
United States Census 2020

2020 Census Non-ID Processing Project Plan WBS 4.107

[This page intentionally left blank]

Table of Contents

1. INTRODUCTION.....	1
2. TEAM MEMBERSHIP AND ORGANIZATION.....	1
2.1 Project Information	1
2.2 Team Communication Plan.....	2
2.3 Skill Sets Required.....	2
2.4 Roles and Responsibilities	3
2.5 High-Level Project Metrics.....	4
3. PROJECT SCOPE.....	5
3.1 Strategic Goals, Objectives, Strategies, and Performance Measures.....	5
3.2 High-Level Project Purpose	5
3.3 Research Questions	6
3.4 Scope Statement and Assumptions	6
3.4.1 Assumptions.....	6
3.4.2 Scope Statement.....	6
3.4.3 Project Objectives and Activities that are within Scope.....	7
3.4.4 Project Activities that are Out of Scope.....	7
3.4.5 Prioritized Scope.....	7
3.5 Project Dependencies	8
3.5.1 Predecessors.....	8
3.5.2 Interdependencies	8
4. METHODOLOGY AND TESTING	9
4.1 High-Level Methodology.....	9
4.2 Field Test Summary	10
5. PROJECT MANAGEMENT AND SYSTEMS ENGINEERING	13
6. PROJECT CLOSEOUT.....	13
7. APPROVAL SIGNATURES	14
8. DOCUMENT LOGS.....	15
8.1 Sensitivity Assessment.....	15
8.2 Review/Approval	15
8.3 Version History	16

1. Introduction

To meet the strategic goals and objectives for the 2020 Census, the Census Bureau must make fundamental changes to the design, implementation, and management of the decennial Census. These changes must build upon the successes and address the challenges of the previous Censuses while also balancing challenges of cost containment, quality, flexibility, innovation, and disciplined and transparent acquisition decisions and processes.

This document presents the plan for managing the 2020 Census Non-ID Processing Project for the Research and Testing (R&T) Phase of the 2020 Census (FY2012 - FY2014). This document is in accordance with the *2020 Census Program-Level Research and Testing Strategy*, the *2020 Census Program-Level Research and Testing Plan*, the *2020 Census Program Management Plan*, the *2020 Census Systems Engineering Management Plan*, the *2020 Census Acquisition Manual*, and other relevant 2020 Census and industry research and program standards and guidelines.

The purpose of the Project Plan is to describe key management and technical plans for the research project. A Study Plan supplements the Project Plan, describing the technical and analytical approach in further detail. The Project Plan enables the 2020 Research and Planning Office (20RPO) and the Integrated Product Team (IPT) to validate that the project planning baseline fits within the broader program needs (e.g., scope, schedule, relationship to other projects) and provides an early opportunity for the 20RPO and IPT Leads to identify synergies and potential overlaps and disconnects among projects.

The Project Plan serves as a vehicle for both communication and coordination by providing a broad outline of project objectives, tasking, and timelines to team members and stakeholders during early project organization stages. Once approved, this document includes the project control baseline for this project. Changes to the approved baseline need to go through the program-level approved change management processes (see the *2020 Census Change Control Management Plan*).

2. Team Membership and Organization

2.1 Project Information

Project Name	Non-ID Processing
Project Budget #	6410x07*
Work Breakdown Structure (WBS) #	4.107
IPT Lead and Division	Frank McPhillips, DMD
Investment Area	Enumeration
Research Track	Expanded, Automated, and Tailored Response
IPT Member Divisions	CARRA, CSM, DMD, DSSD, GEO

*“x” denotes last digit of current fiscal year

2.2 Team Communication Plan

Meeting Time and Frequency	Team: Thursdays at 11am MTdb Processing subteam: Tuesdays at 11am AR subteam: Wednesdays at 10am Location-Based Technologies subteam: as needed
Meeting Locations	3H145
Project Folder Location	https://collab.ecm.census.gov/div/20rpo/R3-R8_OPS/4.107_Non_ID/SitePages/Home.aspx
Group Email Distribution List(s)	N/A
Status Reporting	Bi-weekly reports to Census Management, weekly detailed reports to Program Manager

2.3 Skill Sets Required

Decennial Census Programs

Project Management

- Scheduling
- Budget Formulation and Execution
- Communications Planning
- Risk Management
- Quality Management
- Change Control
- Requirements Management

Administrative Records

Matching, Record Linkage, and Data Mining

Input and Output File Development

IT Security

Data File Transfers

Workload Management

MAF/TIGER Database (MTdb) Management

Specification Development

Contract Management

Geography

Field Operations

Study Design

Field Test Design and Management

Qualitative Data Collection

Cost/Benefit Analysis

Statistical Analysis

Report Writing

Application of Location-based Technologies

2.4 Roles and Responsibilities

Role	Responsibilities	Name(s) and Division(s)
IPT Lead / Project Manager	Set meeting agendas, facilitate team discussions, arrange integration activities with other teams, ensure all deliverables are on time.	Frank McPhillips, DMD
Note Taker	Capture a summary of discussions at team meetings, action items assigned, and other details as needed.	Rotation among team members
Schedule Point of Contact	Facilitate assembly of tier 3 and tier 4 schedules; status schedule once baselined; submit change requests as needed.	Ryan King, CSM
Budget Point of Contact	Monitor compliance with budget plan and provide variance explanations to 2020 management.	Frank McPhillips, DMD
Risk and Issues Management Point of Contact	Facilitate the team's formulation of risk statements, probability and impact assessment, mitigation strategies, and contingency planning, as well as monitoring and updating risks as appropriate during the course of the project.	Michael Niosi, DMD
Field Test Point of Contact	Coordinates and communicates the team's requirements and objectives to the field test team	Frank McPhillips, DMD Mekia Turner, DMD
Change Control Point of Contact	Represent team in pursuing changes to scope, objectives or other aspects of project plan and other documentation.	Keith Wechter, DSSD
Requirements Point of Contact	Facilitate operational and software requirements development and documentation for the team.	Mekia Turner, DMD
Documentation Management Point of Contact	Manages the team SharePoint site.	Mekia Turner, DMD
Acquisitions Point of Contact	Lead the development of documents associated with the team acquisition needs. Serves as liaison with 20RPO acquisitions staff.	Sean Kinn, GEO
Knowledge Management Point of Contact	Oversees the review and disposition of knowledge management entries. Serves as liaison with 20RPO knowledge management staff.	Michael Niosi, DMD
Data Access Point of Contact	Identifies and coordinates data needs from the team and works with DMD data access coordinator to secure access to datasets not readily available to team.	Matt Bouch, CARRA

2.5 High-Level Project Metrics

Project Performance Measures

Project metrics provide information on the day-to-day activities of individual projects by quantifying how well investments are meeting business and technical requirements and performance against the established cost, schedule and technical baseline.

Measures highlight performance that will help project teams to:

- monitor project status relative to budget, schedule, and technical performance
- make decisions affecting the project design using quantitative data

Measures will also help management and/or stakeholders to:

- effectively assess progress against the strategic plan and predetermined performance goals
- improve processes and achieve operational efficiency to include cost reduction
- make decisions affecting program design using quantitative data
- control resources and apply them to projects that best support the 2020 Census goals, objectives and strategies

Metrics will be defined as the project team begins to implement project activities and identify appropriate outputs/outcomes to support goals and objectives. These measures will align with the appropriate organizational strategies, objectives, and goals as outlined within the 2020 Census Strategic Plan

3. Project Scope

3.1 Strategic Goals, Objectives, Strategies, and Performance Measures

The objective of the Research and Testing Phase is to develop preliminary designs based on solid evidence and a trade-off analysis aimed at achieving the goal of conducting the 2020 Census at a lower cost than the 2010 Census (per housing unit on an inflation-adjusted basis). The table below identifies the 2020 Census strategic goals, objectives, strategies, and performance measures that the project supports.

2020 Census Strategic Goals Supported	2020 Census Strategic Objectives Supported	2020 Census Strategies	2020 Census Performance Measures
Goal 1: A Complete and Accurate Census	Objective 1.1: All living quarters associated with physical locations	Strategy 1.1.2: Identify better methods to link persons to physical locations to improve geocoding	1. Net coverage error rate at the national level 2. Differential over- or under-coverage by group or by geographic area (e.g., Type of Enumeration Area TEA) 3. Numbers and rates of erroneous enumerations and omissions
	Objective 1.3: People accurately enumerated	Strategy 1.3.4: Improve the tracking capabilities of the enumeration universe	
Goal 3: An efficient 2020 Census	Objective 3.1: An efficient design	Strategy 3.1.1: Contain the cost per housing unit by enhancing the effectiveness of census operations through new approaches and innovation, including agile decision-making.	Cost per housing unit
		Strategy 3.1.3: Leverage strategic partnerships (e.g., USPS, Internet service providers, and administrative records providers such as SSA) to improve the 2020 Census design, lower cost, and improve quality.	

3.2 High-Level Project Purpose

Identify, test, and refine options that readily link (match) addresses supplied by Census respondents and enumerators to an existing MTdb record, or in the case of a non-match, determine the correct physical location and associated geographic codes.

3.3 Research Questions

1. *Via automated processing, how can the Census Bureau more effectively match and geocode addresses lacking a pre-assigned Census ID?*
2. *For cases not matched, but successfully geocoded during automated Non-ID Processing, how can the Census Bureau significantly reduce the Field Verification workloads?*

3.4 Scope Statement and Assumptions

3.4.1 Assumptions

- 2020 Non-ID Processing will account for response modes including, but not limited to the following:
 - Internet (via web and mobile device apps)
 - Telephone interviews (inbound and outbound)
 - Paper questionnaires received at data capture centers that lack a Census ID
 - Living quarters adds captured by enumerators during fieldwork
- Mechanisms for real-time matching and geocoding will be developed or are readily available, including the following:
 - Real-time address standardization
 - Real-time access to administrative record data for address data supplementation
 - Real-time access to the MTdb for matching and geocoding Non-ID cases
- Applications and infrastructure associated with Non-ID Processing will safeguard Title 13 address information as well as personally identifiable information such as respondent name, phone number, and date of birth.
- Administrative record data of sufficient quality will be available to automated Non-ID Processing to enable research in enhancing matching and geocoding in order to reduce workloads for clerical Non-ID Processing and Field Verification
- It is feasible to integrate a map interface or other self-geocoding mechanism into the internet response interfaces (i.e., the technology exists and may be adapted for use for Census Bureau purposes)

3.4.2 Scope Statement

The 2020 Non-ID Processing team will conduct research and testing regarding the enhancement of automated Non-ID Processing, explore alternative methods for assigning a census block geocode to a response, and pursue options for decreasing operational cost.

3.4.3 Project Objectives and Activities that are within Scope

- Determination of options for enhancements to automated Non-ID matching and geocoding processes, including exploration of the following:
 - Implementation of real-time and near real-time matching
 - Use of administrative records data to supplement or validate respondent-provided address data prior to matching and geocoding
 - Applying quality control measures, such as edits to detect input file errors prior to matching and geocoding
- Establishment of options for operational cost savings by decreasing workload for clerical Non-ID Processing and any associated field verification, including:
 - Utilization of address and contact information from administrative records in lieu of fieldwork to verify the existence and location of living quarters addresses not matched to a census record
 - Implementation of a mechanism by which a respondent could identify and confirm the geographic location of their living quarters at the time of response, such as a map-based interface
 - Exploration of the use of Location Based Services and similar technologies for the purposes of deriving a respondent's geographic location
- Conduct testing and analyze test results
- Report test results, analysis, conclusions, and recommendations

3.4.4 Project Activities that are Out of Scope

While the Non-ID team will contribute to discussions regarding address data collection during response, implementation of changes to the address components in the response modes is not in scope for the team; instead, this is part of the research work of the Optimizing Self Response (4.102) and Questionnaire Content Design and Mode Study (4.105) teams.

3.4.5 Prioritized Scope

The highest priority for the project will be testing methods for enhancing automated Non-ID Processing. In 2010 Non-ID Processing, approximately two thirds of the incoming cases were resolved during automated processing. The remaining cases were resolved through clerical processing, which is much more costly than automated processing. In addition, fieldwork to verify the non-matching, geocoded addresses is costly. Therefore, the focus for 2020 Non-ID research will be improving in the resolution rate during

automated Non-ID Processing to reduce clerical and field verification workload, and thus reduce overall operational cost. Assuming that *near* real-time processing (defined here as occurring within 24 hours of receipt of delivery of Non-ID cases from Workload Management) is determined to be sufficient to meet goals for improving case resolution rates, it will take priority over real-time processing as an in-scope activity for the planned field tests. However, at a minimum the Team can simulate real-time processing in separate testing efforts, similar to the planned approach for the exploration of location-based technologies.

3.5 Project Dependencies

3.5.1 Predecessors

- Evaluation of Address List Maintenance Using Supplemental Data Sources (Evaluation 11)
- Non-ID Processing Assessment
- Be Counted Assessment
- Avoid Follow-up Assessment

3.5.2 Interdependencies

- 6410101/6411101/4.101 Automating Field Activities (Infrastructure and Operations)
- 6410103/4.103 Optimizing Self-Response
- 6410104/6411104/4.104 Workload Management Systems
- 6410105/4.105 Questionnaire Content, Design and Mode Study
- 6410189/4.189 Enumeration: 2020 Tests and Experiments
- 6810104/8.104 Privacy and Confidentiality Study
- 6810105/8.105 Matching Process Improvement
- 6810106/8.106 Contact Frame
- 6810189/8.189 Infrastructure: 2020 Tests and Experiments
- GSS Initiative
- 2011 NCRN grant to Carnegie Mellon; Stephen Fienberg, principal investigator
- Policy considerations
- CR26: Citizens Self-Report of Housing Unit Additions and Deletes via Citizen Owned Handheld Devices

4. Methodology and Testing

4.1 High-Level Methodology

The general flow of research activities will be to start with a literature review to determine if other work can inform the determination of the scope of the team's research. Having identified the research goals and objectives, the team will develop a study plan for research. Next, the team will work to develop requirements for formal testing that provide opportunities to pursue all of the research goals. After each formal test, the team will assemble summary documentation of the results, including any conclusions. Once the last formal test is complete, the team will produce a report that assembles the cumulative results of the testing efforts.

To date, the following documents have been part of the literature review:

Bauder, Mark and Judson, D.H., "Administrative Records Experiment in 2000 (AREX 2000) Household Level Analysis", Census 2000 Experiment, April 17, 2003.

Berning, Michael, "Administrative Records Experiment in 2000 (AREX 2000) Request for Physical Address Evaluation", Census 2000 Experiment, April 17, 2003.

Childs, Jennifer Hunter (2011). "Study Plan for Cognitive Testing of Roster, Coverage, and Address Questions for the 2020 Census." September 2011. DRAFT.

Decennial Management Division, Geographic Programs Branch (2012); "2010 Census Field Verification (FV) Assessment Report," 2010 Census Planning Memoranda Series No. 162, U.S. Census Bureau, February 23, 2012.

Gibson, Christine and Nusser, Sarah. "2006 Census Test Evaluation 3: Assessing the Role of Spatial Abilities in Address Canvassing Activities," December 11, 2006.

Jackson, Geoff; Wechter, Keith; and Winder, Susanna, "2010 Census Be Counted and Questionnaire Assistance Centers Assessment," 2010 Census Planning Memoranda Series No. 194, May 22, 2012.

Niosi, Michael (2012); "2010 Census Non-ID Processing Assessment Report," 2010 Census Planning Memoranda Series No. 201, U.S. Census Bureau, July 13, 2012.

Tomaszewski, Christine (2012); "2010 Census Evaluation of Address List Maintenance Using Supplemental Data Sources," August 14, 2012. DRAFT

4.2 Field Test Summary

The following tables describe the goals, objectives and success criteria for planned testing, which the team will conduct to meet all research goals.

Ongoing Testing (Oct 2012 – Sep 2014)

Goal	Related Objectives	Success Criteria
Demonstrate the application of location-based technologies for respondent geocoding	Working via existing systems as well as prototypes developed in the Census Bureau’s Center for Applied Technology (CAT), the team will participate in “proof of concept” testing of location-based technologies, such as collection of GPS coordinates from mobile devices and IP addresses from internet submissions.	<p>Demonstration of the ability to utilize the technology will be the initial measure of success.</p> <p>Assuming resolution of policy issues regarding collecting location data directly from respondent devices, a further measure could be the use of the technology in one of the Field Tests.</p>
Demonstrate respondent self-geocoding via map interface	Working with CAT staff or potentially contractor(s), the team will conduct “proof of concept” testing to demonstrate the use of technology such as an internet-based map through which a respondent could indicate the geographic location of their living quarters.	<p>As with the testing of location-based technology, a demonstration of the ability to utilize the “self-geocoding” technology will be the initial measure of success.</p> <p>Similarly, assuming resolution of policy issues regarding display of Title 13 address and location data on a response interface, a further measure could be use of the technology in one of the Field Tests.</p>
Utilize 2010 Non-ID Processing assessment data to test proposed enhancements to 2020 Non-ID Processing	Using known data from 2010 Non-ID Processing, the team will modify datasets to simulate multiple scenarios for matching in attempt to overcome obstacles encountered during 2010 Non-ID Processing	Demonstration of improvements in resolution rate via automated processing during address standardization, matching and geocoding.

Field Test 15 (Jan 2013)

Goal	Related Objectives	Success Criteria
Demonstrate the use of Administrative Records to support matching and geocoding	Use Administrative Records to supplement (or corroborate) respondent-provided address information prior to conducting automated matching and geocoding.	Data provided by Administrative Records is used for matching and the quality of the match is corroborated by an alternate source (e.g., FastData).
Utilize a matching and geocoding service without having to apply updates to the MTdb	Demonstrate the use of a Matching & Geocoding application for automated matching/geocoding of Test 15 Non-ID addresses.	Matching and Geocoding occurs in sufficient time for team to analyze results to inform subsequent tests. Also, all address records obtain a valid outcome ¹ via automated processing.

Field Tests 17, 18 and 20 (May-Jun 2013)

Goal	Related Objectives	Success Criteria
Continue testing the use of Administrative Records to support matching and geocoding	Continue use of Administrative Records to supplement/corroborate respondent-provided address information prior to conducting matching and geocoding. This may include modifications to the methods used in Test 15.	Administrative Records provide supplementation and corroboration at comparable or improved rates when compared to the results from Test 15.
Demonstrate <i>near real-time matching and geocoding</i> of address inputs from multiple modes of response	Test the near real-time matching and geocoding of Non-ID cases from internet. <i>Potentially</i> test near real-time processing for other modes, such as computer aided telephone interviewing (CATI), telephone Questionnaire Assistance (TQA) and during non-response follow-up (NRFU).	The establishment of a direct data delivery to GEO for processing on a flow. All address records obtain a resolution upon completion of GSSMG processing, and occur within 24 hours of GEO's receipt of the input file.
Utilize Internet Protocol (IP) address from internet responses to derive a census block	Derive census block using IP address and compare against results from standard geocoding methodologies.	Use IP address to derive a census block geocode for all internet responses. Compare with results from standard geocoding to provide analysis data for consideration in future testing.

¹ Valid outcomes could include rejection for insufficient data, matched and/or geocoded, or unresolved.

Test 26 (Oct 2013)

Goal	Related Objectives	Success Criteria
Continue testing the use of Administrative Records to support matching and geocoding	Continue use of Administrative Records to supplement respondent-provided address information prior to conducting matching and geocoding. This may include modifications to the methods used in preceding tests.	Administrative Records provide supplementation and corroboration at comparable or improved rates when compared to the results from preceding tests.
Demonstrate real-time matching and geocoding of address input from multiple modes of response	Test the matching and geocoding of Non-ID cases from internet, CATI, TQA, and NRFU enumerator devices.	Demonstrate the use of Administrative Records and GSSMG processing during the CATI interview or Internet response.
Examine how real-time matching and geocoding impact internet and CATI response	Integrate real-time matching and geocoding into both the CATI and Internet response vehicles to prompt respondents for further information if they have provided incomplete or insufficient address or contact information.	Obtain data for analysis to determine the degree to which real-time processing was able to increase successful matching and geocoding by prompting additional input/clarification from respondents regarding address/location information.

Test 29 (April 2014)

Goal	Related Objectives	Success Criteria
Continue testing the use of Administrative Records to support matching and geocoding	Continue use of Administrative Records to supplement respondent-provided address information prior to conducting matching and geocoding. This may include modifications to the methods used in preceding tests.	Administrative Records provide supplementation and corroboration at comparable or improved rates when compared to the results from preceding tests.
Demonstrate real-time matching and geocoding of addresses from multiple response modes	Test the matching and geocoding of Non-ID cases from internet, CATI, TQA, and NRFU enumerator devices.	Demonstrate the use of Administrative Records and GSSMG processing during the CATI interview or Internet response.
Reduce Field Verification (FV)	Utilization of address and contact information from administrative records in lieu of fieldwork to verify the existence and location of living quarters addresses not matched to a census record.	Analysis indicates that use of Administrative Records is a reliable and cost effective alternative to an actual FV.

5. Project Management and Systems Engineering

Key management strategies to ensure that the results of the research inform the design decisions include:

- Employ appropriate systems engineering, program management, risk management, and governance best practices to ensure research objectives, key milestones, and outcomes are met;
- Manage research and testing activities as an integrated program;
- Work collaboratively with other Census Bureau initiatives that are providing inputs into the 2020 Census research projects and/or design decisions;
- Provide research project leads with the training, resources, and support they need to get the job done and hold them accountable for results; and
- Engage selected internal and external stakeholders in the review process, make connections with relevant research and initiatives outside of the Census Bureau, and provide transparency to these stakeholders about the decision-making process.

Project activities related to acquisition and sourcing, budget, schedule, risk and issues, and requirements are maintained outside of this document.

6. Project Closeout

The team will identify lessons learned and recommendations that will be included for future consideration and maintained in the Knowledge Management database. The team will conduct other closeout activities, such as contract closeout, as required by the 20RPO.

7. Approval Signatures

Burton H. Reist

8/15/2014

Burton H. Reist
Chief, 2020 Research and Planning Office

Date

8. Document Logs

8.1 Sensitivity Assessment

This table specifies whether the document contains any administratively restricted information.

Verification of Document Content	
This document does not contain any: <ul style="list-style-type: none"> • Title 5, Title 13, Title 26, or Title 42 protected information; • Procurement information; • Budgetary information; and/or, • Personally identifiable information. 	
Document Author/Team Lead: Frank McPhillips, 4.107 Lead	Date: 6/17/2014

8.2 Review/Approval

This table documents the review level and/or approval authority.

Document Review and Approval Tier: Program Document		
Name	Area Represented	Date
Jill O'Brien	Program Manager	6/17/2014
Kimberly Higginbotham	Program Management Manager	8/13/2014
Burton Reist	Chief, 2020 Research and Planning Office	8/15/2014

8.3 Version History

The document version history recorded in this section provides the revision number, the version number, the date it was issued, and a brief description of the changes since the previous release. Baseline releases are also noted.

Rev #	Version	Date	Description
	0.1	6/14/12	Initial draft for 30-day checkpoint. Includes Sections 2.1, 2.2, 3.1 through 3.4.4
	0.2	7/05/12	Second draft for 60-day checkpoint. Includes updates from 30-day checkpoint as well as Sections 2.3 and 3.4.5.
	0.3	8/16/12	Third draft for 90-day checkpoint. Includes updates since 60-day checkpoint.
	0.4	9/20/12	Final draft for 120-day checkpoint/baselining
	1.0	10/19/12	Baselined version following incorporation of feedback from program management review
	2.0	6/17/2014	Annual Project Update