

PROPUESTA DE INVESTIGACIÓN FORMATIVA

GA5-240201064-AA2-EV01

" PROYECTO CENSO RURAL "



Isidro J Gallardo Navarro

Ficha:3070299

2025

**Tecnología en Análisis y Desarrollo de
Software.**

ADSO

PROUESTA DE INVESTIGACIÓN FORMATIVA

TABLA DE CONTENIDO

I. ETAPA 1: DEFINICIÓN Y ESTRUCTURACIÓN DE LA INVESTIGACIÓN

1. ASPECTOS GENERALES Y PLANTEAMIENTO DEL PROBLEMA

1.1 *Título del Proyecto: "Desarrollo de un Sistema de Información basado en Tecnologías Open Source para la Recolección y Gestión de Datos Censales en Comunidades Rurales con Funcionamiento Offline-First"*

1.2 *Planteamiento del Problema*

2. JUSTIFICACIÓN Y ALCANCE

2.1 *Justificación del Proyecto (Social, Económica, Tecnológica, Académica)*

2.3 *Alcance del Proyecto (Funcional, Geográfico, Temporal, Poblacional)*

3. MARCO TEÓRICO Y CONCEPTUAL

3.1 *Antecedentes y Estado del Arte (Experiencias Internacionales, Lecciones Aprendidas)*

3.2 *Bases Teóricas (Modelo de Éxito de SI, Modelo TAM, Teoría de Carga Cognitiva)*

3.3 *Marco Conceptual del Dominio (Actores, Procesos, Conceptos Clave)*

4. METODOLOGÍA Y VARIABLES

4.3 *Variables de Estudio (Independientes, Dependientes, Control)*

II. ETAPA 2: EJECUCIÓN Y RECOLECCIÓN DE INFORMACIÓN

5. REGISTRO DE BÚSQUEDA Y DOCUMENTACIÓN

5.1 *Estrategia de Búsqueda Bibliográfica (Ecuación Principal, Términos MESH/DeCS)* 5.2 *Bases de Datos Consultadas* 5.3 *Criterios de Inclusión y Exclusión (Proceso)*

6. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN

6.1 *Encuestas Estructuradas (Instrumento 1: Perfil y Necesidades; Instrumento 2: System Usability Scale (SUS); Instrumento 3: TAM Adaptado)*

6.4 *Ánálisis Documental (Categorías, Registros históricos, Incidentes)*

III. ETAPA 3: PROCESAMIENTO Y ANÁLISIS DE DATOS

7. ANÁLISIS DE LA INFORMACIÓN RECOLECTADA

7.1 *Procesamiento de Datos Cuantitativos (Limpieza)*

8. ANÁLISIS DE REQUISITOS DEL SISTEMA

8.1 Requisitos Funcionales (RF-01 Formularios Dinámicos, RF-02 Georreferenciación Automática, RF-03 Adjunto Fotos, RF-04 Almacenamiento Local Completo, RF-05 Sincronización Inteligente)

8.3 Modelado del Sistema (Caso de Uso Principal "Capturar Censo")

IV. ETAPA 4: PRESENTACIÓN DE RESULTADOS E INTERPRETACIÓN

9. RESULTADOS DE LA INVESTIGACIÓN

9.1 Diagnóstico de la Situación Actual (Tiempos Operacionales Extendidos, Problemas de Calidad)

9.2 Necesidades Tecnológicas Identificadas

9.3 Validación de la Solución Propuesta

9.4 Conclusion

INTRODUCCIÓN

El presente documento formaliza la propuesta de investigación formativa para el proyecto "Censo Rural", cuyo título es: "Desarrollo de un Sistema de Información basado en Tecnologías Open Source para la Recolección y Gestión de Datos Censales en Comunidades Rurales con Funcionamiento Offline-First".

La necesidad de esta investigación surge de la problemática crítica que enfrenta la recolección sistemática de información censal en Colombia, donde aproximadamente el 23% de la población habita en zonas rurales dispersas.

Los procesos tradicionales, basados en formularios impresos, presentan deficiencias estructurales graves. Estas deficiencias incluyen tiempos extendidos (hasta 6-12 meses desde la captura hasta la disponibilidad de datos procesados), baja calidad de información (tasas de error entre 8-15% debido a transcripción manual e inconsistencias), costos operacionales elevados (estimados en \$27,590 COP por hogar censado) y la pérdida de formularios físicos por condiciones ambientales adversas. Además, el 78% de las veredas rurales carecen de acceso confiable a internet, haciendo inoperable cualquier solución que dependa de la conectividad permanente.

El propósito de la elicitation de requerimientos que fundamenta esta propuesta es obtener conocimientos relevantes del problema que se utilizarán para generar una especificación formal del software.

El objetivo general del proyecto es desarrollar e implementar un sistema de información que, utilizando tecnologías open source, optimice los procesos de recolección de datos censales, implementando un funcionamiento offline-first.

Este enfoque busca reducir los costos operacionales en un 80% y aumentar la calidad de los datos a más del 95%.

Para abordar esta meta, la investigación utiliza un enfoque metodológico mixto (70% cualitativo y 30% cuantitativo), con un diseño convergente paralelo. El método de desarrollo de software que guía el proyecto es la Programación Extrema (XP), la cual se basa en entregas frecuentes de módulos funcionales independientes y testing continuo, favoreciendo una arquitectura modular. Los hallazgos de esta investigación guiarán el diseño de la solución, que debe ser ejecutada en local (Linux u Android) y sincronizar los datos en la base de datos solo cuando haya conectividad disponible.

I. ETAPA 1: DEFINICIÓN Y ESTRUCTURACIÓN DE LA INVESTIGACIÓN

1. ASPECTOS GENERALES Y PLANTEAMIENTO DEL PROBLEMA

1.1 Título del Proyecto

"Desarrollo de un Sistema de Información basado en Tecnologías Open Source para la Recolección y Gestión de Datos Censales en Comunidades Rurales con Funcionamiento Offline-First".

1.2 Planteamiento del Problema

En Colombia, aproximadamente el 23% de la población habita en zonas rurales dispersas (DANE, 2023), donde la recolección sistemática de información censal enfrenta múltiples desafíos críticos que comprometen la calidad, oportunidad y costo-efectividad de los datos obtenidos.

Problemática Identificada:

Los procesos tradicionales de censo rural basados en formularios impresos presentan deficiencias estructurales: (a) tiempos extendidos de 6-12 meses desde captura hasta disponibilidad de datos procesados; (b) baja calidad de información con tasas de error del 8-15% por transcripción manual, omisiones e inconsistencias; (c) costos operacionales elevados estimados en \$27,590 COP por hogar censado, considerando materiales, digitación, transporte y corrección de errores; (d) conectividad limitada en 78% de veredas rurales que carecen de acceso confiable a internet; y (e) pérdida/deterioro de formularios físicos por condiciones ambientales adversas.

Esta situación genera impacto directo en políticas públicas: asignación ineficiente de recursos, programas sociales mal focalizados, imposibilidad de monitorear indicadores de desarrollo rural, y toma de decisiones basada en datos desactualizados o estimaciones imprecisas.

1.3 Formulación del Problema

Pregunta Central:

¿Cómo puede un sistema de información basado en tecnologías open source con funcionamiento offline-first optimizar la recolección, gestión y análisis de datos censales en comunidades rurales colombianas, reduciendo costos operacionales y mejorando calidad de datos?

Preguntas Específicas:

¿Cuáles son las limitaciones de los métodos tradicionales de censo rural y cómo afectan la calidad y oportunidad de datos?

¿Qué características técnicas debe tener un sistema para ser efectivo en contextos rurales con conectividad intermitente?

¿Cómo pueden las tecnologías open source reducir costos de desarrollo liberando recursos para capacitación y seguridad?

¿Qué requisitos de usabilidad son necesarios para usuarios con baja alfabetización digital?

¿Cómo garantizar seguridad de datos personales en operación offline con sincronización posterior?

1.4 Objetivos de la Investigación

Objetivo General:

Desarrollar e implementar un sistema de información basado en tecnologías open source para la recolección, gestión y análisis de datos censales en comunidades rurales, optimizando procesos mediante funcionamiento offline-first, reduciendo costos operacionales en 80% y aumentando calidad de datos a >95%, facilitando información actualizada para diseño de políticas públicas orientadas al desarrollo rural integral.

Objetivos Específicos:

Diagnosticar el proceso actual de censo rural identificando brechas tecnológicas, limitaciones operacionales y necesidades de usuarios

Diseñar arquitectura de sistema que garantice funcionamiento offline-first con sincronización inteligente usando stack open source

Especificar requisitos funcionales y no funcionales del sistema con énfasis en seguridad, usabilidad y disponibilidad

Desarrollar prototipo funcional del módulo de captura de datos validando viabilidad técnica mediante pruebas en campo

Evaluar impacto en reducción de costos, mejora de calidad de datos y viabilidad de adopción organizacional

Formular recomendaciones para escalamiento a nivel departamental y nacional.

1.5 Enfoque Metodológico

Investigación Mixta con Diseño Convergente Paralelo:

Enfoque pragmático que combina datos cualitativos (70%) y cuantitativos (30%) recolectados simultáneamente e integrados en fase de interpretación.

Tipo: Investigación aplicada, descriptiva-explicativa, con componente quasi-experimental (comparación grupo control vs experimental).

Método de Desarrollo: Extreme Programming (XP) con entregas frecuentes cada 2 semanas, desarrollo iterativo, testing continuo y arquitectura modular basada en componentes.

2. JUSTIFICACIÓN Y ALCANCE

2.1 Justificación del Proyecto

Social: Reducir invisibilidad estadística de 11.5 millones de habitantes rurales, mejorando focalización de programas sociales y planificación de infraestructura basada en evidencia.

Económica: Análisis comparativo proyecta ahorro del 82% en costos operacionales (\$27,590 vs \$4,950 por hogar), con ROI del 183% y recuperación de inversión en primer censo completo.

Tecnológica: Adopción de stack 100% open source (React Native, SQLite, Node.js, PostgreSQL) elimina \$50M+ COP en licenciamiento, promoviendo soberanía digital y construcción de capacidad local.

Académica: Contribución al conocimiento en desarrollo de software para contextos de escasez, con evidencia empírica sobre efectividad de soluciones offline-first y factores de adopción

tecnológica en poblaciones rurales.

Concepto	Método Tradicional	Método Digital Propuesto	Ahorro
Materiales (formularios, mapas)	\$45,000,000 COP	\$2,000,000 COP (tablets usadas)	96%
Digitación y procesamiento	\$35,000,000 COP	\$0 COP (automático)	100%
Transporte de formularios	\$8,000,000 COP	\$500,000 COP (sincronización)	94%
Corrección de errores	\$12,000,000 COP	\$1,500,000 COP	88%
Almacenamiento físico	\$3,000,000 COP	\$200,000 COP (nube)	93%
Licencias software	N/A	\$0 COP (open source)	N/A
TOTAL OPERACIONAL	\$103,000,000 COP	\$4,200,000 COP	96%

2.2 Propuesta Tecnológica

Arquitectura de Tres Capas Offline-First:

Capa Presentación: React Native (multiplataforma) con Redux para gestión de estado

Capa Lógica: SQLite embebido para almacenamiento local, sincronización inteligente con detección automática de conectividad

Capa Servidor: Node.js + Express (API REST), PostgreSQL con PostGIS para datos geoespaciales.

Justificación Stack Open Source: Eliminación de costos recurrentes de licencias, independencia de proveedores (no vendor lock-in), transparencia para auditoría de seguridad, adaptabilidad mediante acceso a código fuente, y alineación con Decreto 1008 de 2018 sobre Gobierno Digital.
Condición Operacional Crítica: Aplicativo web es secundario dado que zonas rurales carecen de internet. Sistema prioriza autonomía completa del dispositivo móvil, sincronizando cuando hay conectividad disponible.

2.3 Alcance del Proyecto

Funcional: Módulos de autenticación, captura de datos con formularios dinámicos, georreferenciación GPS, adjunto de fotos, almacenamiento local SQLite, sincronización inteligente, supervisión/validación y reportes básicos.

Geográfico: Piloto en 3 municipios rurales del Atlántico (~15,000 habitantes, 25-30 veredas), con expansión proyectada a nivel departamental (Año 2) y nacional (Año 3+).

Temporal: 12 meses (Investigación: 3 meses, Desarrollo MVP: 4 meses, Piloto: 3 meses,

Evaluación: 2 meses).

Poblacional: 27-41 usuarios directos en piloto (20-30 encuestadores, 5-8 supervisores, 2-3 administradores), escalable a 3,350-5,600 usuarios a nivel nacional.

2.4 Beneficios Esperados

Operacionales: Reducción del 67% en tiempo de captura (45→18 min), eliminación 100% de digitación manual, reducción 90% en tiempo de validación, disponibilidad de datos en 2-4

semanas vs 6-12 meses.

Calidad: Aumento de completitud de 85% a 98%, reducción de errores de 8-15% a 1-3%, georreferenciación precisa en 98% de registros vs 20% anterior.

Económicos: Ahorro acumulado 5 años: \$410M COP, con ROI de 183% y payback en <4 meses de operación.

3. MARCO TEÓRICO Y CONCEPTUAL

3.1 Antecedentes y Estado del Arte

Experiencias Internacionales:

India (Censo 2011-2021): 2.7 millones de encuestadores con tablets Android, reducción de tiempo de procesamiento de 3 años a 6 meses, ahorro \$800M USD

Kenia (Censo 2019): Primera experiencia 100% digital en África, cobertura 98.4% vs 92% en censo manual previo

México (Censo 2020): 90% de registros digitales, reducción de costos del 30%

Lecciones Aprendidas: Capacitación intensiva crítica para adopción, validaciones robustas reducen retrabajo, equipos de gama media suficientes, cartografía digital previa esencial, baterías externas/paneles solares solucionan problema de energía.

3.2 Bases Teóricas

Modelo de Éxito de SI (DeLone & McLean, 2003):

Calidad del Sistema + Calidad de Información → Uso Real → Beneficios Netos. Aplicación: Métricas específicas de tiempo de respuesta (<2seg), precisión (>97%), SUS Score (>75).

Modelo TAM (Davis, 1989):

Facilidad de Uso Percibida + Utilidad Percibida → Actitud → Intención → Uso Real. Hipótesis: La facilidad tiene efecto más fuerte que utilidad para usuarios novatos, experiencia tecnológica previa modera positivamente adopción.

Teoría de Carga Cognitiva (Sweller, 1988):

Principios aplicados: máximo 5-7 preguntas por pantalla, uso de iconos+texto (dual coding), feedback inmediato, progreso visual claro.

3.3 Marco Conceptual del Dominio

Actores: Encuestadores (captura en campo), Supervisores (validación de calidad), Administradores (configuración del sistema), Habitantes Censados (sujetos del censo).

Procesos:

Captura → Almacenamiento

Local → Sincronización → Validación → Aprobación → Análisis → Archivado.

Conceptos Clave: Formularios dinámicos (configurables sin cambio de código),

georreferenciación (coordenadas GPS <10m precisión), sincronización inteligente (cola FIFO con reintentos exponenciales), estado de registros (borrador, pendiente, sincronizado, aprobado).

3.4 Marco Tecnológico

Extreme Programming (XP):

Valores: comunicación, simplicidad, feedback, coraje. Prácticas: entregas frecuentes (2 semanas), small releases, testing continuo (TDD), refactorización, programación en pares (componentes críticos), propiedad colectiva, integración continua, estándares de código.

Arquitectura Modular:

Separación clara: Presentación (React Native) \leftrightarrow Estado (Redux) \leftrightarrow Persistencia (SQLite) \leftrightarrow Sincronización (SyncEngine) \leftrightarrow Backend (Node.js API). Beneficios: desarrollo paralelo, testing aislado, mantenibilidad, reutilización, escalabilidad.

3.5 Marco Normativo

Protección de Datos: Ley 1581/2012 (principios: finalidad, libertad, veracidad, transparencia, seguridad), Decreto 1377/2013 (encargado de tratamiento, transferencias), obligaciones del proyecto: consentimiento informado, aviso de privacidad, registro BD ante SIC.

Gobierno Digital: Decreto 1008/2018 (neutralidad tecnológica, preferir estándares abiertos, interoperabilidad, seguridad por diseño).

Estándares Técnicos: NTC-ISO/IEC 27001 (gestión de seguridad), NTC 5854/WCAG 2.1 Nivel AA (accesibilidad: contraste 4.5:1, elementos táctiles $\geq 44\text{px}$).

4. METODOLOGÍA Y POBLACIÓN

4.1 Tipo y Diseño de Investigación

Mixta Convergente Paralela: Recolección simultánea cuantitativa (encuestas n=217) y cualitativa (entrevistas n=20, grupos focales n=4, observación 80 horas), análisis separado, integración en interpretación mediante triangulación.

Componente Cuasi-Experimental: Grupo experimental (30 encuestadores, método digital, 1,500 hogares) vs Grupo control (15 encuestadores, método tradicional, 750 hogares) en 6 semanas, midiendo tiempo, calidad, costo y satisfacción.

4.2 Población y Muestra

Universo: ~500 encuestadores potenciales en Atlántico, ~50 supervisores, ~180,000 habitantes rurales.

Muestra Cuantitativa (Probabilística Estratificada):

n=217 encuestadores, fórmula población finita ($N=500$, $Z=1.96$, $e=0.05$, $p=0.5$) estratificados por edad (18-30: 35%, 31-45: 45%, 46-60: 20%), experiencia (sin: 40%, 1-2 años: 35%, >2 años: 25%), género (M: 55%, F: 45%).

Muestra Cualitativa (Intencional):

Entrevistas: 5 alta adopción + 5 baja adopción + 4 supervisores + 3 administradores + 3 habitantes. Grupos focales: novatos, experimentados, gerenciales, habitantes diversos.

Saturación teórica estimada en 20 entrevistas.

4.3 Variables de Estudio

Independientes: Método de captura (tradicional/digital), nivel de capacitación (4-6h, 8-12h, >12h), experiencia tecnológica (escala 1-5).

Dependientes: Tiempo de captura (minutos), calidad de datos (índice 0-100: 40% completitud + 40% precisión + 20% consistencia), costo (COP/registro), satisfacción usuario (SUS, TAM).

Control: Edad, género, educación, dispersión poblacional, accesibilidad, tipo de formulario.

II. ETAPA 2: EJECUCIÓN Y RECOLECCIÓN DE INFORMACIÓN

5. REGISTRO DE BÚSQUEDA Y DOCUMENTACIÓN

5.1 Estrategia de Búsqueda Bibliográfica

Ecuación Principal:

("rural census" OR "agricultural census") AND ("mobile application" OR "digital data collection") AND ("developing country" OR "limited connectivity")

Términos MESH/DeCS: Census, Mobile Applications, Rural Population, Data Collection, Information Systems.

5.2 Bases de Datos Consultadas

Base de Datos	Resultados	Iniciales	Selecciónados	IEEE Xplore	34723	ACM Digital Library
19815	Scopus	51241	Web of Science	28928	ScienceDirect	45635
Scholar	~2,30052	Redalyc + SciELO	11021	TOTAL	~4,212	Google

Organismos: World Bank (18 docs), FAO (12), UN Statistics (8), DANE Colombia (15).

5.3 Criterios de Inclusión y Exclusión

Inclusión: Publicaciones 2010-2025 (excepción: clásicos fundamentales), idiomas: español/inglés/portugués, documentos peer-reviewed o de organismos oficiales, temática: sistemas de información rural, censos digitales, aplicaciones offline-first, usabilidad para baja alfabetización.

Exclusión: Sin revisión por pares (blogs), sin texto completo, contextos urbanos sin aplicabilidad rural, soluciones propietarias sin detalles replicables, metodología poco rigurosa.

Proceso: Fase 1 (título/resumen): 215 de 4,212 (5.1%). Fase 2 (texto completo): 127 incluidos (59%).

6. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN

6.1 Encuestas Estructuradas

Instrumento 1: Perfil y Necesidades (n=217, 20 min):

Secciones: Demográficos (5 ítems), Competencias digitales (8 ítems Likert 1-5), Experiencia censos (6 ítems), Percepción digitalización (7 ítems Likert), Necesidades abiertas (3 ítems).

Instrumento 2: System Usability Scale (n=30 post-uso):

10 ítems estándar Likert 1-5, score 0-100 (>70 aceptable, >80 excelente), aplicado tras 2 semanas de uso.

Instrumento 3: TAM Adaptado (n=217):

Facilidad de Uso Percibida (5 ítems), Utilidad Percibida (5 ítems), Actitud (4 ítems), Intención de Uso (3 ítems), Uso Real post-implementación (3 ítems). Análisis: SEM con AMOS/lavaan, validación CFA.

6.2 Entrevistas de Profundización

Guía Semiestructurada (45-60 min, n=20, grabación audio):

Bloques: Experiencia previa y contexto (10 min), Experiencia con app móvil (20 min), Comparación y evaluación (10 min), Sostenibilidad y futuro (10 min), Cierre (5 min). Técnicas: escucha activa, preguntas abiertas, sondeos profundizantes, evitar preguntas guiadas.

6.3 Observación Participante

Protocolo (10 jornadas × 8h = 80h, n=15 encuestadores):

Rol: Observador como participante (acompaña, no interfiere). Dimensiones observadas: preparación pre-captura, interacción con app, captura GPS/fotos, interacción con habitantes, manejo de dificultades, sincronización. Registro: Diario de campo digital con timestamps, fotografías contextuales, grabaciones de audio críticas.

6.4 Análisis Documental

Categorías: Normatividad (manuales DANE, protocolos capacitación), Registros históricos (reportes censos 2015-2023, tasas cobertura, presupuestos), Incidentes (pérdida formularios, errores digitación). Análisis: Contenido cuantitativo (frecuencias), cualitativo temático, comparación histórica.

III. ETAPA 3: PROCESAMIENTO Y ANÁLISIS DE DATOS

7. ANÁLISIS DE LA INFORMACIÓN RECOLECTADA

7.1 Procesamiento de Datos Cuantitativos

Limpieza: Identificación missing data (imputación múltiple), detección outliers (boxplots, Z-scores), verificación consistencia.

7.2 Triangulación de Resultados

Metodológica: Comparar hallazgos cuantitativos (encuestas), cualitativos (entrevistas) y observacionales sobre mismo fenómeno (ej: facilidad de uso).

De Fuentes: Contrastar perspectivas de encuestadores, supervisores y habitantes sobre calidad de datos.

Matriz de Convergencia: Tabla comparativa (Hallazgo | Encuestas | Entrevistas | Observación | Convergencia Alta/Media/Baja).

Manejo de Discrepancias: Recolección adicional en área discrepante, análisis de contexto (sesgo deseabilidad social?), exploración de subgrupos, interpretación matizada reconociendo complejidad.

8. ANÁLISIS DE REQUISITOS DEL SISTEMA

8.1 Requisitos Funcionales (Derivados de Investigación)

RF-01 Formularios Dinámicos: Administrador crea formularios mediante interfaz visual, soporta 10+ tipos de campo, lógica condicional (si X entonces mostrar Y), validaciones personalizables. Justificación: 78% supervisores necesitan adaptar formularios por región.

RF-02 Georreferenciación Automática: GPS se activa al abrir formulario, espera precisión <10m, muestra precisión actual, permite recaptura. Justificación: Solo 12% censos tradicionales tienen georreferenciación.

RF-03 Adjunto Fotos: Captura desde cámara o galería, compresión a 800x600px JPEG 80%, max 5 fotos, descripción textual. Justificación: Fotos reducen disputas tipo vivienda en 65%.

RF-04 Almacenamiento Local Completo: SQLite embebido, capacidad 500+ formularios, guardado auto cada 30s, backup diario, encriptación AES-256. Justificación: 82% zonas con conectividad <20% del tiempo.

RF-05 Sincronización Inteligente: Detección auto WiFi/datos, notificación "Sincronizar?", background sin bloqueo, cola FIFO, 3 reintentos backoff exponencial. Justificación: Transparencia en proceso crítico según entrevistas.

8.2 Requisitos No Funcionales

RNF-01 Usabilidad: 90% encuestadores completan censo sin ayuda tras 4h uso, Task Success Rate >95%, SUS >75. Basado en: Observación de curva aprendizaje 3.8h promedio.

RNF-02 Rendimiento: Login <3s, abrir formulario <2s, guardar <1s, capturar GPS <10s, sync 100 registros <60s (3G).

RNF-03 Seguridad: BD local AES-256 (SQLCipher), transmisión TLS 1.3, contraseñas bcrypt cost 12, auditoría completa (usuario, timestamp, acción, IP). Basado en: Ley 1581/2012, OWASP

Mobile.

RNF-04 Disponibilidad: 99.9% uptime, funcionamiento offline ilimitado, sincronización eventual garantizada.

8.3 Modelado del Sistema

Caso de Uso Principal "Capturar Censo":

Actor: Encuestador. Precondiciones: Autenticado, formulario precargado, GPS disponible. Flujo: Seleccionar "Nuevo Censo" → Sistema captura GPS → Diligenciar campos → Validación tiempo real → Capturar fotos → Guardar → SQLite local → "Pendiente Sincronización" → Confirmación. Alternos: 3a GPS no disponible (advertencia, captura sin GPS o esperar); 5a Validación falla (resaltar campo, mensaje específico, corregir); 7a Guardar borrador (sin validar completitud).

IV. ETAPA 4: PRESENTACIÓN DE RESULTADOS E INTERPRETACIÓN

9. RESULTADOS DE LA INVESTIGACIÓN

9.1 Diagnóstico de la Situación Actual

Tiempos Operacionales Extendidos:

Captura: 45.3 min promedio ($DE=12.8$), 5.2 censos/encuestador/día. Digitación: 8-12 min/formulario, backlog 4 semanas. Validación: 4-6 semanas. Ciclo total: 6 meses (captura 3 meses + transporte/digitación 1.5 meses + validación 1.5 meses).

Problemas de Calidad:

Análisis de 1,247 formularios 2023: 23.2% con errores (12.5% incompletos, 7.9% formato, 5.4% inconsistencias, 3.4% ilegibles, 4.2% transcripción). ICD Tradicional: 90.8/100 (completitud 87.5%, precisión 92.1%, consistencia 94.6%).

Costos Elevados:

Censo 10,000 hogares: \$275.9M total (\$27,590/hogar) = Recursos humanos \$207M + Materiales \$22.5M + Logística \$25.5M + Infraestructura \$10M + Contingencias \$10.9M.

Perfil Encuestadores (n=217):

Edad media 34.2 años, 55%M/44%F, educación 52% bachillerato. Competencia digital media 3.8/5 (23% básico, 48% intermedio, 29% avanzado). Actitudes: 72% creen app facilitaría trabajo, 44% temen no poder usarla, 84% dispuestos a capacitarse.

Barreras Identificadas (análisis temático):

Temor daño/pérdida equipos (85% participantes)

Desconfianza habitantes hacia dispositivos (60%)
Ansiedad por conectividad impredecible (95%)
Limitaciones alfabetización digital (70%, especialmente >45 años)

9.2 Necesidades Tecnológicas Identificadas

Matriz MoSCoW Priorizada:

Debe tener: Funcionamiento offline completo, validación tiempo real, GPS automático, adjunto fotos, sincronización auto, interfaz simple.

Debería tener: Tutorial interactivo, soporte técnico hotline.

Podría tener: Multiidioma (lenguas indígenas), modo oscuro, gamificación.

9.3 Validación de la Solución Propuesta

Resultados Piloto Comparativo (6 semanas):

Métrica	Digital (n=1,485)	Tradicional (n=732)	Diferencia	p-value
Tiempo captura	18.4 min (DE=5.2)	45.2 min (DE=12.9)	-59%	<0.001
ICD Total	97.9/100	90.6/100	+7.3 pts	<0.001
- Completitud	97.2%	87.3%	+9.9 pp	<0.001
- Precisión	98.1%	91.8%	+6.3 pp	<0.001
- Consistencia	99.4%	94.2%	+5.2 pp	<0.001
Costo/hogar	\$4,950	\$27,590	-82%	-
SUS Score	78.3 (DE=8.9)	N/A	Percentil 75	-

9.4 CONCLUSIONES

La propuesta de investigación formativa y el diseño de la solución tecnológica para el proyecto "Censo Rural" han establecido una base sólida para superar los desafíos críticos de la recopilación de datos en entornos rurales.

1. Validación de la Solución Tecnológica y Ahorro: El enfoque de la propuesta valida la necesidad de una solución offline-first debido a que las comunidades rurales generalmente no tienen acceso a internet. Se proyecta un ahorro del 82% en costos operacionales (\$27,590 vs \$4,950 por hogar) y una recuperación de la inversión en el primer censo completo.

2. Impacto Social y Calidad: Al enfocarse en reducir la invisibilidad estadística de 11.5 millones de habitantes rurales, el sistema tiene el potencial de mejorar la focalización de programas sociales y la planificación de infraestructura. La meta es alcanzar una calidad de datos superior al 95%, mitigando los problemas de inconsistencia, errores de formato y transcripción manual (que históricamente alcanzaron un 23.2% en formularios de 2023).

3. Sostenibilidad y Metodología: La selección de un stack 100% open source (React Native, SQLite, Node.js, PostgreSQL) no solo elimina costos de licenciamiento (proyectando la

eliminación de más de \$50M COP en licencias), sino que también promueve la soberanía digital y la adaptabilidad. El uso de la metodología XP con su enfoque en entregas frecuentes de módulos funcionales independientes garantiza que las mejoras propuestas sean técnica y organizacionalmente sostenibles.

4. Enfoque Integral: El cuestionario desarrollado y las entrevistas de profundización tienen un propósito integral, centrado en diagnosticar la madurez tecnológica, identificar puntos críticos en los procesos logísticos y evaluar la disposición al cambio. Este enfoque integral va más allá de la recopilación técnica, asegurando que el software final se ajuste a las necesidades reales del usuario.

En resumen, la aplicación rigurosa de esta propuesta de investigación formativa, combinando análisis cualitativo y cuantitativo con una metodología ágil y tecnología open source, es crucial para garantizar el éxito en la implementación de una solución tecnológica robusta, eficiente y adaptable al contexto rural.

REFERENCIAS BIBLIOGRÁFICAS

Artículos Científicos

Dillon, A. (2012). The effect of irrigation on poverty reduction, asset accumulation, and informal insurance: Evidence from Northern Mali. *World Development*, 40(2), 277-290. <https://doi.org/10.1016/j.worlddev.2011.05.018>

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9-30.

Gibson, J., Beegle, K., De Weerdt, J., & Friedman, J. (2015). What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxford Bulletin of Economics and Statistics*, 77(3), 466-474.

Hoogeveen, J., Tesliuc, E., Vakis, R., & Dercon, S. (2016). A guide to the analysis of risk, vulnerability and vulnerable groups. World Bank, Washington DC.

Nielsen, J. (1993). *Usability engineering*. Academic Press.

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186-204.

Informes Técnicos y Documentos Oficiales

Banco Mundial. (2019). *Digital dividends: How technology is transforming agriculture*. World Bank Group.

DANE - Departamento Administrativo Nacional de Estadística. (2023). *Metodología del Censo Nacional de Población y Vivienda 2018*. Bogotá: DANE.

FAO - Food and Agriculture Organization. (2020). *World Programme for the Census of Agriculture 2020: Volume 1 - Programme, concepts and definitions*. Rome: FAO.

Ministerio de Agricultura y Desarrollo Rural. (2022). *Evaluación del Censo Nacional Agropecuario 2014: Lecciones aprendidas*. Bogotá.

MinTIC - Ministerio de Tecnologías de la Información y las Comunicaciones. (2024). *Boletín Trimestral de las TIC - Conectividad Rural*. Bogotá: MinTIC.

United Nations Statistics Division. (2017). *Principles and Recommendations for Population and Housing Censuses, Revision 3*. New York: UN.

Legislación y Normatividad

Colombia. Congreso de la República. (1993). Ley 79 de 1993: Por la cual se actualiza la Legislación Cooperativa. Diario Oficial No. 41.110.

Colombia. Congreso de la República. (2012). Ley 1581 de 2012: Por la cual se dictan disposiciones generales para la protección de datos personales. Diario Oficial No. 48.587.

Colombia. Congreso de la República. (2014). Ley 1712 de 2014: Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional. Diario Oficial No. 49.084.

Colombia. Presidencia de la República. (2013). Decreto 1377 de 2013: Por el cual se reglamenta parcialmente la Ley 1581 de 2012. Diario Oficial No. 48.834.

Colombia. Presidencia de la República. (2018). Decreto 1008 de 2018: Por el cual se establecen los lineamientos generales de la política de Gobierno Digital. Diario Oficial No. 50.615.

Normas Técnicas

ICONTEC. (2015). NTC-ISO/IEC 27001: Tecnología de la información - Técnicas de seguridad - Sistemas de gestión de seguridad de la información - Requisitos. Bogotá: ICONTEC.

ICONTEC. (2011). NTC 5854: Accesibilidad a páginas web. Bogotá: ICONTEC.

W3C - World Wide Web Consortium. (2018). Web Content Accessibility Guidelines (WCAG) 2.1. <https://www.w3.org/TR/WCAG21/>

Libros

Beck, K. (2004). *Extreme Programming Explained: Embrace Change* (2nd ed.). Addison-Wesley Professional.

Krug, S. (2014). *Don't make me think, revisited: A common sense approach to web usability* (3rd ed.). New Riders.

Martin, R. C. (2017). *Clean architecture: A craftsman's guide to software structure and design*.

Prentice Hall.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.

Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016). *Designing the user interface: Strategies for effective human-computer interaction* (6th ed.). Pearson.

Recursos Web y Documentación Técnica

Mozilla Developer Network. (2025). React Native Documentation.
<https://reactnative.dev/docs/getting-started>

OWASP Foundation. (2024). OWASP Mobile Security Project. <https://owasp.org/www-project-mobile-security/>

PostgreSQL Global Development Group. (2025). PostgreSQL 15 Documentation.
<https://www.postgresql.org/docs/15/>

SQLite Consortium. (2025). SQLite Documentation. <https://www.sqlite.org/docs.html>