**ALTERNATIVA LIBRE PARA LA RECOLECCIÓN DE DATOS DE CAMPO PROPUESTA PARA EL GRUPO DE INVESTIGACIÓN GEOMATICA**

El presente trabajo de investigación es elaborado por Daniel Carvajal Patiño y dirigido por Gustavo Diaz en el marco del 8° Diplomado DUNTI.

**RESUMEN**

Se planea encontrar, determinar y configurar una alternativa software para ArcGIS, un software de recolección de datos utilizado actualmente para el grupo de investigación Geomatica. Dado que se trata de un grupo de investigación, el cual presenta una alta tasa de elaboración de proyectos que requieren de trabajo de campo, necesita una metodología eficiente para la recolección de datos que posteriormente serán usados en la elaboración de dichos proyectos. Para Geomatica ‘ArcGIS’ es una herramienta que presenta características poco favorables y desean cambiar dicho software por uno que les sea más favorable.

**PALABRAS CLAVE**

Recolector de Datos, Software Libre, Datos.

**INTRODUCCIÓN**

Para cualquier grupo de investigación la recolección datos es un proceso muy importante1, ya que el análisis, uso o tratamiento de estos datos permitirá comprobar el objetivo de dicha investigación. Por ende, deben tener una metodología o herramientas que permitan la captura de datos de una manera eficiente. Cabe resaltar que no todas las investigaciones requieren de una misma manera para recolectar los datos, las investigaciones sobre el espacio y las investigaciones en microbiología exigen métodos y herramientas distintas de recolección. En lo que a este trabajo compete y dado que Geomatica es un grupo orientado a hacer proyectos con “Trabajo de campo”2, se hablará e investigará sobre herramientas software para la recolección de datos en campo. Respecto a metodologías3 no se indagará mucho dado que para este tipo de proyecto se debe salir a campo a hacer la recolección, ya sea recolectando en hojas de papel, llenando formularios o recolectando con dispositivos portátiles. También actualmente Geomatica utiliza un software instalado en dispositivos portátiles para realizar las recolecciones. Lo que requiere es encontrar una alternativa que les presente más autonomía y control. ArcGIS no es una herramienta libre lo que representa un control muy básico sobre la herramienta y sus componentes, los cuales Geomatica quisiera adaptar aún más a sus necesidades, también el hecho de que el uso de ArcGIS requiere de un costo, el cual Geomatica preferiría usar en desarrolladores que mejoren la herramienta libre que se planea buscar.

**OBJETIVOS**

**OBJETIVO GENERAL:**

* Solventar la necesidad de una metodología estándar para la recolección de datos en las salidas de campo del grupo de investigación Geomatica.

**OBJETIVO ESPECIFICO:**

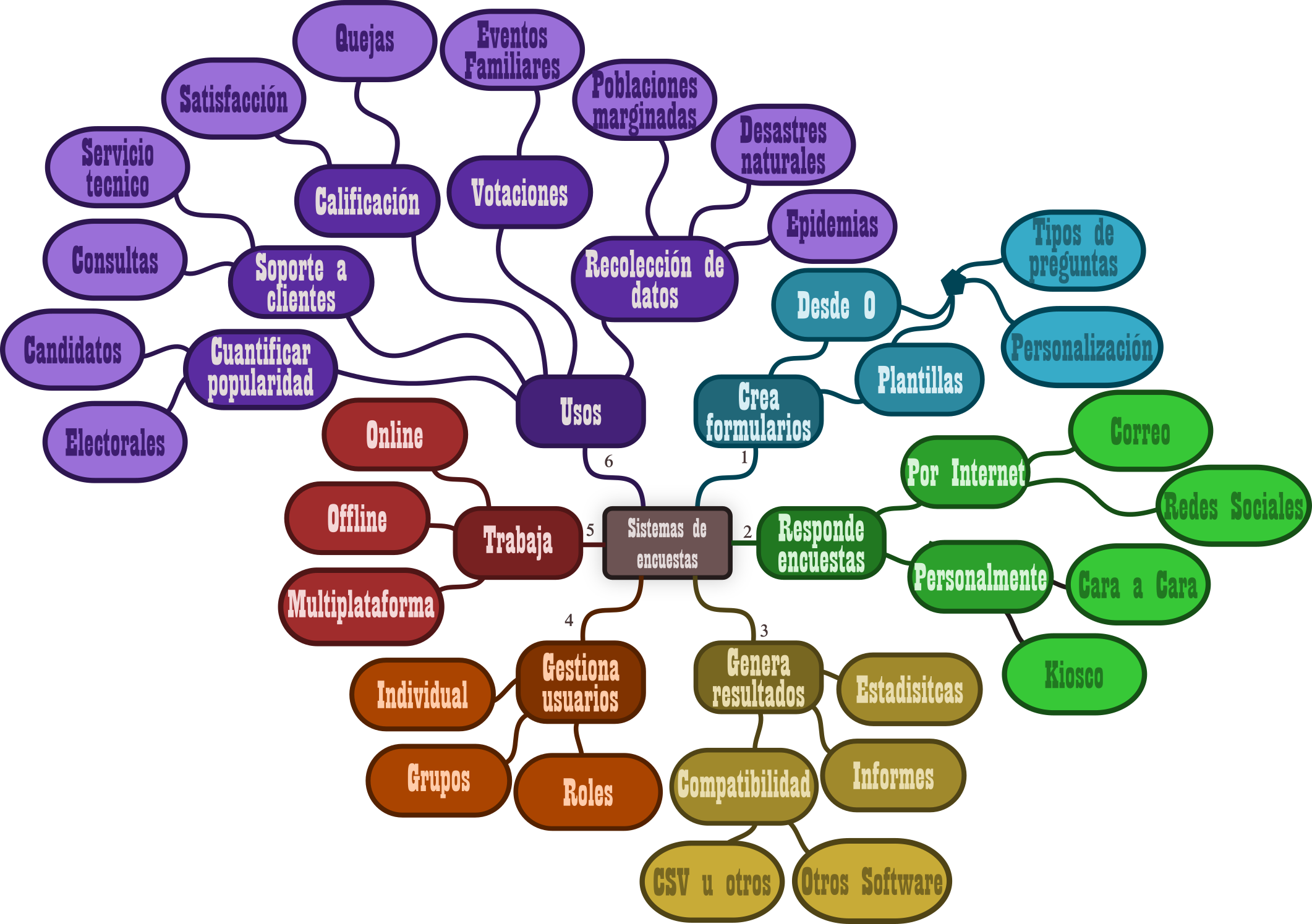
* Realizar un estudio de requerimientos para la problemática presente en Geomatica respecto a la recolección de datos.
* Consultar las distintas herramientas existentes en el mercado para la recolección de datos a gran escala y que cumplan con los requerimientos definidos.
* Seleccionar una de las distintas herramientas consultadas teniendo en cuenta las ventajas y desventajas presentada por cada una de ellas.
* Instalar la herramienta seleccionada en el servidor del grupo de investigación Geomatica.

**DESARROLLO**

A Continuación, se mostrará el proceso que se llevó acabo para completar los objetivos.

**MARCO TEÓRICO**

Como se ha mencionado anteriormente los datos son una parte importante en cualquier tipo de proyecto o investigación, y cada tipo de investigación tiene sus propios métodos y herramientas para esta tarea. Para los proyecto que requieren trabajo de campo, es decir, proyecto o investigaciones que requieren de “acciones encaminadas a obtener en forma directa datos de las fuentes primarias de información, es decir, de las personas y en el lugar y tiempo en que se suscita el conjunto de hechos o acontecimientos de interés para la investigación”1,, es necesario tener herramientas que faciliten el trabajo para el recolector pues en ciertas ocasiones debe recorrer grandes cantidades de terreno y cargar con un montón de papel, como se solía hacer, es muy engorroso. Por fortuna y gracias a los avances de la tecnología, esta recolecta se puede realizar sobre los dispositivos móviles utilizando software especializado.

Existe una gran variedad de software especializado en la recolecta de datos, y como todos los softwares, unos más específicos que otros, unos con más o distintas funcionalidades que otros, unos más conocidos que otros. En general todos estos recolectores de datos presentan un árbol de funcionalidades de la siguiente manera.

**Figura 1.** Mapa general de funcionalidades para un sistema de encuestas o recolector de datos mediante encuestas.

**CONSULTA DEL MERCADO**

Se puede asegurar que existe una gran cantidad de software para la recolección de datos dado que se realizó una búsqueda que así lo demostró. Por mencionar algunas. Kobotoolbox, Portaldeencuestas, Surveymonkey, Eencuestas, Evalmaster, Google form, Quicktapsurvey, Suonper, ONA, Rotatorsurvey, Inty360, Moreapp, Formhub, Smap, Mfieldwork, ODK, Datawinners, ArcGIS Collector y más. La búsqueda se realizó sin ningún criterio de filtrado, todas las herramientas mencionadas poseen algunas de las funcionalidades mostradas en el árbol anteriormente descrito.

Pero no todas tienen las funciones que más requiere Geomatica, y llegados aquí surge una inquietud, ¿Que requiere Geomatica? Mediante distintas reuniones con ingenieros del grupo se obtuvieron los siguientes requisitos.

1. La aplicación permita la recolección de datos de manera offline.
2. La aplicación permitirá la observación de los datos recogidos en tiempo real, para evitar errores en los mismos.
3. permita la toma de posición de los datos y la visualización de los datos en un mapa de los mismos.
4. Por ultimo y la más importante, se debe de poseer control sobre la aplicación. Con control se refiere a poder ingresar al código de la aplicación y cambiarlo si es necesario, para que de esta manera se pueda optimizar el proceso.
5. La aplicación permitirá la observación de los datos recogidos en tiempo real, para evitar errores en los mismos.
6. La aplicación permita la toma de posición de los datos y la visualización de los datos en un mapa de los mismos.

Teniendo en estos criterios y en especial el 4°, se realizó un proceso de selección entre las distintas herramientas consultadas.

**COMPARACIÓN**

Como primer filtro se descartaron aquellas herramientas que no cumplieran con los 3 primeros requisitos, para ser más específicos las que no cumplieran con el primer requisito, ya que, al tratarse de trabajos de campo, la funcionalidad offline es muy importante. Con este primer filtro quedan únicamente las siguientes herramientas. ONA, Rotatorsurvey, Inty360, Moreapp, Formhub, Smap, Mfieldwork, ODK, Datawinners, Kobotoolbox. Luego se prosiguió a hacer la comparación con estas herramientas.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Software | Encuestas | Offline | Análisis | Open Source | Funciones de Pago | Pago | Plataformas | Lenguaje |
| ONA | **✔** | **✔** | **✔** | **BSD** | **✔** | **Mensual** | **WEB, Linux, Mac** | **Python** |
| Rotatorsurvey | **✔** | **✔** | **✔** | **✘** | **✔** | **Anual** | **Windows, Android** | **✘** |
| Inty360 | **✔** | **✔** | **✔** | **✘** | **✔** | **Mensual** | **Windows, Mac, Android, iOS** | **✘** |
| Moreapp | **✔** | **✔** | **✔** | **✘** | **✔** | **Créditos** | **WEB** | **✘** |
| Formhub | **✔** | **✔** | **✔** | **BSD** | **✘** | **✘** | **Linux** | **Python – JS** |
| Smap | **✔** | **✔** | **✔** | **GPL** | **✘** | **✘** | **Linux** | **JS – Java** |
| Mfieldwork | **✔** | **✔** | **✔** | **✘** | **✔** | **Créditos** | **WEB** | **✘** |
| ODK | **✔** | **✔** | **✔** | **Apache** | **✘** | **✘** | **Linux, Mac, Windows** | **Java – JS** |
| Datawinners | **✔** | **✔** | **✔** | **✘** | **✔** | **Anual** | **WEB** | **✘** |
| Kobotoolbox | **✔** | **✔** | **✔** | **AGPL** | **✘** | **✘** | **Linux** | **Python** |

**Tabla 1.** Comparación de las funcionalidades de varias herramientas recolectoras de datos.

Dado que lo que se requiere es un control sobre la herramienta y este control solo puede ser posible si la herramienta es “Libre”4, la primera selección queda de la siguiente manera.

|  |  |  |  |
| --- | --- | --- | --- |
| Repositorio | Lenguaje | GitHub | Ultimo commit |
| [ONA](https://github.com/onaio/onadata) | Python - JS | ✔ | Hace 3 días |
| [Formhub](https://github.com/SEL-Columbia/formhub) | JS - Python | ✔ | 15-abr-15 |
| [Smap](https://github.com/smap-consulting/smapserver) | JS - Java | ✔ | Hace 4 días |
| [ODK](https://github.com/opendatakit/collect) | Java - JS | ✔ | Hace pocas Horas |
| [Kobotoolbox](https://github.com/kobotoolbox) | Python - JS | ✔ | Hace pocas horas |

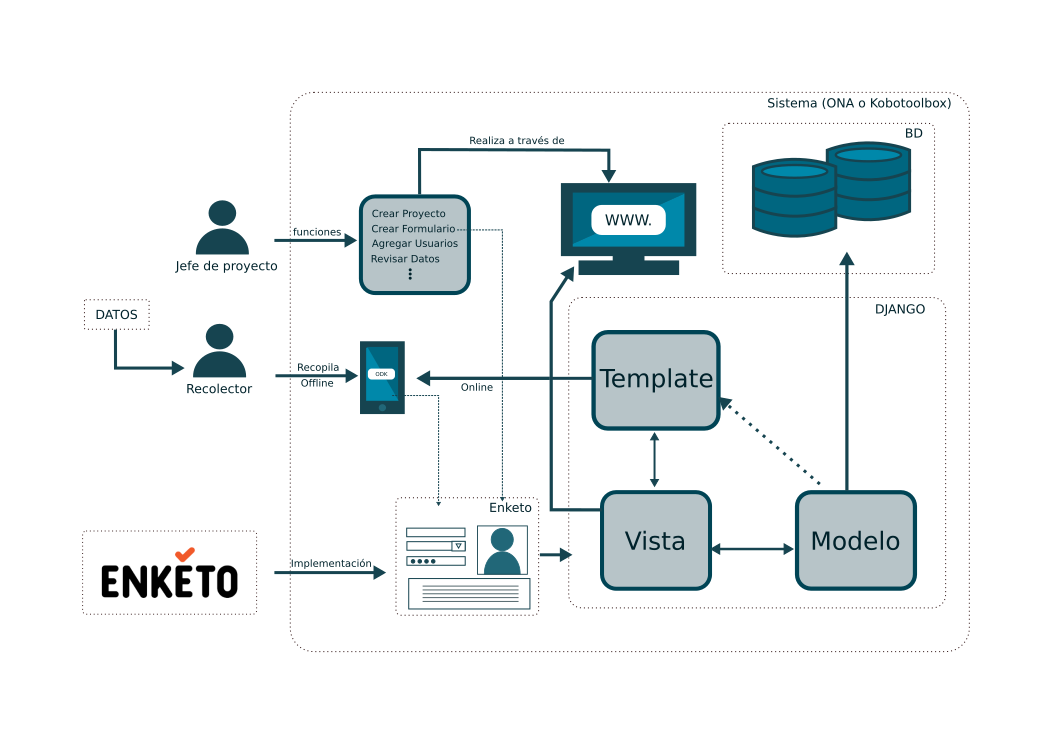
**Tabla 2.** Lista de herramientas recolectoras de datos libres. **Hoy:** 17 – Marzo - 2017

En este punto se realizó un filtro basado en la vigencia del software y en su lenguaje. Primero dado que un proyecto vigente presenta mayor compatibilidad con las nuevas tecnologías que uno abandonado hace varios años. Como lenguaje de la herramienta se escogió Python, dado que es un lenguaje que ha tomado gran importancia en el mundo y por qué varios de los ingenieros de Geomatica prefieren este lenguaje. De esta manera quedan dos herramientas a escoger ONA o Kobotoolbox. Cabe resaltar ambas herramientas cumplen con los 6 requisitos que solicita Geomatica.

**SELECCIÓN DE HERRAMIENTA**

El plan que seguir será escoger alguna de ellas para su instalación. Para ello se realizó una consulta minuciosa en cada uno de los repositorios de dichas plataformas, ambas disponibles en GitHub, se revisaron los contribuidores, la licencia, los Issues, el método de instalación, las dependencias, la comunidad, los commits, entre otras cosas. De esto se concluye que ambos proyectos están muy relacionados entre sí y esto hace difícil la decisión a tomar. Ambas están siendo desarrollados muy de la mano, pues se encontraron conexiones y relaciones entre ellas, por ejemplo, el recolector de datos de Kobotoolbox está basado en onadata, el cual es el recolector de datos de ONA, ya que ONA es la empresa detrás de onadata. En los repositorios de ONA existen Forks de los repositorios de Kobotoolbox lo que podría significar algún tipo de contribución de ONA a Kobotoolbox, también ambos proyectos comparten contribuidores, Respecto a Issues y commits la cosa no ayuda pues en ciertos aspectos onadata presenta ventaja, mientras que en las otras Kobotoolbox es quien gana, incluso ambas herramientas funcionan en una instalación típica o en Dockers. Cabe resaltar, nuevamente, que esta información fue tomada de lo que mostraban los repositorios de GitHub de ambos proyectos.

Se debe tomar una decisión, onadata o kobotoolbox, y la única manera de decidir es comparar la estructura de ambas herramientas y poner a prueba su funcionamiento. Al estar desarrolladas en Python, más específicamente en Django, ambas cuentan con la siguiente estructura.



**Figura 2.** Estructura del funcionamiento de onadata y Kobotoolbox.

Es necesario hacer énfasis en una de las tantas similitudes que presentan onadata y kobotoolbox y es el uso de Enketo,ya que esta librería presenta una gran variedad de opciones y robustez a la hora de hacer toma de datos offline.

**INSTALACION**

La estructura es similar, queda comprar su funcionamiento. Primer factor que comprar es su facilidad de instalación, y la documentación respecto. Onadata presenta un proceso de instalación complicado e incluso presenta una documentación incompleta y desactualizada para dicho proceso, a tal punto que existen varias guías de terceros con pequeñas modificaciones para cada caso, e incluso se realizaron pequeñas modificaciones no documentadas en el proceso de instalación. Para la instalación de onadata se tuvieron en cuenta las recomendaciones de las distintas guías disponibles5. En caso contrario Kobotoolbox6 es muy sencilla de instalar un par de líneas en la terminal, una buena conexión a Internet, un par de minutos o una hora y listo, Kobotoolbox instalado. Las instalaciones anteriormente mencionadas se realizaron en local, ya que al servidor solo se montará una de ellas.

**FUNCIONAMIENTO**

Comparando su funcionamiento y experiencia de usuarios Kobotoolbox es la mejor herramientas, el funcionamiento de onadata se puede probar creando una cuenta en la página web [https://odk.ona.io](https://odk.ona.io/) y el funcionamiento de Kobotoolbox se puede comprobar creando una cuneta en la pagina web [https://kf.kobotoolbox.org](https://kf.kobotoolbox.org/). Las cuentas creadas en las plataformas web de ambas entidades son cuentas limitadas. Onadata Classic (que es la herramienta instalada, pues existe una opción con muchas más funcionalidades disponible en la web [https://ona.io/](https://ona.io/home/), pero no es útil para Geomatica dado que tendrían que incurrir a gastos, no de licencia pero si de servicios, ya que la cuenta básica es muy limitada) presenta una gran desventaja frente a Kobotoolbox puesto que no deja de ser un CRUD para encuestas XLSFORM hechas a mano en Excel o cualquier hoja de cálculo. La estructura mencionada anteriormente la cumple onadata (ona.io), pero no la cumple onadata classic (odk.ona.io). En cambio, kobotoolbox si presenta las mismas funcionalidades tanto la versión que se instaló como la versión en la plataforma oficial, a excepción a un cambio en la GUI.

Por estas razones se elige instalar Kobotoolbox en los servidores de Geomatica. Y aquí surgen nuevos inconvenientes, Kobotoolbox presenta tres maneras para ser instalada en un servidor. Mediante Dockers, Manual, o usando una máquina virtual con vagrant(Usada para las pruebas). La instalación de Dockers y la Manual se descartan debido a su complejidad, dado que para ambas se requiere de conocimientos profesionales en servidores y en Linux. La instalación mediante Vagrant y máquina virtual es las más sencilla de realizar, pero presenta un inconveniente y es que kobotoolbox ya no le da soporte a esta metodología. Y la documentación y foros sobre esta metodología en la actualidad son nulas.

**COMPARACION CON ARCGIS**

Igual que con todas las demás herramientas ArcGIS cuenta con una plataforma Web donde se pueden observar una breve explicación de sus funcionalidades. Para empezar ArcGIS es todo un “ecosistema” de aplicaciones, Apis, funcionalidades, cursos, personas, etc. que comparados con Kobotoolbox, se queda pequeño. Para empezar ArcGIS cuenta con una mayor cantidad de personas desarrollando y dando soporte, y esto se puede ver en el GitHub de ArcGIS comprado con el GitHub de Kobotoolbox. Como segundo punto de comparación. Kobotoolbox y ArcGIS tienen un enfoque diferente. Kobotoolbox es una herramienta para la recolección de datos y ArcGIS es una herramienta cartográfica, cabe resaltar que en el mundo de ArcGIS existe también una herramienta recolectora de datos, la cual es libre de usar, ArcGIS Collector. En otras palabras, ArcGIS no cobra por la recolección de los datos en sí, si no por el tratamiento o uso de estos datos en su plataforma y con las distintas funcionalidades sobre mapas que presenta, pues como herramienta cartográfica presenta una amplia cantidad de funciones sobre mapas. ArcGIS en el mercado es comparado con Google Maps, ya que con ambas herramientas se pueden crear distintos tipos de mapas.

**INSTALACION EN SERVIDOR**

\*\* EN PROCESO con problemas\*\*

**CONCLUCIONES**

* Se realizaron varias entrevistas con ingenieros del grupo logrando una lista completa de requerimientos con los cuales fue posible escoger la herramienta más óptima para Geomatica.
* Se realizo la investigación y arrojo una lista de distintas alternativas para el software ArcGIS.
* Se decide usar Kobotoolbox debido a su ventaja sobre las demás herramientas consultadas.
* \*\*EN PROCESO con problemas\*\*

**REFERENCIAS**

1. SOTO-LESMES Virginia Inés, DURAN María Mercedes de Villalobos. El trabajo de campo clave en la investigación cualitativa. AQUICHAN VOL. 10 No 3 - CHÍA, COLOMBIA - DICIEMBRE 2010 p. 253-266.
2. OMONTE Rivero Abraham. Ciencias Sociales y administrativas. Investigación académica. [en línea] <<http://www.mailxmail.com/curso-ciencias-sociales-investigacion-administrativas-academica/trabajo-campo-concepto-diseno>> [citado: 4-Nov-17]
3. UNIVERSIDAD nacional autónoma de México. Aspectos teóricos de la medición y métodos de recolección de datos. [en línea] <<http://www.psicol.unam.mx/Investigacion2/pdf/METO6F.pdf>> [citado: 4-Nov-17]
4. Free Software Foundation. ¿Qué es el software libre? [en línea] <<https://www.gnu.org/philosophy/free-sw.html>> [citado: 4-Nov-17]
5. Guais de Instalación onadata. <<https://github.com/badili/classic_ona/blob/master/installation-configuration.md>>, <<https://gist.github.com/oguya/8c7b81634d648500a8c00400abab327e>>, <<https://github.com/bibbox/app-azizi-amp>>, <<https://api.ona.io/static/docs/install.html>> [en línea] [citados: 04-Nov-17]
6. Guías de instalación Kobotoolbox [en línea] [citado: 4-Nov-17] <<http://support.kobotoolbox.org/customer/portal/articles/1691105-using-kobotoolbox-offline>>.