



Población y Salud en Mesoamérica

E-ISSN: 1659-0201

revista@ccp.ucr.ac.cr

Universidad de Costa Rica

Costa Rica

Hidalgo-Céspedes, Jeisson; Rosero-Bixby, Luis; Antich-Montero, Daniel  
Mejora de la calidad y disminución de costos de censos y encuestas utilizando  
computadores de mano (PDA). Una aplicación en Costa Rica  
Población y Salud en Mesoamérica, vol. 5, núm. 1, julio-diciembre, 2007  
Universidad de Costa Rica  
San José, Costa Rica

Disponible en: <http://www.redalyc.org/articulo.oa?id=44659907>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



## Población y Salud en Mesoamérica

Revista electrónica publicada por el  
Centro Centroamericano de Población  
Universidad de Costa Rica, 2060 San José, Costa Rica  
<http://ccp.ucr.ac.cr>

---

### **Población y Salud en Mesoamérica**

#### **Revista electrónica**

Volumen 5, número 1, informe técnico 2

**Número especial** *CRELES – Costa Rica: Estudio de Longevidad y Envejecimiento Saludable*

Julio – diciembre, 2007

Publicado 1 de julio, 2007

<http://ccp.ucr.ac.cr/revista/>

### **Mejora de la calidad y disminución de costos de censos y encuestas utilizando computadores de mano (PDA). Una aplicación en Costa Rica**

***Jeisson Hidalgo-Céspedes***

***Luis Rosero-Bixby***

***Daniel Antich-Montero***

## Mejora de la calidad y disminución de costos de censos y encuestas utilizando computadores de mano (PDA). Una aplicación en Costa Rica.<sup>1</sup>

*Jeisson Hidalgo-Céspedes<sup>2</sup>, Luis Rosero-Bixby<sup>3</sup>, Daniel Antich-Montero<sup>4</sup>*

### RESUMEN

Los datos sobre la población en países en desarrollo se obtienen principalmente de las visitas a los hogares. Ejércitos de entrevistadores recolectan esos datos usando cuestionarios en papel (*PAPI, paper and pencil method*), pero tal información se infesta de errores no aleatorios. La reciente tecnología de las entrevistas asistidas por computadora (*CAPI*) es un avance importante sobre *PAPI* en lo que respecta a reducción de errores y los tiempos de espera en la disponibilidad de los datos. Pero la tecnología *CAPI* ha sido de difícil implementación en los países en desarrollo debido a restricciones ergonómicas, costos y delincuencia asociada a las computadoras portátiles (*laptops*). Por esto, el CCP ha desarrollado un sistema *CAPI* para *PDA* (*Personal Digital Assistant* o computadora de mano), que podría revolucionar la forma en que los datos son obtenidos, reduciendo el fraude, errores, tiempos y costos incluso para encuestas complejas. Este artículo reporta las lecciones aprendidas de usar este sistema en el estudio CRELES en Costa Rica.

**Palabras clave:** Recopilación de datos, cuestionarios, entrevista asistida por computadora, encuestas

### ABSTRACT

Population data in developing countries come mostly from household visits. Armies of interviewers collect those data using paper and pencil (the *PAPI* method). But that information is infested with nonrandom errors. Newly developed “computer aided personal interviewing” technology (*CAPI*) is an important advance over *PAPI* in terms of reducing errors and shortening time lags for data availability. But *CAPI* has proved to be difficult to implement in developing countries because of the ergonomics, costs, and delinquency restrictions of *laptops*. In response, we developed a *CAPI* system for *PDAs* that may revolutionize the way data is gathered, by reducing fraud, errors and costs and by opening opportunities for collecting complex information. At the core of our *PDA-CAPI* is an “electronic questionnaire markup language” (*EQML*) we developed after analyzing thousands of existing questions. This paper reports the lessons learned from using our *PDA-CAPI* in a complex survey in Costa Rica.

**Keywords:** Data collection, questionnaires, computer assistant interviewing, surveys

<sup>1</sup> Artículo presentado en la reunión anual de la Asociación de Población de los Estados Unidos de América (*PAA, Population Association of America*), en Filadelfia, abril de 2005.

<sup>2</sup> Centro Centroamericano de Población de la Universidad de Costa Rica, [jhidalgo@ccp.ucr.ac.cr](mailto:jhidalgo@ccp.ucr.ac.cr)

<sup>3</sup> Centro Centroamericano de Población de la Universidad de Costa Rica, [lrosero@ccp.ucr.ac.cr](mailto:lrosero@ccp.ucr.ac.cr)

<sup>4</sup> Centro Centroamericano de Población de la Universidad de Costa Rica, [dantich@ccp.ucr.ac.cr](mailto:dantich@ccp.ucr.ac.cr)

## 1. INTRODUCCIÓN

La encuesta como medio de investigación en las ciencias sociales, representa el avance metodológico más sustancial en los últimos 15 años. Pero no sólo las disciplinas de las ciencias sociales usan encuestas para obtener datos sobre las poblaciones humanas. Virtualmente todas las áreas de análisis de políticas públicas confían en los datos de encuestas para tomar decisiones informadas. La combinación de métodos de muestreo eficientes, confiables y rentables más los diseños de sofisticados cuestionarios, hacen la encuesta el método de recolección de datos de preferencia para, literalmente, miles de estudios cada año, tanto académicos como orientados a políticas.

En los países en desarrollo, todos o al menos la mayoría de los datos provienen de encuestas hechas en la puerta de la vivienda, ya que la telefonía tiene una cobertura limitada y los registros administrativos –tales como seguridad social, padrón electoral, nacimientos y defunciones– son inexistentes o de calidad pobre. Ejércitos de entrevistadores, trabajadores sociales y de salud están recolectando esos datos usando cuestionarios en papel (*PAPI, paper and pencil interviewing method*). Aunque las recientes tecnologías de entrevista asistida por computadora; concretamente la entrevista personal (*CAPI, computer assisted personal interviewing*) y la entrevista telefónica (*CATI, computer assisted telephone interviewing*) son avances importantes sobre el papel, se ha encontrado que su implementación es difícil para los países en desarrollo (Hewett 2003).

Este artículo reporta el desarrollo de una aplicación para recopilar datos estadísticos utilizando computadoras de mano, también conocidas como *PDA (personal digital assistants)*. Aunque la aplicación pretende ser de uso genérico, amigable y gratuito, actualmente se está empleando un prototipo en un complejo estudio de adultos mayores costarricenses, que incluye un cuestionario extenso y complejo para cerca de dos horas de entrevista, otro de dieta en 10 minutos y otros dos asociados a la recolección de muestras y registro de biomarcadores como peso, altura y presión arterial.

Este artículo incluye primero una breve descripción de los errores en la recolección de datos estadísticos de los países en desarrollo, como fundamento para el desplazamiento del *PAPI* al *CAPI*. Se justifica seguidamente la necesidad de una aplicación *CAPI* basada en *PDA*. Se examinan las aplicaciones *PDA* existentes en el mercado contra una lista de características deseables que un sistema *CAPI* debería tener. El resto de secciones en el artículo describen la aplicación desarrollada en la Universidad de Costa Rica (UCR), en particular, enfocada en la experiencia del trabajo de campo con computadores de mano. Se reportan resultados considerando problemas de fiabilidad, uso de Graffiti, hardware, fallos de software y otras lecciones aprendidas.

## 2. ERRORES EN ENCUESTAS

A pesar de la explosión en el uso de las encuestas, existen buenas razones para preocuparse de su calidad en el mundo en desarrollo, ya que los estándares rutinariamente empleados en Estados Unidos y Europa para asegurar la calidad de los datos, no son considerados en Asia, África y Latinoamérica. El tercer mundo invierte considerables cantidades de recursos para obtener información a través de encuestas, la cual sirve de guía para tomar decisiones programáticas y para evaluar sus impactos. Si los datos tienen deficiencias, estos programas y evaluaciones podrían ser inadecuados.

Datos de calidad deben ser tan libres de errores como sea posible. Los errores aleatorios reducen la precisión de las estimaciones, usualmente debilitando las asociaciones verdaderas (error tipo II) y en ocasiones, generando asociaciones falsas (error tipo I). Los estadísticos hacen un buen esfuerzo manejando y documentando los errores aleatorios a través del diseño muestral y análisis inferencial, incluyendo intervalos de confianza. Los errores no aleatorios son más problemáticos debido a que influyen las estimaciones hacia rumbos y magnitudes desconocidas.

En el tercer mundo, donde los mapas, los medios de transporte y las telecomunicaciones son a menudo rudimentarios, los prospectos de cometer errores no aleatorios son muy altos. Mapas desactualizados y pobremente trazados pueden, por ejemplo, guiar a los entrevistadores hacia un punto de la muestra equivocado, y aún cuando los mapas fuesen precisos, la ausencia de señales de información en las vías, números de calles y viviendas, a menudo hacen imposible para los entrevistadores estar seguros de que están visitando la unidad que se pretende. La mayoría de los hogares en los países en desarrollo no disponen de una línea telefónica, especialmente en las áreas rurales; y los teléfonos celulares pese a que cada vez son más populares, aún no son lo suficientemente confiables, lo cual limita el uso de la telefonía para mejorar la calidad de las entrevistas y la supervisión.

La carencia de transportes apropiados y la lejanía geográfica, frecuentemente hacen imposible o simplemente no viable, tener la adecuada supervisión de los entrevistadores tanto para asegurar que la vivienda correcta ha sido encontrada y, más importante, que el enumerador esté realmente realizando la entrevista por completo y no falsificando el total o parte de las respuestas. Otro efecto de los problemas de transporte surge porque a menudo es difícil, sino imposible, retornar los cuestionarios llenos hasta la oficina central, por lo que se toman atajos poco recomendables. Por ejemplo, en algunos países a los entrevistadores se les provee sólo una copia del cuestionario y se les pide registrar las respuestas en una hoja separada, la cual es regresada a la oficina central. Esta práctica propicia los errores debido a la facilidad de marcar la fila o columna equivocada. Tales hojas de códigos son leídas por máquinas en algunos países, y se sabe que éstas son notoriamente poco confiables, aún con equipo de *escaneo* nuevo. En ellas, las respuestas borradas podrían causar errores de lectura. Más aún, las entrevistas realizadas en hogares con poca luz también pueden llevar a preguntas mal leídas o mal contestadas. Es muy común encontrar entrevistadores que inadvertidamente han omitido una pregunta o una serie de ellas por malentender las instrucciones complejas o porque las páginas no estuvieron en el cuestionario por fallos de fotocopiado.

Los errores en los datos recolectados son incontrolables. Existen programas excelentes y ampliamente accesibles para la entrada de datos, como el *CSPro*, disponible gratuitamente (*US Census Bureau*, 2005); aunque poseen robustos mecanismos para limitar la introducción de errores como el chequeo de complejas pautas de consistencia o la verificación de datos; estos se hacen mucho después de que las respuestas, y por ende, los errores hayan sido registrados en el campo. Los retrasos en la entrada de datos son comunes en los países del tercer mundo debido a su imposibilidad de adquirir suficiente equipo tecnológico, creciendo las acumulaciones de trabajo y retrasando la producción de datos por meses o incluso años.

### **2.1 Errores de datos en PAPI y CAPI**

Aunque algunos errores pueden ser detectados con una adecuada revisión de los cuestionarios, un problema inevitable en la recolección de datos con papel (el método *PAPI*) es que la detección y corrección de errores ocurre en tiempos y espacios distintos a la entrevista, con el consecuente costo, dificultad o imposibilidad de retornar al campo para obtener el dato correcto (Hidalgo 2004). El resultado final se traduce en una sustancial cantidad de observaciones con imputaciones o con información faltante o equivocada. Las entrevistas asistidas por computadora (el método *CAPI*) pueden reducir o eliminar esos problemas, haciendo el error evidente en el campo, cuando es fácil repetir una pregunta o solicitar una aclaración al informante (Couper 1998).

El uso de computadoras de escritorio para entrevistas por teléfono (el método *CATI*) es también simple, con la ventaja adicional de que existen aplicaciones disponibles para llevarlo a cabo. Pero, la entrevista telefónica no constituye una alternativa para los países en desarrollo por la baja cobertura de este servicio, lo cual es inadecuado para los diseños muestrales de encuestas poblacionales y de salud.

En principio, el método *CAPI* para recolección de datos en viviendas permite lograr simultáneamente los siguientes objetivos 1) reducir los costos del trabajo de campo y la entrada de datos; 2) incrementar la calidad de los datos obtenidos; y 3) acelerar el trabajo de campo y reducir los tiempos de espera para poder hacer análisis estadísticos con los datos. El auge de las computadoras portátiles ha hecho el *CAPI* posible en las visitas a las viviendas. Sin embargo, el uso de estos equipos en los países en desarrollo ha resultado ser problemático, por lo que los computadores de mano o *PDA* emergen como la opción para reemplazar la entrevista tradicional con papel.

## **3. FUNDAMENTOS PARA UN SISTEMA CAPI CON PDA**

Todos o la mayoría de los datos poblacionales en los países en desarrollo provienen de entrevistas hechas en las puertas de las viviendas. El método *CATI* no puede ser usado porque la mayoría de los hogares carecen de servicio telefónico. A su vez, una barrera sustancial para el uso de computadoras portátiles en el *CAPI* es su costo, que escapa del alcance presupuestario de los países menos desarrollados, en contraposición de aquellos ricos en los que se considera una mercancía común.

Más allá del costo se encuentran otras limitaciones para el uso de *laptops* en la recolección de datos en las naciones menos desarrolladas. Un entrevistador provisto de una *laptop* se convierte de un blanco deseable para los ladrones. Las *laptops* envuelven serios problemas ergonómicos que dificultan la interacción natural con los informantes, y en ciertas circunstancias es simplemente imposible lograr una entrevista con ellas, como ocurre en la puerta de la vivienda, que es el lugar más frecuente en países de alto conflicto o barrios de mucha delincuencia, donde los informantes temen admitir extraños en sus viviendas. En países con tradiciones conservadoras donde a las mujeres les está prohibido ingresar hombres desconocidos en sus hogares, las entrevistas deben realizarse en el umbral, imposibilitando también el uso de *laptops*.

La confiabilidad de las computadoras portátiles también es problemática en el trabajo de campo, en especial en regiones con condiciones difíciles, como calles empolvadas o en mal estado, de alta precipitación tropical, con tormentas de arena o similares, que pueden dañar incluso a las *laptops* más resistentes, cuyo costo puede ser muy elevado para una adopción numerosa. La duración de la batería de estas computadoras también es una restricción a considerar.

Las computadoras de mano o *PDA*, tienen notables ventajas sobre la confiabilidad de las portátiles, en especial por carecer de partes móviles como discos duros, la facilidad con que pueden ser protegidas de las condiciones ambientales y su bajo consumo de energía; sumadas con su bajo costo unitario y la eliminación de los problemas ergonómicos mencionados previamente. La forma en que la información es ingresada en la *PDA* –presionando sobre la pantalla– es otra ventaja importante sobre el teclado de las *laptops*.

El bajo costo y la alta capacidad de almacenamiento de las *PDA* más recientes en el mercado han permitido implementar un sistema *CAPI* en ellas, capaz incluso de trabajar con cuestionarios extensos y complejos. Por ejemplo, un dispositivo con un costo cercano a los \$200 permite hacer entrevistas de 2 horas de duración, almacenar miles de esas entrevistas, identificar los informantes entre las unidades de muestreo, ayudar con listas grandes de categorías y revisar la consistencia de las respuestas ingresadas. Es simple incrementar la capacidad de estas pequeñas computadoras agregando tarjetas de memoria, dispositivos de *GPS* (*global positioning system*), capacidades de fotografía digital, telefonía celular, comunicación inalámbrica, grabación de voz, y otras cualidades; aunque con un costo adicional. El CCP ha estimado que el costo total de adquirir un conjunto de *PDA* para recolección de datos podría equiparar el costo de la entrada de datos de una encuesta con papel y la impresión de cuestionarios, con la ventaja de que el equipo *PDA* es reutilizable para futuras encuestas produciendo un ahorro económico considerable.

Las aplicaciones *CAPI* para *PDA* comparten las ventajas de aquellas para *laptop*. Por citar algunas: control automático del flujo de llenado (pases y filtros), eliminación de inconsistencias durante la entrevista, chequeo de valores fuera de rango, imposibilidad de dejar preguntas sin contestar, reducción de costos por impresión de cuestionarios en papel, codificación y entrada de datos, disponibilidad inmediata de los datos para análisis, entre muchas otras (Hidalgo 2004). Sumadas a esas ventajas, con ayuda de *GPS* se puede evitar el fraude monitoreando que la entrevista se haga en el lugar correcto o en un radio cercano; que el tiempo de cada pregunta y la entrevista total sean razonables; y se pueden registrar en una bitácora todas las acciones que haga el enumerador con la aplicación, lo que permite estudiar en la oficina su comportamiento.

### 3.1 Aplicaciones CAPI con PDA en el mercado

La **Tabla 1** muestra las aplicaciones CAPI existentes en el mercado que emplean PDA para recolección de datos estadísticos identificadas en el momento de escribir este artículo. Se ha encontrado que los países ricos prefieren recurrir a los registros administrativos o dirigir sus encuestas por medios telefónicos, y en algunos casos, utilizan *laptops* para entrevistas personales. La mayoría de las aplicaciones encontradas y listadas en la **Tabla 1** son apropiadas únicamente para investigaciones comerciales sencillas, como las encuestas de mercadeo y opinión, descuidando aquellas de mayor complejidad como censos y encuestas realizadas por los institutos de estadística, universidades y otros entes.

### 3.2 Características deseables en una aplicación CAPI

Un sistema CAPI no consiste únicamente en desplegar un cuestionario en la pantalla y proveer mecanismos para ingresar respuestas. Se ha identificado doce características deseables que éstos sistemas deben proveer, sea tanto con PDA como *laptop*.

1. *Al menos cuatro tipos de preguntas.* Típicamente las encuestas emplean al menos cuatro tipos de preguntas: (1) de respuesta simple, (2) de respuesta múltiple, (3) de respuesta numérica (las fechas y cantidades monetarias pueden considerarse como casos especiales), y (4) de respuesta abierta (textual). Esta es una característica mínima que todo CAPI debe ofrecer, de hecho, todas las aplicaciones encontradas la proveen.
2. *Subpreguntas.* La capacidad de anidar preguntas dentro de otras preguntas. Para ilustrar esta característica, tómese por ejemplo la pregunta de respuesta múltiple en la **Figura 1**. Si el informante dice poseer teléfono residencial, se le pedirá de inmediato cuántos de ellos, lo cual representa otra pregunta pero dependiente de la primer. La capacidad de permitir subpreguntas agrega versatilidad al sistema CAPI. De no tenerla, las preguntas deben manejarse en forma independiente a través de saltos. Cuatro de las aplicaciones en el mercado tienen una versión limitada de esta característica.
3. *Largas listas de categorías.* Algunas preguntas tienen una larga lista de posibles respuestas, como ocurre con la ocupación, enfermedades y el país de nacimiento. La lista es tan extensa que no es práctico mostrarla en una pantalla de PDA o *laptop*, incluso ni en papel, ya que el entrevistador tardaría mucho tratando de encontrar una categoría, que quizá no sea la adecuada. En cuestionarios tradicionales con papel, se registra textualmente la respuesta del informante y luego en la oficina, expertos tratan de asociar con la categoría correspondiente. Pero hay varias formas en que los programas CAPI pueden ayudar a realizar la codificación en el campo, entre ellas, buscar con palabras clave en la larga lista de categorías, discriminar jerárquicamente (por ejemplo, escoger provincia, luego el cantón y finalmente el distrito), o proveer un sistema experto (Hidalgo 2004). Sólo un par de los sistemas CAPI analizados provee un mecanismo sencillo de búsqueda con palabras clave.
4. *Saltos y pautas de consistencia.* Tanto la capacidad de controlar automáticamente el flujo de llenado de la entrevista con saltos y filtros, como la capacidad de chequear que las respuestas registradas sean consistentes, son dos de las más notables ventajas del CAPI sobre el PAPI, ya que reduce los errores en los datos obligando a su corrección en el lugar óptimo –el campo– lo cual mejora la calidad de los datos y reduce costos por revisitas. La mayoría de



aplicaciones en el mercado permiten saltos, pero sólo tres permiten algún tipo de reglas de consistencia y para una esta es su característica más fuerte.

5. *Jerarquía de secciones.* La capacidad de organizar el cuestionario en secciones y subsecciones de preguntas. Esto es especialmente necesario cuando un bloque de preguntas debe reiterarse para un grupo de ocurrencias; por ejemplo, los miembros del hogar, hijos, uniones maritales, trabajos, y redes de apoyo. Poder organizar las preguntas en secciones y subsecciones, facilita programar el flujo de llenado, probarlo y mejorarlo. De las aplicaciones en el mercado, sólo 1 de ellas posee un concepto similar.
6. *Jerarquía de entidades.* Los cuestionarios complejos suelen requerir información sobre varias unidades de observación o entidades, las cuales se organizan jerárquicamente (en niveles). La **Figura 2** muestra una jerarquía de 5 niveles y 6 entidades. Esta característica ahorra memoria, ayuda a mantener los datos ordenados y facilita el procesamiento de las variables. Sólo 1 aplicación en el mercado implementa parcialmente esta característica, limitando a un máximo de tres niveles de entidades.
7. *Identificación jerárquica de observaciones.* La identificación de sujetos y puntos muestrales usualmente se organiza en forma jerárquica. Por ejemplo: país, provincia/estado, municipio/cantón, ayuntamiento/distrito, área de enumeración/*cluster*, vivienda y hogar (Antich 1997). Esta estructura de datos, que usualmente se obtiene del marco muestral, es útil para organizar el trabajo de campo y hacer control procesos. Sólo una aplicación tiene esta cualidad.
8. *Control del proceso de entrada de datos.* Ayudar al personal de supervisión y de oficina tanto en la asignación de entrevistas como su regreso e integración en el computador central, lo que incluye el chequeo de duplicaciones, entrevistas pendientes, revisitas, y similares. Esta característica es necesaria para tener un conjunto de datos limpios, sin duplicaciones u observaciones incompletas que afecten las estadísticas. Cuatro de las aplicaciones analizadas parecen tener estas funcionalidades.
9. *Tiempos, coordenadas espaciales y otros valores precisos.* Integrar a los datos de la entrevista, las horas exactas en que se hicieron las distintas preguntas, las coordenadas geográficas (usando GPS) donde la entrevista se realizó, el dispositivo en que se hizo la entrevista y otros valores útiles para reducir el fraude o retroalimentar al personal de oficina para mejorar el proceso de recolección de datos o su calidad. Sólo una de las aplicaciones en el mercado ofrece esta característica.
10. *Bitácora de eventos.* Mantener registros recuperables de todas las acciones hechas por el entrevistador, como son, selección de una categoría, realizar una corrección, emplear Graffiti o el teclado en pantalla. Esta característica es útil para control de calidad del trabajo de campo, detectar y corregir errores sistemáticos, recuperar información como una entrevista eliminada accidentalmente por el enumerador, y otras posibilidades. Ninguna aplicación en el mercado parece ofrecerla.
11. *Codificación de preguntas abiertas.* Para preguntas que no se conoce todas sus posibles respuestas, se pueden listar las más frecuentes y permitir agregar nuevas en el campo en forma textual. Un sistema CAPI inteligente, puede ayudar al entrevistador a agrupar y codificar nuevas categorías a partir de los textos, y regresar éstas al campo antes de que termine la recolección de datos, así la pregunta se va “cerrando” en el tiempo. Por ejemplo, se puede incluir una pregunta abierta para registrar las medicinas tomadas por los entrevistados. Pronto se notará que la mayoría de personas toman las mismas medicinas, éstas se codifican y se integran a las opciones de la pregunta periódicamente, lo cual acelera el trabajo del

entrevistador y de los codificadores. Ninguna aplicación en el mercado parece disponer de esta característica.

12. *Información externa.* Agregar a cada observación información externa como una imagen capturada o sonido grabado durante la entrevista. Sólo una aplicación en el mercado lo ofrece.

#### 4. LA APLICACIÓN CAPI DEL CCP

El Centro Centroamericano de Población (CCP) de la Universidad de Costa Rica, ha desarrollado un sistema *CAPI* llamado *EQ-Software*, el cual se ha usado en varias encuestas y se mantiene en uso, aunque no se ha implementado aún toda su funcionalidad (**Figura 3**).

La aplicación para *PDA* es sólo una parte del *EQ-Software*, un complejo sistema para recolección de datos estadísticos cuya arquitectura se aprecia en la **Figura 5**. El *EQ-Software* provee aplicaciones para ayudar en distintas actividades del proceso de recolección de datos, pero todas ellas utilizan un elemento común, el cuestionario escrito en una notación especial diseñada para este fin que lleva el nombre de *Electronic Questionnaire Markup Language (EQML)*.

El CCP desarrolló el *EQML* tras analizar matemáticamente miles de preguntas en boletas de censos y encuestas complejos de varios países latinoamericanos (Hidalgo 2004). Con miras de estandarización se desarrolló sobre *XML* (Goldfarb 1999) y puede representar formalmente cuestionarios complejos, con un número arbitrario de entidades y niveles (ej. Viviendas, hogares, personas, hijos, vehículos), secciones, subsecciones, preguntas, subpreguntas, variables de diferentes tipos, saltos, pautas de consistencia, instrucciones al enumerador y más. La **Figura 4** muestra una pregunta simple de la Encuesta de Hogares 2005 de Costa Rica (INEC 2006) escrita en *EQML*.

Los tres módulos que componen el *EQ-Software* de la **Figura 5**, utilizan cuestionarios electrónicos y son los siguientes.

1. *EQ-Design.* Para poder recolectar datos, el investigador debe escribir un cuestionario electrónico en *EQML*, lo cual podría hacerse en un editor de texto. El *EQ-Design* ayuda al investigador a preparar cuestionarios válidos en forma visual y amigable en una computadora de escritorio (PC). Hasta la fecha, este módulo no se ha implementado.
2. *EQ-Control.* El cuestionario no es suficiente para iniciar la recolección de datos. Se requiere definir la lista de entrevistadores y supervisores que trabajarán en el campo, qué sitios recorrerá cada cual (áreas de enumeración) y qué equipo usará. El personal de oficina necesita una aplicación que ayude a hacer estas asignaciones, así como extraer los datos que hayan recopilado los entrevistadores, acumularlos en una base de datos, verificar que no hayan observaciones duplicadas ni faltantes y exportar tales datos para análisis estadístico a aplicaciones como *Stata*, *SPSS*, *CSPRO* o similares. El *EQ-Control* es el módulo del *EQ-Software* que se encarga de estas actividades.
3. *EQ-Collector.* Esta aplicación recibe el cuestionario y las áreas donde debe trabajar el entrevistador, y recorre dicho cuestionario registrando las respuestas provistas por los informantes. Automáticamente controla el flujo de llenado, aplica las pautas de consistencia mientras almacena las respuestas. A la fecha, el CCP ha programado un *EQ-Collector* para

*Palm OS* (Foster 2002) en el lenguaje C++ (Stroustrup 1997), capaz de trabajar con cuestionarios complejos y extensos en forma muy eficiente pese a las limitaciones de estos dispositivos. El *EQ-Software* admite incorporar *EQ-Collectors* para otras plataformas, como podrían ser *Pocket PC*, *web* e incluso, *papel*.

#### 4.1 Experiencias del EQ-Software

Los programas parciales del *EQ-Software* se han empleado con resultados exitosos en la recolección de datos de cinco encuestas hasta la fecha. Se describen a continuación.

La Encuesta de Observación de Vecindarios 2004, a cargo del CCP entre el 29 enero al 10 de marzo del 2004, con un equipo de 4 enumeradores. Se recolectaron 456 observaciones (Hidalgo 2004).

El CRELES (Costa Rica: Estudio de Longevidad y Envejecimiento Saludable). También del CCP. Inició su trabajo de campo el 3 de noviembre del 2004 y terminará en el 2008 para aplicar un complejo cuestionario con más de 600 preguntas, 2 horas de entrevista, a 3000 adultos mayores de todo el país, en dos rondas distanciadas 2 años aproximadamente.

La Encuesta Sobre Corrupción 2005. Realizada entre el 19 de noviembre al 16 de diciembre del 2005 en República Dominicana a cargo de la Pontificia Universidad Católica Madre y Maestra, y el Centro de Estudios Sociales y Demográficos. Se recolectaron 1512 observaciones.

El *Latin American Public Opinion Project (LAPOP) 2006*. Encuesta realizada simultáneamente en 17 países, de los cuales 5 cambiaron el papel por *palm* entre junio y julio del 2006; en concreto, El Salvador, Honduras, Costa Rica, Panamá y República Dominicana. Se recopilieron 1500 observaciones en cada país.

La Encuesta de Exclusión Social. Realizada en Venezuela, Colombia, Uruguay, Perú, Argentina y Costa Rica. En este último con *palms* en abril del 2007.

De la experiencia acumulada en más de 10,000 horas de entrevista, se extrae el siguiente reporte, especialmente enfocado en el estudio CRELES.

#### 4.2 Problemas de hardware

La aplicación de entrada de datos, el *EQ-Collector*, se ha empleado con los siguientes equipos y modelos: *Palm m125*, *Sony Clié SJ22*, *Zire 22*, *Zire 31*, *Zire 71*, *Zire 72*, *Tungsten E* y *Tungsten E2*, cuyo precio varía –o varió– entre US \$100 a \$200 en el mercado. En más de 10,000 horas de entrevista, se han observado diferentes comportamientos referentes al fallo de hardware. Se describen estos en tres categorías, los debidos a un modelo particular, errores de fábrica y el daño normal por uso.

Con un equipo de 8 *palms* se ha encontrado una tasa de fallos severos estimada a 1 por cada 1500 horas de entrevista. Se dice severo porque el dispositivo tuvo que dejar de usarse y por ende,

sustituirse. De los 6 dispositivos *Zire 71* adquiridos por el CCP, en 4 se reiteró un defecto de pantalla que impedía la introducción de datos. No hubo pérdidas de información por este inconveniente, y no reincidió en ningún otro modelo, lo que hace sospechar que era deficiencia únicamente en los usados por el CCP.

Tanto los dispositivos para sincronizar las *Zire 71*, a veces llamados “cunas” o *cradles* en inglés (**Figura 6**), como las mismas *palms*, sufren desgaste en los pines donde ambos hacen contacto. Al cabo de dos años de uso, el mecanismo dejó funcionar y estos dispositivos no pueden sincronizarse por este medio al computador de escritorio. Posiblemente los ingenieros de *Palm Inc.* visualizaron este defecto porque sus modelos posteriores aparecieron con mejores conectores, que a la fecha continúan trabajando apropiadamente en el CCP.

Sin embargo, la tasa de fallos presentada está fuertemente influenciada por las *Zire 71*. Los otros equipos adquiridos por el Centro han mostrado mucha mejor confiabilidad. Por ejemplo, las 2 *Sony Clié SJ22* no han fallado hasta la fecha y fueron adquiridas varios meses antes que las *Zire 71*. Si se ignoraran las fallas de las *Zire 71* los errores a considerar serían los provocados por desgaste normal.

La mayoría de fallos físicos encontrados en la *palms* se pueden considerar como debidos al desgaste natural por su uso, el cual es el factor más importante. Este tipo de fallos no se consideran severos porque el dispositivo puede continuar usándose aunque con un poco de dificultad. Pero el deterioro es progresivo en el tiempo y tarde o temprano se tendrá que reemplazar el dispositivo, lo cual se convierte en un fallo severo. El reemplazo no se hace en forma gradual sino en ráfagas, ya que normalmente los dispositivos se adquieren en grupo y se deterioran en forma similar durante el trabajo de campo.

La región más afectada del dispositivo es su pantalla sensible, la cual se desgasta por el contacto con el lápiz. Se aconseja fuertemente la adquisición de plásticos adherentes (de la misma marca que el dispositivo), que cubren, protegen y mejoran la escritura en ella. Estos deben reemplazarse periódicamente (cada uno o dos meses) y se debe educar a los enumeradores para que no golpeen grotescamente la pantalla.

Hay otra forma de fallo a considerar en dispositivos que nunca fueron utilizados. En la primera oportunidad que el CCP adquirió un grupo de *palms*, 6 unidades de *Zire 71*, una de ellas presentó dificultad para cargarse. Dos años después se compraron 12 *Tungsten E*, y una de ellas presentaba una región insensible de la pantalla. Un comportamiento similar ocurrió en el Centro de Estudios Sociales y Demográficos (CESDEM) de República Dominicana, al adquirir 20 *Zire 31* para la encuesta LAPOP; una de ellas no funcionaba adecuadamente. Esta reincidencia hace pensar en la probabilidad del equipo defectuoso de fábrica, como un problema para adquisiciones que se hacen en el extranjero o por Internet sin garantía local.

#### 4.2.1 Capacidad de la memoria y pila

Mucha curiosidad gira alrededor de las capacidades de almacenamiento y horas de trabajo que puede brindar las *PDA*. Las *palms* existentes a la fecha cuentan con un promedio de 28MB de memoria libre para almacenar entrevistas. La cantidad de observaciones que se pueden acumular

varía dependiendo del número de variables que tenga el cuestionario. En el caso del extenso cuestionario CRELES un dispositivo puede acumular unas 9500 entrevistas, y en el de LAPOP unas 57,500. Pese a estas capacidades, no es recomendable acumular las entrevistas en los dispositivos. Estas deben moverse a la computadora de escritorio con frecuencia y en ella mantener respaldos.

Si las *palm* son correctamente cargadas durante la noche, la duración de la pila alcanza para unas 4 ó 5 horas de entrevista aproximadamente. Lo cual puede ser o no suficiente dependiendo de la encuesta. En el caso del CRELES cada entrevistador realiza en promedio dos visitas diarias, cada una de dos horas aproximadas, por lo que basta con la carga nocturna. Pero para otras encuestas las *PDA* debieron usarse por 7 o más horas diarias. Se decidió proveer al entrevistador del cargador que acompaña de fábrica al dispositivo. Los enumeradores recargaban los equipos mientras pausaban para almorzar. En pocas ocasiones se solicitó corriente eléctrica en el lugar de entrevista. Si el equipo de trabajo se transportaba en algún medio vehicular, el uso de cargadores adaptados al encendedor de cigarrillos resultó de utilidad. Desdichadamente sólo modelos discontinuados como las *Zire 71*, cuentan con útiles baterías externas distribuidas como accesorios, las cuales permiten cargar el dispositivo mientras se está empleando en la palma de la mano.

Si la *palm* llega a alcanzar un muy bajo nivel de energía, ésta se auto apaga y espera a ser cargada. En los modelos más antiguos, si no se recarga en unas horas, llega a perder los datos que en su memoria interna tenga acumulados. En la encuesta LAPOP esto ocurrió accidentalmente a un supervisor, quizá porque alguno de los botones de la *palm* fue presionado continuamente por otro objeto en el contenedor que le transportaba, provocando un voraz consumo de energía. Los datos fueron recuperados de inmediato gracias al respaldo periódico en tarjetas de memoria.

#### 4.3 Problemas de software

También se ha experimentado con problemas en el software, principalmente debidos a errores en el diseño del cuestionario que son detectados durante el trabajo de campo. Ejemplo son, saltos errados, rangos numéricos muy estrechos o pautas de consistencia muy fuertes. Estos aparecen con más frecuencia durante los primeros días de trabajo y son corregidos rápidamente en forma conjunta entre investigadores e informáticos (Antich 1997).

Una vez que se ha iniciado el trabajo de campo, los cambios que se hagan en el cuestionario pueden resultar muy peligrosos, ya que la estructura de los datos recolectados (en las *PDA* o la computadora de escritorio) podría ser incompatible con los cambios que quieren introducirse, exigiendo una reestructuración en miles o millones de registros. Existe controversia si esto debe permitirse o no. La mayoría de aplicaciones lo prohíben. Pero de la experiencia adquirida con las encuestas realizadas por el CCP, se evidencia que es una necesidad con mucha importancia, proporcional a la duración del trabajo de campo, como ocurre en el estudio CRELES con dos fases de recolección cada una de 2 años.

El *software* de recolección también ha presentado errores (en inglés, *bugs*), los cuales han afectado ligeramente la información recolectada. Una entrevista se perdió por completo y un par han tenido que completarse con papel. Estos errores son corregidos tan pronto como son

detectados. Debe recordarse que los programas actualmente usados están aún en etapa de desarrollo.

Hasta la fecha, los cuestionarios han sido diseñados para papel por los investigadores, y los informáticos del CCP los traducen a *EQML*. Encuesta tras encuesta, en este proceso se hace evidente una enorme cantidad de errores en el diseño, gracias a la rica estructura del *EQML* para representar cuestionarios bien definidos.

Un error en el hardware, el cuestionario o los programas, puede ocurrir inesperadamente, por lo que en todas las encuestas se ha provisto de cuestionarios en papel a los entrevistadores, para situaciones de emergencia. Cuando se ha entrevistado con el papel, las respuestas son posteriormente introducidas por el enumerador a su *palm* de trabajo y finalmente llevadas al computador principal en formato electrónico como de costumbre.

#### **4.4 Problemas de aceptabilidad**

Todas las innovaciones usualmente generan resistencia. Se ha dicho que tanto los informantes como entrevistadores podrían rechazar la nueva tecnología. En la práctica se ha notado lo opuesto. Los enumeradores se han expresado intensamente a gusto con el dispositivo, en especial porque éste controla la secuencia de llenado, y han dicho no querer regresar a las entrevistas con papel. De igual forma, no se han recibido reportes de rechazo de los informantes por el uso del *PDA*. Por el contrario, a menudo los entrevistados se muestran intrigados y maravillados con el dispositivo, en especial al mirar su fotografía en la pequeña pantalla, como se hace en CRELES.

La capacitación de los entrevistadores en el uso del *PDA* es normalmente sencilla. La mayoría de ellos ya ha tenido experiencias con encuestas en papel, pero ninguno con *palms* o *CAPI*. Las operaciones básicas con el dispositivo, como marcar en la pantalla o incluso, escribir texto, les resultan intuitivas, en especial a las personas más jóvenes. Quizá esto se deba a la familiaridad con otros elementos tecnológicos como videojuegos o teléfonos celulares.

Existe mucha preocupación asociada a la posibilidad de asaltos o robos por el uso de *PDA*. Hasta la fecha, el CCP no ha sufrido este inconveniente, pero sí ocurrió en la encuesta *LAPOP* de República Dominicana con uno de los dispositivos. Gracias al respaldo en tarjetas externas de memoria, los datos se recuperaron por completo; lo cual es una práctica recomendada. En caso de entrevistar en lugares de mucha delincuencia, si al enumerador no se le permite ingresar a las viviendas, se hace la entrevista con papel como medida preventiva, y luego en un lugar seguro se transcriben los datos del papel a la *palm*.

#### **4.5 Ingreso de los datos**

Hubo ciertas preocupaciones sobre posibles errores con la entrada de datos en los dispositivos *PDA*, como por ejemplo marcar la respuesta equivocada en la pantalla. En general, esta posibilidad de error mantiene la misma probabilidad tanto al marcar en el papel, hacer clic con el ratón de una *laptop* o presionar una tecla equivocada en una máquina de escritorio.

Para comparar la confiabilidad del ingreso de datos en *PDA* contra papel, en cada entrevista de la prueba piloto del CRELES se incluyeron dos entrevistadores, uno registraba respuestas en papel y el otro en *palm* simultáneamente. En más de 5,000 repuestas se encontró una tasa de confiabilidad, o coincidencia, del 96% (**Tabla 2**). La discrepancia fue de un 3% en preguntas con dos o tres categorías u opciones de respuesta, las cuales son las más comunes en los cuestionarios (Hidalgo 2004). Las discrepancias se duplicaron para preguntas con muchas categorías o con respuestas numéricas. No hay forma de determinar con esos datos si los errores provienen mayoritariamente del *PDA*, del papel o de ambos. Considerando que durante esta prueba piloto se estaba entrenando a los entrevistadores y se estaba probando el cuestionario, es muy probable que en una encuesta real la tasa de discrepancia baje y la confiabilidad del ingreso de datos en la pantalla sea superior al 99%.

La *Palm* provee dos mecanismos para ingresar texto en el dispositivo. El primero es escribiendo directamente en un área especial de la pantalla, de la misma forma que se hace sobre un papel. *Palm* interpreta los trazos hechos y los traduce en las letras o números correspondientes. Este mecanismo se conoce como *Graffiti*. El segundo corresponde a invocar un teclado en pantalla, similar al de una computadora de escritorio, en el cual el enumerador puede ir presionando cada tecla para formar palabras.

El uso de *Graffiti* para ingresar texto o información numérica en el *PDA* también fue una fuente de inquietudes, referentes a la curva de aprendizaje de los entrevistadores, la eficiencia de la escritura en pantalla y su precisión. La existencia de bitácoras que registran las acciones emitidas por el enumerador al usar el *PDA*, permite examinar algunos de estos comportamientos. Los resultados siguientes están basados en datos de dos preguntas, una textual y otra numérica en alrededor de 300 entrevistas realizadas durante 8 semanas de prueba de campo del CRELES. Las preguntas examinadas corresponden al nombre de todos los miembros del hogar (alfanumérica, en promedio 3 nombres por vivienda) y la edad de cada uno de ellos (numérica).

Los entrevistadores usaron *Graffiti* el 57% del tiempo de acuerdo con las bitácoras citadas (**Tabla 3**). El uso de *Graffiti* aparece un 25% más alto para números que para texto. Contrario a las expectativas, no hubo una “curva de aprendizaje”, la cual habría incrementado el uso de *Graffiti* en el tiempo. Parece que la escogencia de *Graffiti* o teclado se hace desde un inicio y cada entrevistador se mantiene firme a su preferencia. Tres entrevistadores prefieren el *Graffiti*, uno lo utiliza principalmente para números y otro lo emplea en raras ocasiones (**Tabla 3**).

Los datos crudos sugieren que el *Graffiti* es menos eficiente que el teclado en pantalla, ya que requiere un 22% más de tiempo para ingresar un carácter (**Tabla 4**). Aunque tal asociación es confundida por el uso diferencial de *Graffiti* por entrevistadores y el tipo de información (número o texto) ingresado. Después de ajustar esos modificadores con una regresión de Poisson, la escritura de *Graffiti* muestra ser un 33% más eficiente.

Aunque no se tienen datos precisos sobre la confiabilidad o precisión de la escritura con *Graffiti*, las bitácoras de la *PDA* permiten determinar la tasa de corrección de caracteres introducidos con este método, la cual es de 37 correcciones por cada 100 caracteres ingresados (**Tabla 5**). Las correcciones hechas a la hora de introducir un texto son 4 veces más frecuentes que cuando se

ingresa un número. Los datos también muestran una pequeña mejora en la tasa de correcciones en el tiempo, pero está restringido a las dos primeras semanas únicamente.

#### 4.6 Transferencia de información

Un componente, a menudo descuidado, del proceso de recolección de datos con *CAPI* es el que se encarga de asignar entrevistas a los enumeradores, acumular la información recolectada por los *PDA*, respaldarla y exportarla a otras aplicaciones como *Stata* o *SPSS*. Este módulo se le ha llamado *EQ-Control* en la arquitectura del *EQ-Software* (**Figura 3**).

Existe una muy justificable preocupación de que el riesgo de pérdida de datos podría incrementar al pasar del papel al *CAPI*, en especial por pérdidas catastróficas como fallos de hardware o software en el equipo central donde se acumulan los datos, errores humanos, accidentes o robos. Aunque estos peligros también se encuentran presentes en la recolección con papel, en el computador es menos visible ya que una persona podría no notarlo.

La aplicación de control puede evitar desastres al proveer mecanismos automáticos de redundancia de información (respaldos), lo cual es una notoria ventaja sobre los cuestionarios en papel. De igual forma, los datos recolectados en *PDA* pueden respaldarse en tarjetas de memoria que los enumeradores deben transportar en forma independiente al dispositivo, así se evita la pérdida de trabajo en caso de robo o fallo de hardware y software. Hasta la fecha, la pérdida de datos con *CAPI* ha sido prácticamente nula.

### 5. RESULTADOS Y DISCUSIÓN

Este artículo reporta la experiencia en el uso de computadoras de mano, también conocidas como *Personal Digital Assistants (PDA)*, para recopilar información en varias encuestas, principalmente del complejo estudio CRELES en Costa Rica. Se ha encontrado que las *PDA* ofrecen una alternativa para mejorar sustancialmente la calidad de los datos y reducir tanto el costo como el tiempo de espera para disponer de ellos, en encuestas hechas en los países en desarrollo. Dadas las crecientes capacidades de procesamiento, memoria y escalabilidad de las *PDA*, sumado con su progresiva caída de precio, se estima que tales aplicaciones podrían transformar la forma en que los datos poblacionales son recopilados en los países en desarrollo.

La tecnología *CAPI*, de uso común con *laptops* en países desarrollados, es una conocida mejora sobre la recolección con papel. La ventaja más importante del *CAPI* se debe a que la detección y corrección de errores ocurren en el mismo tiempo y lugar: la entrevista; sumado con el ahorro de costos por digitación y la disponibilidad inmediata de los datos para análisis estadístico. Por el alto costo, ambiente adverso y limitaciones ergonómicas, las *laptop* no han sido adoptadas en los países en desarrollo. Las *PDA* parecen ser la alternativa.

Aunque el mercado ofrece varias aplicaciones *CAPI* con *PDA*, todas ellas tienen limitaciones no apropiadas para los requerimientos de las encuestas sociodemográficas o de salud. Estos



programas son comúnmente costosos y adecuados para encuestas de mercado, con cuestionarios y diseños muestrales sencillos.

El Centro Centroamericano de Población (CCP) de la Universidad de Costa Rica está desarrollando una aplicación que llena las necesidades de las encuestas complejas. Se ha usado un prototipo en varios estudios con resultados muy satisfactorios. Este artículo comparte esta experiencia, en especial, la debida al ambicioso estudio longitudinal de envejecimiento CRELES, financiado por la fundación *Wellcome Trust*.

El sistema en desarrollo por el CCP es una compleja arquitectura que provee aplicaciones para ayudar al investigador en el diseño del cuestionario, administrar la lista de áreas de enumeración, repartirla entre los entrevistadores, dirigir automáticamente la entrevista, registrar las respuestas y verificar su consistencia, concentrar la información recolectada en una base sin redundancias ni faltantes, y exportar tales datos a programas de análisis estadístico. Todas las aplicaciones emplean un rico y robusto lenguaje para representar el cuestionario llamado *EQML*, construido por el CCP tras analizar miles de preguntas de cuestionarios de censos y encuestas latinoamericanos.

El uso de prototipos de las aplicaciones durante más de 10,000 horas de trabajo de campo ha demostrado que, (1) el sistema es factible; (2) es bien aceptado por entrevistadores e informantes; (3) cometer errores de ingreso de datos (escoger una respuesta incorrecta) es un problema infrecuente; (4) tanto respuestas textuales como numéricas pueden ser ingresadas eficientemente usando *Graffiti* o el teclado en pantalla; (5) el hardware es confiable pero sufre de desgaste normal; (6) la duración de la pila es suficiente para 4 ó 5 horas de entrevista y puede recurrirse a estrategias por si se requiere más; (7) la pérdida de datos es 100% evitable.

Aunque el sistema esté trabajando adecuadamente, es claro a partir de las experiencias reportadas que considerables refinamientos necesitan implementarse para disponer de una aplicación para uso general. El CCP está actualmente trabajando en esta dirección. Esta es una de esas raras ocasiones en las que tecnología nativa ha sido desarrollada localmente, aunque esto no debe sorprender, ya que la exportación de software en Costa Rica está creciendo aceleradamente. Tal tecnología está ideada para responder a las necesidades específicas de los países en desarrollo; puede incrementar la calidad, disminuir los costos y el tiempo de espera para disponer de datos; con su consecuente impacto positivo en la investigación y toma de decisiones.

## 6. REFERENCIAS

- Antich Montero, Daniel. *Guía metodológica para el procesamiento de censos y encuestas por computador*. Instituto Tecnológico de Costa Rica. Cartago, Costa Rica. 1997. Tesis para optar por el grado de Magíster Scientiae en Computación
- Couper, Mick P., et al. *Computer Assisted Survey Information Collection*. Ley-Interscience, United States of America, 1998
- Foster, Lonnon. *Palm OS Programming Bible, Second Edition*. Wiley Publishing, United States of America, 2002
- Goldfarb, Charles y Prescod, Paul. *Manual de XML*. Prentice Hall, Madrid, España, 1999
- Hewett, Paul C., et al. *The feasibility of Computer-Assisted Survey Interviewing in Africa: Experience form Two Rural Districts in Kenya*. Population Council, Working Papers No. 168, 2003.
- INEC, Instituto Nacional de Estadística y Censos. *Encuesta de Hogares de Propósitos Múltiples, Julio 2006; instructivo del entrevistador*. San José, Costa Rica, 2006.
- Hidalgo Céspedes, Jeisson. *Diseño de un modelo computacional y desarrollo de un prototipo para la recolección de datos de censos y encuestas con computador de mano*. Universidad de Costa Rica, San José, Costa Rica, 2004. Tesis de maestría en Computación en Informática
- Naciones Unidas. *Manual de procesamiento de datos de censos de población con microcomputadoras*. New Work, 1982.
- Stroustrup, Bjarne. *The C++ Programming Language*. 3rd ed. Addison-Wesley, United States of America, 1997
- U.S. Census Bureau. *Census and Survey Processing System*. 2005. URL: <http://www.census.gov/ipc/www/csprol>
- U.S. Bureau of the Census. *1994 Annual Research Conference and CASIC Technologies Interchange: Proceedings*. USA, 1994.

## **RECONOCIMIENTOS**

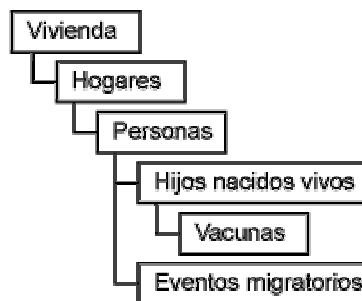
El proyecto "Costa Rica: estudio de longevidad y envejecimiento saludable (CRELES)" es una investigación de la Universidad de Costa Rica, ejecutada por el Centro Centroamericano de Población con la colaboración del Instituto de Investigaciones en Salud y con subvención de la Fundación Wellcome Trust.

- Investigador Principal: Luis Rosero-Bixby.
- Co-investigadores: Xinia Fernández y William H. Dow.
- Investigadores colaboradores: Ericka Méndez, Guido Pinto, Hannia Campos, Kenia Barrantes, Floribeth Fallas, Gilbert Brenes, Fernando Morales.
- Personal de informática y apoyo: Daniel Antich, Aaron Ramírez, Jeisson Hidalgo, Juanita Araya y Yamileth Hernández.
- Personal de campo: José Solano, Julio Palma, Jenny Méndez, Maritza Aráuz, Mabelyn Gómez, Marcela Rodríguez, Giovanni Salas, Jorge Vindas y Roberto Patiño.

**Figura 1. Ejemplo de subpreguntas (INEC 2006)**

16. ¿Tienen en esta vivienda...	¿Cuántos?	
	Sí	No
...teléfono residencial? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...refrigeradora? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...horno de microondas? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...ducha para agua caliente? ...	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...tanque para agua caliente? ..	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...lavadora? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...tanque para almacenar agua? ...	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...computadora? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...fax (no de computadora)? ...	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...equipo de sonido? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...VHS ó DVD? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...vehículo (no de trabajo)? ....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...televisor a color? .....	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...televisión por cable ó satélite? .	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>
...radio (no equipo de sonido)? ...	O1 <input type="checkbox"/>	O2 <input type="checkbox"/>

**Figura 2. Una jerarquía de 5 entidades**



**Figura 3. Una pantalla del EQ-Collector para Palm**

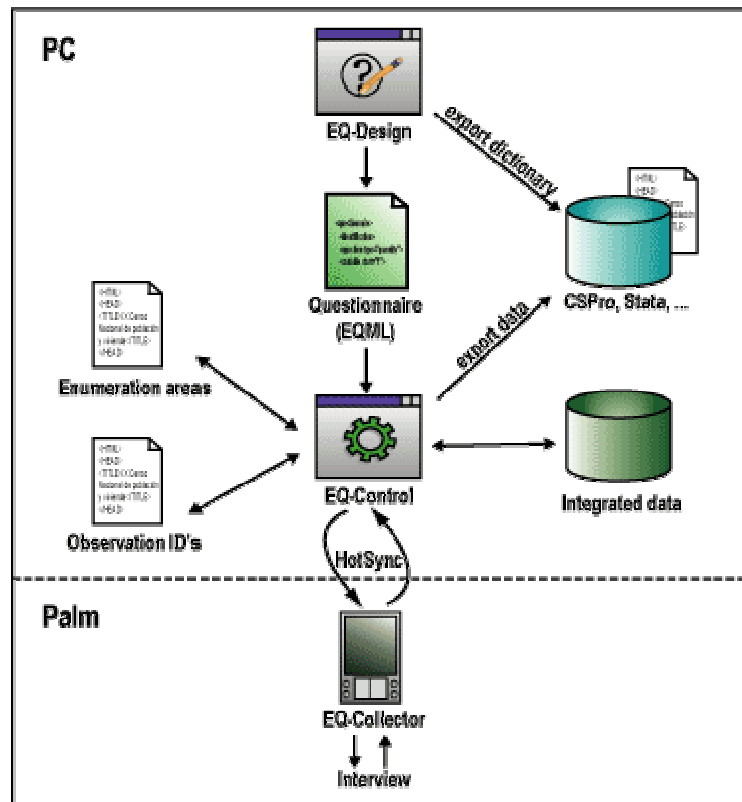


**Figura 4. Ejemplo de una pregunta simple en EQML y su representación en la Encuesta de Hogares (EHPM2005)**

```
<question name="pV12">
  <text>¿Tiene esta vivienda servicio sanitario...</text>
  <categories>
    <category code="1">conectado a alcantarilla o cloaca?</category>
    <category code="2">conectado a tanque séptico?</category>
    <category code="3">de pozo negro o letrina?</category>
    <category code="4">con otro sistema?</category>
    <category code="5">No tiene</category>
  </categories>
  <variable name="V12" label="Servicio sanitario" type="cat"/>
</question>
```

12. ¿Esta vivienda tiene servicio sanitario...	
...conectado a alcantarilla o cloaca? .	01
...conectado a tanque séptico? .....	02
...de pozo negro o letrina? .....	03
...con otro sistema? .....	04
No tiene .....	05

**Figura 5. Arquitectura del EQ-Software**



**Figura 6. Cuna para Zire 71**



**Tabla 1. Comparación de aplicaciones CAPI con PDA existentes en el mercado. La lista fue elaborada hacia inicios del 2005 y no pretende ser exhaustiva**

Aplicación	Plataformas	Costo	Características	Referencia
Survey View	Web, PocketPC, <del>Palm</del> , paper	\$2000	1, 4	<a href="http://www.surveyview.com/">http://www.surveyview.com/</a>
Snap	Web, E-mail, Paper, Kiosk, Phone, PDA, Scanning, TabletPC	\$1300 + \$1000 (primeros 5 usuarios) + \$600*(siguientes grupos de 5 usuarios)	1, 2, 4	<a href="http://www.mercatorcorp.com/">http://www.mercatorcorp.com/</a>
Pendragon Forms	Palm	\$230 + \$55 additional user	1, 4, 9, 12	<a href="http://www.pendragonsoftware.com/">http://www.pendragonsoftware.com/</a>
Ethnos CAPI	Paper (scanner), Web, CATI, PocketPC	5600€ + 1000€ cada grupo de 5 licencias	1, 2, 4, 8	<a href="http://www.soft-concept.com/ethnos">http://www.soft-concept.com/ethnos</a>
Entryware Pro	Palm	Depende de cada proyecto	1, 4	<a href="http://www.techneos.com/">http://www.techneos.com/</a>
SyncSurvey	Palm	Tarifas mensuales dependiendo de la cantidad de usuarios e informantes	1	<a href="http://www.syncsurvey.com/">http://www.syncsurvey.com/</a>
MobileSurvey	Paper, Web, Email, Palm, PocketPC, Symbian	\$1000 + \$195 por PDA	1, 2, 4	<a href="http://www.perseus.com/">http://www.perseus.com/</a>
EZSurvey	Paper, Web, Email, Palm	\$400	1, 4	<a href="http://www.raosoft.com/">http://www.raosoft.com/</a>
AccessPoint	Palm, PocketPC, Tablet	No encontrado	1, 4, 8	<a href="http://www.globalbay.com/">http://www.globalbay.com/</a>
MobileMemoir	PocketPC, Tablet, Laptop, PC	No encontrado	1, 2, 4	<a href="http://mobilememoir.com">http://mobilememoir.com</a>
CSPProX	Papel, PC, Web, PocketPC	El cliente debe consultar	1, 3, 4, 6, 7, 8	<a href="http://www.serpro.com/">http://www.serpro.com/</a>

Nota: los números en la columna de características se explican a continuación en el texto

**Tabla 2. Discrepancia PDA vs. Papel**

Tipo de pregunta	Discrepancia	Repuestas comparadas
Con 2 a 3 categorías	3.1%	3,276
4 o más categorías	6.2%	1,170
Numéricas	7.0%	884
Total	4.4%	5,330

**Tabla 3. Porcentaje de quién uso Graffiti**

	Textos	Números	Ambos
<b>En el campo</b>			
1-2 semanas	48%	82%	57%
3-6 semanas	51%	74%	57%
7-8 semanas	55%	69%	58%

<b>Entrevistador</b>			
Entrevistador 1	3%	11%	5%
Entrevistador 2	100%	100%	100%
Entrevistador 3	98%	100%	99%
Entrevistador 4	4%	76%	21%
Entrevistador 5	85%	94%	88%
<b>Total</b>	<b>52%</b> <b>(906)</b>	<b>75%</b> <b>(308)</b>	<b>57%</b> <b>(1,304)</b>

**Tabla 4. Duración de la entrada de datos con Graffiti y teclado en pantalla (segundos por carácter)**

Mecanismo	Observado	Ajustado*
Teclado en pantalla	1.85	2.06
Graffiti	2.25	1.37
<b>Tasa Graffiti / Teclado</b>	<b>1.22</b>	<b>0.67</b>

\*Controlado (con regresión de Poisson) por entrevistador, texto/número y período de trabajo de campo.

**Tabla 5. Tasa de correcciones en Graffiti (por cada 100 caracteres)**

En el campo	Textos	Números	Ambos	Ajustado*
1 a 2 semanas	43%	8%	38%	41%
3 a 6 semanas	40%	11%	36%	36%
7 a 8 semanas	43%	9%	39%	36%
<b>Total</b>	<b>41%</b>	<b>10%</b>	<b>37%</b>	<b>37%</b>

\* Controlado con Regresión de Poisson por entrevistador y texto/número.