal CDC decision making purposes and to provide suggestions for improving data hub services in support of development of public health solutions. The customer interviews with Public health researchers at CDC (epidemiologists, statisticians, Health Scientist who are new or experience users of scientific data) and potentially external partners such as data providers and state and local health departments.

### 3. Use of Improved Information Technology and Burden Reduction

The interviews will be conducted in person, on-site or by virtual video conferencing like Skype for Business or Adobe Connect (Att. 1 - Interview guide). Using formative interview protocols allows the interviewer to follow the respondent's lead during in-person conversations. This wouldn't be possible if a list of fixed questions were used. This also is not possible if automated, technological-based collection techniques, such as a web-based survey, are used. On-site, in-person interviews allow interviewers to establish rapport with respondents and produce visual cues for interpreting responses that may require further probing or clarification. However, there are instances where teams can use improved information technology such as Skype or video conferencing for interviews to reduce the burden and provide flexibility in responder's schedule.

### 4. Efforts to Identify Duplication and Use of Similar Information

This is a unique I-Catalyst project and a new proposed solution. Other than proprietary business databases, there is no existing database that can provide the level of detail about the customer experiences, wants, and needs necessary to support innovations on services of the Data Hub.

### 5. Impact on Small Businesses or Other Small Entities

No small businesses will be involved in this project.

### 6. Consequences of Collecting the Information Less Frequently

Data is collected once at this stage in the discovery process, respondents will participate in an interview once lasting no more than 30 minutes.

### 7. Special Circumstances Relating to the Guidelines of 5 CFR 1320.5

This request fully complies with the regulation 5 CFR 1320.5. There are no special circumstances.

### 8. Comments in Response to the Federal Register Notice and Efforts to Consult Outside the Agency Not Applicable

### 9. Explanation of Any Payment or Gift to Respondents

There is no exchange of payment or gifts to respondents for the voluntary interviews.

### 10. Assurance of Confidentiality Provided to Respondents

Activities for this request do not involve the collection of Individually Identifiable Information.

### 11. Justification for Sensitive Questions

There are no sensitive data items to be asked of individual respondents. CDC Human Research Protection Office determined that data /IC is not research involving human subjects and IRB is not required - OADS Project Determination approval (Att. 5).

12. Estimates of Annualized Burden Hours and Costs

Project SUIDIRF project team will interview 50 respondents for this ICR. The project will interview medical examiners/forensic pathologists, law enforcement, death investigators, and child death review Projects for an average of 30 minutes and maximum of 1 responses per respondent. Annualized burden will be 25 hours and an estimated annualized burden cost of \$900.00.

**Estimated Annualized Burden Hours**

**Table A: Estimated Annualized Burden Hours**

Type of Respondents	Form Name	No. of Respondents	No. of Responses per Respondent	Avg. Burden per Response (in hrs.)	Total Burden (in hrs.)
Public health researchers at CDC (epidemiologists, statisticians, Health Scientist)	Interview Guide	50	1	30/60	25
<b>Total</b>					<b>25</b>

**Table B: Estimated Annualized Burden Costs**

Type of Respondents	Form Name	No. of Respondents	No. of Responses per Respondent	Avg. Burden per Response (in hrs.)	Total Burden (in hrs.)	Hourly Wage Rate*	Total Respondent Costs
Public health researchers at CDC (epidemiologists, statisticians, Health Scientist)	Interview Guide	50	1	30/60	25	Average 36.00	\$900.00
							<b>\$ 900.00</b>

\*Average of hourly wage from <http://www.bls.gov/home.htm>

13. Estimates of Other Total Annual Cost Burden to Respondents or Record Keepers

There are no projected cost burdens for reporting.

14. Annualized Cost to the Government

- a. The project cost is associated with the CDC project team members responsible for conducting the interviews. These figures were estimated as the sum of the anticipated direct labor; fringe and burden on direct labor.

<b>Project Staff Oversight</b>	<b>Annual Cost</b>
CDC Cost: Health Scientist (5% of Time)	\$5,000.00
CDC Cost: PH Advisor (2% of Time)	\$3,600.00
<b>Total</b>	<b>\$8,600.00</b>

15. Explanation for Program Changes or Adjustments

This information collection request is a new submission.

16. Plans for Tabulation and Publication and Project Time Schedule

The proposed interviews will be conducted within 2-3 months after approval of GenIC. There is no planned publication from this information collection. Interim reports will be developed, which will incorporate data collected from these sources in 2017 and 2018.

17. Reason(s) Display of OMB Expiration Date is Inappropriate

The display of the OMB expiration date is not inappropriate.

18. Exceptions to Certification for Paperwork Reduction Act Submissions

There are no exceptions to the certification statement.