



Herramienta para la construcción de cuestionarios DDI para encuestas y censos

PROYECTO DE GRADO – FACULTAD DE INGENIERÍA

FEDERICA FERREIRA | KARINA MATTEI
TUTORES: ADRIANA MAROTTA | SEBASTIÁN GARCÍA

RESUMEN

Al momento de interpretar los datos obtenidos a partir de la realización de un censo o encuesta, es necesario contar con información extra sobre estos datos, para de esta forma lograr una mejor interpretación de ellos.

Por tal motivo, cuando se define el cuestionario a realizar, se debe guardar información que nos permita luego mejorar la calidad de las conclusiones que derivan de los datos. Es por ello que resulta importante almacenar el flujo del cuestionario, las variables definidas, las preguntas, el dominio y las posibles respuestas de las mismas, etc.

Para este propósito existen varios estándares definidos y utilizados por organizaciones estadísticas que definen cómo se almacena toda esa información relacionada con el cuestionario. Respetar un estándar aporta a la interoperabilidad de sistemas y colabora a un mejor entendimiento de la información recolectada. El presente proyecto centra su análisis en el estándar DDI [1], en sus dos versiones Codebook y Lifecycle.

El objetivo del proyecto consiste en implementar una herramienta para la construcción de cuestionarios bajo el estándar DDI.

Una de las motivaciones del proyecto fue que al comenzar la investigación del estándar DDI, no existían herramientas libres para poder definir el flujo del cuestionario, ni herramientas que manejaran las dos versiones de DDI, ni mucho menos la migración de una versión a otra.

Los principales resultados obtenidos se pueden resumir en los siguientes puntos:

- Herramienta intuitiva y de fácil manejo que permite la definición de preguntas, variables, dominios y lo más importante, permite definir el flujo de un cuestionario.
- Dicha herramienta permite además el trabajo con ambas versiones del estándar, teniendo la posibilidad de migrar archivos de una versión a otra.
- El análisis y estudio de los metadatos asociados a los cuestionarios realizados en este proyecto, sirve como punto de partida para la verificación de la calidad de los datos obtenidos con el cuestionario.

ÍNDICE

Resumen	1
Índice	2
1 Introducción	5
1.1 Objetivos	5
1.2 Organización del informe	6
2 Conceptos básicos	7
2.1 Microdatos	7
2.2 Metadatos	7
2.2.1 Organizaciones y Metadatos	8
2.3 Estándares	11
2.3.1 DDI	11
2.3.2 SDMX	12
2.3.3 Diferencias entre SDMX y DDI	13
2.3.4 Dublin Core	13
2.4 Técnicas de recolección de datos	14
3 DDI	16
3.1 DDI-Codebook	17
3.1.1 Grupos de atributos	17
3.1.2 Descripción del documento <docDscr>	19
3.1.3 Descripción del estudio	20
3.1.4 Descripción del archivo <fileDscr>	21
3.1.5 Descripción de variables <dataDscr>	22
3.1.6 Otro material relacionado con el estudio <otherMat>	23
3.2 DDI-Lifecycle	23
3.2.1 Uso Simple vs Complejo	24

3.2.2	Relación con otros estándares.....	25
3.2.3	Etapas del ciclo de vida	26
3.2.4	Identificación, versionado, mantenimiento y referencia.....	29
3.2.5	Instancias DDI	30
3.2.6	Instrumento	30
3.3	DDI-Codebook vs DDI-Lifecycle	32
4	Herramientas relacionadas	33
4.1	Colectica	33
4.2	Blaise	34
4.3	Nesstar	35
4.3.1	Nesstar Publisher	35
4.3.2	Nesstar Server	36
4.3.3	Nesstar WebView	36
5	Análisis	37
5.1	Requerimientos.....	38
6	Diseño	40
6.1	Funcionalidades de la aplicación.....	40
6.2	Estructura de la aplicación	40
6.3	Interfaz	41
6.4	Casos de Uso	44
7	Implementación	45
7.1	Herramientas utilizadas	45
7.2	Patrón MVVM	45
7.3	Módulos de la aplicación.....	46
7.4	Principales funcionalidades.....	47
7.4.1	Importación/Exportación.....	47
7.4.2	Conversión de DDI2 a DDI3.....	47

7.4.3	Creación de instrumento	48
7.4.4	Exportación a RTF	48
7.5	Manual de Usuario	49
8	Ejemplos.....	50
8.1	CodeBook	50
8.2	DDI 3.....	50
9	Validación Instituto Nacional de Estadística	52
10	Desarrollo del proyecto	54
10.1	Cronogramas	54
10.2	Etapas.....	54
11	Conclusiones	57
12	Trabajos futuros.....	58
12.1	Repositorio	58
12.2	Exportación a HTML	58
12.3	Aplicación WEB	58
12.4	Validación Microdatos	58
12.5	Publicación de API.....	58
13	Bibliografía	60
Anexo 1 – Casos de uso		62
Anexo 2 – Manual de usuario		71

1 INTRODUCCIÓN

Luego de realizada una encuesta y/o censo se obtiene un conjunto de datos, a partir del cual se sacan conclusiones, resúmenes, derivaciones, publicaciones, etc.

Cada encuesta tiene asociado un cuestionario, se trata del orden de las preguntas y el cómo es preguntada la encuesta, cada cuestionario además tiene asociada una documentación que hace posible la interpretación correcta de los resultados obtenidos.

Los datos que no contienen la información detallada de cómo éstos fueron relevados y procesados, puede llevar a una interpretación errónea y conclusiones equivocadas. Por este motivo existen estándares internacionales que permiten utilizar una terminología común para la interpretación y descripción de datos.

Uno de los estándares elegido por varios organismos como el Instituto Nacional de Estadística de Uruguay y U.S. Census Bureau, es DDI [1]. Dicho estándar tiene dos líneas de desarrollo, una versión simple, DDI – Codebook y una versión extendida, DDI – Lifecycle diseñada para documentar y manejar los datos en todo su ciclo de vida.

Actualmente son escasas las herramientas que trabajan con la generación de metadatos en el estándar DDI en cualquiera de sus dos versiones. Esta es entonces una de las principales motivaciones del proyecto, crear una herramienta que facilite la creación de metadatos de acuerdo al estándar DDI, para los cuestionarios de censos y/o encuestas.

La generación de metadatos asociados a los cuestionarios es muy importante para poder validar información recolectada a través de los mismos. Esos metadatos permiten una mejor interpretación de los microdatos de encuestas y censos, así como también la verificación de la calidad de los mismos. Existe en este marco una tesis de maestría en curso "Calidad de datos aplicada a encuestas y censos" la cual se centra en la verificación de la calidad de los microdatos a partir de la metadata correspondiente. El análisis y estudio de los metadatos asociados a los cuestionarios realizados en este proyecto, sirve como punto de partida para mencionado estudio.

1.1 OBJETIVOS

El objetivo del proyecto consiste en construir una herramienta que permita la creación y la manipulación de cuestionarios con sus respectivos metadatos de acuerdo al estándar DDI.

Los objetivos específicos son los siguientes:

- Estudio del estándar DDI en sus dos versiones (DDI – Codebook, DDI – LifeCycle)
- Análisis de las diferencias entre ambas versiones del estándar.
- Implementación de una herramienta que permita la creación de metadatos asociados a cuestionarios.

1.2 ORGANIZACIÓN DEL INFORME

El informe está dividido en trece capítulos.

Luego de la introducción, en el segundo capítulo se brinda una visión general sobre conceptos recurrentes a lo largo del proyecto. Tiene como objetivo contextualizar el trabajo y servir de base para las siguientes secciones, en donde el enfoque deja de ser global para ser más detallado.

En el tercer y cuarto capítulo se detalla el estándar DDI, se describen sus versiones y se mencionan herramientas relacionadas con dicho estándar.

En el quinto y sexto capítulo se detalla la especificación de los requerimientos funcionales, el diseño conceptual, el diseño lógico y la visión global del sistema.

El séptimo y octavo capítulo describen la implementación del sistema. Se mencionan las herramientas utilizadas y la explicación a alto nivel de implementación de cada una de las funcionalidades del sistema. Además se presentan dos ejemplos de archivos utilizados para probar la herramienta.

En el noveno capítulo se presenta la validación realizada por el INE a través de dos encuentros donde se mostró el sistema implementado.

En el capítulo once, se presenta como fue desarrollado el proyecto, describiendo sus etapas y detallando el cronograma.

Por último, en los capítulos doce y trece se incluyen las conclusiones y se presentan lineamientos para trabajos futuros.

Al final del presente informe se incluye una sección de Anexos conteniendo el detalle de los casos de uso del sistema y el manual de usuario.

2 CONCEPTOS BÁSICOS

Actualmente existen infinidad de herramientas que publican datos y otras tantas herramientas y/o personas que consumen y utilizan dichos datos. De esa forma se genera un gran intercambio de información que nos lleva a cuestionarnos si los datos están siendo bien interpretados, o si tal vez se necesite algún tipo de información extra para saber interpretar los datos que se encuentran disponibles.

En este capítulo se presentan conceptos básicos y distintos estándares utilizados para interpretar y documentar datos.

2.1 MICRODATOS

Luego de una encuesta o censo, podemos decir que el conjunto de datos obtenido se puede separar en dos grupos de acuerdo a su estructura y a su procesamiento:

- **Datos agregados:** Aquellos con algún tipo de procesamiento (contabilizados, ordenados, etc.)
- **Microdatos:** Datos sin procesar.

Los microdatos son la información recabada a partir de un determinado individuo.

Dada la cantidad de datos que se pueden encontrar en un archivo de microdatos, los resultados de censos y/o encuestas son publicados generalmente como datos agregados. [1]

Sin embargo, al utilizar los microdatos como punto de partida para presentar datos agregados, generalmente ocurre una pérdida de información. Por ejemplo, si las estadísticas de educación y empleo se agregan por separado, no pueden ser utilizadas para explorar la relación entre estas dos variables. El acceso a los microdatos permite a los investigadores mucho más libertad para investigar estas interacciones y realizar un análisis detallado. [2]

Los usuarios pueden generar sus propias estadísticas de los datos de cualquier manera que se desee.

2.2 METADATOS

Para una correcta interpretación, resulta necesario que los microdatos se encuentren acompañados de una documentación que nos ayude a interpretarlos. Son los *metadatos* quienes nos aportan esa información extra.

Podríamos definir un metadato como información sobre las características de los datos, que permiten entender lo que representan para que puedan ser compartidos y explotados de manera eficaz por todo tipo de usuarios a lo largo del tiempo.

METADATO: el término fue acuñado por Jack Myers en la década de los 60 para describir conjuntos de datos. La primera acepción que se le dio (y actualmente la más extendida) fue la de dato sobre el dato, ya que proporcionaban la información mínima necesaria para identificar un recurso. En este mismo trabajo se afirma que puede incluir información descriptiva sobre el contexto, calidad y condición o características del dato. [3]

Los datos sin un contexto pueden llegar a ser subutilizados o mal interpretados, lo que puede acarrear el riesgo de obtenerse conclusiones no del todo precisas; por lo tanto, la documentación de éstos tiene gran importancia para todos los usuarios de los datos, además de ser la base para la creación de una cultura estadística.

Los metadatos ayudan al usuario a transformar los datos estadísticos en información, de ahí la importancia de la adopción de una terminología común y un conjunto de elementos estándar, documentados en relación a normas bien establecidas y de aceptación internacional.

De un tiempo a esta parte se ha comenzado a buscar estrategias para mejorar el intercambio de datos estadísticos, son lineamientos para mejorar el mecanismo utilizado para la comparación entre las naciones, pudiendo ser usados para armonizar los procesos de generación, compilación y diseminación de información estadística. Son desarrollados por acuerdo de expertos técnicos de diversos países bajo cooperación internacional y se constituyen en recomendaciones internacionales por su carácter no vinculante.

En resumen, los metadatos proporcionan a los usuarios información sobre los datos existentes, sus procesos de recolección, procesamiento y evaluación que se utilizan en su generación, así como las formas de acceder a ellos, con el fin de que cada usuario pueda identificar, localizar y consultar según sus necesidades de información. Por lo tanto, los metadatos ayudan al usuario a transformar los datos estadísticos en información relacionada con el proceso estadístico.

2.2.1 ORGANIZACIONES Y METADATOS

Existen organizaciones a nivel nacional e internacional que publican y/o consumen metadatos, y además trabajan en conjunto para que la construcción de metadatos sea un proceso más homogéneo.

A nivel internacional presentamos PARIS21 e IHNS, a nivel nacional se presenta el Instituto Nacional de Estadística.

2.2.1.1 PARIS21

PARIS21 (Partnership in Statistics for Development in the 21st Century) es una iniciativa única que tiene como objetivo promover el mejor uso y la producción de estadísticas en el mundo en desarrollo. Desde su creación en 1999, PARIS 21 ha desarrollado con éxito una red mundial de estadísticos, políticos, analistas y profesionales del

desarrollo comprometidos con la toma de decisiones basadas en la evidencia. El objetivo principal de PARIS21 es alcanzar los objetivos nacionales e internacionales de desarrollo y reducir la pobreza en los países de ingresos bajos y medios. PARIS21 intenta facilitar el desarrollo de la capacidad estadística, aboga por la integración de datos fiables en la toma de decisiones además de coordinar las fuentes de datos estadísticos.

El objetivo de PARIS21 es desarrollar una cultura de la gestión para Resultados de Desarrollo. PARÍS 21 persigue este objetivo principalmente alentando y ayudando a países de ingresos bajos y medios a diseñar, implementar y monitorear una estrategia nacional de desarrollo estadístico.

PARIS 21 fue establecido por las Naciones Unidas, la Comisión Europea, la Organización para la Cooperación y el Desarrollo Económico (OCDE), el Fondo Monetario Internacional y el Banco Mundial.

Actualmente un veintidós por ciento de la población mundial vive en extrema pobreza. Aunque esta cifra ha disminuido en los últimos años, más esfuerzos necesitan ser desplegados con el fin de alcanzar los objetivos internacionales de desarrollo.

En este contexto, los estadísticos tienen un enorme papel. Según PARIS 21 la reducción de la pobreza no puede lograrse sin la mejora de las políticas gubernamentales en los países en desarrollo. Y las buenas políticas gubernamentales necesitan buenas estadísticas para alcanzar sus metas. Las estadísticas son esenciales para controlar el progreso y alcanzar los objetivos de desarrollo. Las estadísticas son cruciales para proporcionar un ambiente de toma de decisiones basada en la evidencia y para hacer que los gobiernos rindan cuentas de sus actividades.

La visión de PARIS 21 radica en reducir la pobreza y mejorar la gobernanza en los países en desarrollo mediante la promoción de la integración de las estadísticas y datos fiables en el proceso de toma de decisiones.

PARIS21 está dedicada a fomentar un mejor uso de las estadísticas en los países en desarrollo, mediante el apoyo y el fortalecimiento de sus sistemas nacionales de estadística. [4]

2.2.1.2 IHSN

IHSN (International Household Survey Network) es una red de organismos internacionales creada para promover la disponibilidad, la fiabilidad y la accesibilidad de los datos de encuestas.

La misión de la IHSN es mejorar la disponibilidad, accesibilidad y calidad de los datos de las encuestas en los países en desarrollo, además de fomentar el análisis y uso de estos datos por los tomadores de decisiones nacionales e internacionales de desarrollo, la comunidad científica y otras partes interesadas.

Los objetivos de IHSN son:

- Mejorar la coordinación de los programas de encuestas con patrocinio internacional, con énfasis sobre el calendario, la secuenciación, la frecuencia y coste-efectividad.
- Desarrollar metodologías coordinadas para todas las fases del ciclo de vida de la encuesta.
- Disponibilizar un catálogo de datos central con información de encuestas y censos, como herramienta para los usuarios de datos.
- Desarrollar estándares, herramientas y directrices que permitan a los productores de datos documentarlos, difundirlos de acuerdo a los estándares internacionales y a las mejores prácticas.
- Mejorar la colaboración entre los productores y los usuarios de los datos.

Estas son algunas de las organizaciones que conforman IHSN:

- PARIS21
- Instituto de Estadística de la UNESCO (UIS)
- UNICEF
- United Nations Statistics Division (UNSD)
- Organización Mundial de la Salud: OMS
- Banco Mundial

De acuerdo a sus objetivos, IHSN promueve algunas herramientas que merecen una breve presentación por su importancia en el proceso estadístico y por su utilización a nivel mundial:

Archivo Nacional de Datos, **ANDa** (NADA para su sigla en inglés), es un catálogo de metadatos en un portal web. Permite buscar, consultar y comparar información relevante a censos y/o encuestas. La aplicación es utilizada por un gran número de organizaciones. ANDa utiliza el estándar DDI en su versión 2.x. [5]

Editor de metadatos, también conocido como **Nesstar Publisher**, es un editor para la preparación de metadatos y datos, como también para la publicación de los mismos en un catálogo como por ejemplo el Archivo Nacional de Datos (ANDa). Este editor maneja metadatos compatibles con la versión 2.x del estándar DDI y con el estándar Dublin Core. La aplicación fue desarrollada por Nesstar y está catalogada como un software libre. [6]

Aprovechando las herramientas promovidas por la IHSN, se desarrolla también el Programa Acelerado de Datos (**PAD**), cuyo objetivo es aumentar el uso y el valor de los datos de la encuesta. El programa apoya a los productores y usuarios de datos en los países en desarrollo, ayudándoles a hacer un mejor uso de los datos existentes y mejorando la alineación de los programas de encuestas como también la producción estadística de datos. [7]

2.2.1.3 INSTITUTO NACIONAL DE ESTADÍSTICA

Instituto Nacional de Estadística del Uruguay (**INE**), tiene por objetivo la elaboración, supervisión y coordinación de las estadísticas nacionales.

El INE se inscribe en la iniciativa de Documentación de Datos (Data Documentation Initiative - DDI) a través del Programa Acelerado de Datos (PAD) de la Red Internacional de Encuestas de Hogares (IHSN) buscando así el fortalecimiento en la generación y documentación de operaciones estadísticas. [8]

2.3 ESTÁNDARES

En esta sección son presentadas algunas iniciativas de estándares.

Al momento de documentar datos a publicar, cada herramienta podría hacerlo de la forma que considere más adecuada, pero suena lógico tener un estándar para hacerlo, por ello, muchas herramientas e instituciones documentan la información publicada de acuerdo a estándares internacionales.

2.3.1 DDI

DDI (Documentation Data Initiative) promovido por el Banco Mundial (BM) y la Organización para la Cooperación y el Desarrollo Económico (OCDE), es un esfuerzo por establecer un estándar internacional basado en archivos de formato XML para la documentación, presentación, transferencia, diseminación y preservación de datos. También permite documentar una a una las variables de los proyectos estadísticos y vincular esta información con la base de datos.

La **DDI Alliance** comparte un compromiso en el cumplimiento de los requisitos mundiales de normas de acceso público y de los productos semánticos, brindando documentación e integración de datos de ciencias sociales y otros datos para la comprensión de la condición humana.

Los objetivos de DDI Alliance son:

- Mantener el continuo desarrollo del estándar DDI y de productos semánticos.
- Promover la utilización del estándar.
- Apoyar el desarrollo de programas de formación que fomenten el uso de estos estándares para todos los niveles de habilidad de los posibles adoptantes.
- En la medida de lo posible, garantizar la compatibilidad de las normas de DDI con nuevos estándares de metadatos en otros campos.
- Equilibrar los intereses de una comunidad diversa de partes interesadas a través de un proceso abierto, transparente e impulsado por el consenso.

DDI tiene dos grandes líneas de desarrollo: el estándar original DDI (versiones 1,0 a 2,5, ahora llamados "DDI - Codebook") y DDI basado en el ciclo de vida (versiones 3.0 a 3.2, ahora llamados "DDI - Lifecycle"). Ambas líneas son capaces de describir conjuntos de microdatos con una gran cantidad de metadatos relacionados (preguntas, variables, conceptos, categorías, códigos, etc.).

Un archivo de datos es en última instancia, sólo una cadena de caracteres no comprensible por sí misma. Sólo puede ser interpretado a través del uso de la correspondiente documentación, en ella se indica la ubicación de una variable en el archivo de datos, la relación entre las variables, todas las respuestas posibles a una pregunta, cómo se tomaron muestras de la población de interés, etc. El archivo de datos (microdatos) y su documentación (metadatos) constituyen la información de contenido.

2.3.2 SDMX

El estándar SDMX (Statistical Data and Metadata eXchange) fue lanzado en 2001 por siete organizaciones que trabajan con datos estadísticos a nivel internacional: el Banco de Pagos Internacionales (BIS), el Banco Central Europeo (BCE), Eurostat, el Fondo Monetario Internacional (FMI), la Organización para la Cooperación y el Desarrollo Económico (OECD), la División de Estadística de Naciones Unidas (UNSD) y el Banco Mundial. Estas siete organizaciones actúan como los patrocinadores de la iniciativa SDMX.

El objetivo declarado de SDMX fue el desarrollo y uso más eficientes de los procesos de intercambio y disseminación de datos y metadatos estadísticos entre los organismos internacionales y sus países miembros. Para lograr este objetivo, SDMX ofrece un modelo de información para generar formatos estándar para los datos y metadatos, junto con las directrices de contenido y una arquitectura de tecnología de información, para el intercambio de datos y metadatos. Las organizaciones son libres de hacer uso de cualquiera de los elementos que sean más apropiados para cada caso.

Con Internet y la "World Wide Web", el intercambio electrónico y el intercambio de datos se han convertido en algo más fácil y más común, pero el intercambio a menudo ha tenido lugar utilizando todo tipo de formatos y conceptos no estandarizados. Esto crea la necesidad de normas y directrices comunes que permiten procesos más eficientes para el intercambio y difusión de datos y metadatos estadísticos. Por ello, SDMX busca la armonización conceptual estadística que facilite la integración y la comparación de datos provenientes de diversas fuentes. SDMX garantiza que los metadatos siempre vayan junto con los datos, haciendo que la información sea inmediatamente comprensible y utilizable.

Así mismo, la evolución de SDMX y su implementación a través de Servicios Web, está dando lugar a que la disseminación de datos estadísticos a usuarios finales sea también a través de SDMX, promoviendo el desarrollo de herramientas muy llamativas de visualización, gráficas y representación en mapas.

Sus principales elementos son:

- Modelo de información SDMX: Datos estadísticos (dimensiones, atributos y medidas) y Metadatos estadísticos (metadatos estructurales y metadatos de referencia).
- Lineamientos y pautas orientados al contenido: Conceptos de dominio cruzado (transversales), dominios estadísticos temáticos vocabulario común de metadatos.
- La arquitectura IT para el intercambio de datos: Formatos estándar para el intercambio de datos y metadatos, arquitecturas para el intercambio de datos.

2.3.3 DIFERENCIAS ENTRE SDMX Y DDI

SDMX y DDI han sido intencionalmente diseñados para alinearse entre sí, así como con otras normas de metadatos. Aunque existe un cierto solapamiento en su capacidad descriptiva, pueden caracterizarse como estándares complementarios.

Uno de los puntos de similitud entre SDMX y la versión 3.0 de DDI es la existencia en cada caso de un modelo conceptual, que forma la base de la implementación de XML. En SDMX, existe el modelo de información SDMX, que es un meta-modelo del cambio y de los procesos de difusión en torno a las estadísticas globales. En DDI, el modelo es un modelo de metadatos más específico, una visión particular del ciclo de vida de los datos. Estos modelos tienen muchas similitudes, en parte debido a que ambos están alineados con la norma ISO / IEC 1117910 (un estándar ampliamente aceptado en relación con la semántica y los registros de metadatos).

A modo de resumen, SDMX es similar a DDI, pero las dos normas son en realidad muy diferentes en su alcance. Considerando que los documentos DDI basan su investigación a través de los microdatos y el ciclo de vida de los datos agregados, SDMX se ocupa de la eficiencia de todo el intercambio de datos agregados. SDMX es utilizado principalmente por la comunidad de las estadísticas oficiales para el intercambio de datos de series de tiempo. El hecho de que estas dos normas están bien alineados significa que se pueden combinar en formas potentes. Por otra parte, los usuarios de los dos estándares pueden mover datos de un formato estándar a la otra con bastante facilidad.

2.3.4 DUBLIN CORE

El estándar DCMI (Data Core Metadata Initiative), también conocido Dublin Core, consta de un simple conjunto de elementos para la descripción de recursos digitales. Este estándar es particularmente útil en la descripción de los recursos relacionados a los microdatos, como cuestionarios, informes, manuales, guiones, programas informáticos, etc. Fue creado en 1995 por OCLC (Online Computer Library Center) y NCSA (National Center for Supercomputing Applications) en un taller en Dublin, Ohio. Con los años, se ha convertido en el estándar más utilizado para la descripción de recursos digitales en la Web y fue aprobado como un estándar ISO en 2003. Es mantenido y

desarrollado aún más por la DCMI, una organización internacional dedicada a la promoción estándares de metadatos.

Una de las principales razones detrás del éxito de Dublin Core es su simplicidad. Desde el principio, ha sido el objetivo de los diseñadores para mantener el conjunto de elementos tan pequeño y simple como sea posible para permitir que el estándar sea utilizado por personal no especializado. La norma también hace que sea fácil y barato crear registros descriptivos simples para recursos de información, al tiempo que proporciona para la recuperación eficaz de esos recursos en la Web o en cualquier entorno de red similar.

En su forma más simple el Dublin Core se compone de los siguientes quince elementos de metadatos, todos los cuales son opcionales y repetibles: Title, Relation, Rights, Subject, Coverage, Date, Description, Creator, Format, Type, Publisher, Identifier, Source, Contributor, Language.

La definición del estándar Dublin Core se puede resumir como un elemento simple pero efectivo para describir una amplia gama de recursos en red. Otra forma de ver Dublin Core es como un "pequeño lenguaje para hacer afirmaciones sobre los recursos". En este lenguaje, hay dos clases de términos, los elementos (sustantivos) y los calificadores (adjetivos), ambos pueden combinarse en un modelo simple de declaraciones. En el diverso mundo de la Internet, el Dublin Core puede ser visto como un "traductor de metadatos para los turistas digitales": fácil de entender, pero no necesariamente expresando relaciones o conceptos complejos.

2.4 TÉCNICAS DE RECOLECCIÓN DE DATOS

Una encuesta es un proceso de investigación, en el cual se busca recolectar datos por medio de un cuestionario previamente diseñado. Los datos se obtienen realizando un conjunto de preguntas dirigidas a una muestra representativa o al conjunto total de la población estadística en estudio, integrada a menudo por personas, empresas o entes institucionales, con el fin de conocer estados de opinión, ideas, características o hechos específicos.

Existen varios métodos para recolectar dichos datos, clasificándolos según su medio de captura. A cada uno de ellos se lo conoce con sus siglas en inglés. A continuación se describen alguno de ellos:

- **PAPI:** Las siglas significan Encuesta en lápiz y papel (Paper And Pencil Interviewing). Se refiere al método de realizar una entrevista e ingresar los resultados de la entrevista en un papel.
- **CAPI:** Las siglas significan Encuesta persona a persona, asistida por computadora (Computer Assisted Personal Interviewing). Este método es similar al anterior con la diferencia que los resultados de la entrevista se ingresan directamente en una computadora, en vez de en un formulario en papel.
- **CASI:** Las siglas significan Encuesta auto asistida por computadora (Computer Assited Self Interviewing). A diferencia del método anterior, en este caso es el propio entrevistado quien completa el cuestionario.

- **CATI:** Las siglas significan Encuesta asistida por teléfono (Computer Assited Telephone Interviewing). La tasa de rechazo de una encuesta CATI es comparativamente baja, siempre que se cuenten con encuestadores bien entrenados que logren persuadir al entrevistado para lograr la entrevista y mantener su atención.
- **CAWI:** Las siglas significan Encuesta asistida por web (Computer Assited WebInterviewing). Las encuestas vía WEB son usadas cuando no es requerido un encuestador. El problema de este tipo de encuesta es la baja tasa de respuestas, dado que habitualmente el sujeto no se motiva a responder, a no ser que tenga algún interés en los resultados del estudio o porque está recibiendo un pago o premio.

Se realizaron varios estudios y experimentos que comparan los métodos de recolección de datos que demuestran sus ventajas y desventajas. Dos de los métodos más comparados son PAPI y CAPI. El método CAPI tiene varias ventajas, las más destacadas son la reducción de errores y de tiempo [9]. Sin embargo, el papel sigue siendo el medio más usado a pesar de los avances tecnológicos de las últimas décadas por su bajo costo, versatilidad y seguridad. La tasa de rechazos de una encuesta en papel mediante encuestador es muy baja.

3 DDI

Cómo y dónde se empieza la captura de metadatos depende de los datos que se describen, la aplicación en la que se utiliza, y las necesidades organizativas de los creadores.

Los beneficios de usar DDI son varios para todos los involucrados, productores de datos, archivistas y usuarios de datos. A continuación se listan algunas ventajas que brinda DDI:

- Interoperabilidad: documentos escritos con la especificación DDI se pueden intercambiar y transportar fácilmente. Las aplicaciones pueden ser desarrolladas para trabajar con estos documentos homogéneos.
- Rico en contenido: DDI fue diseñado para fomentar el uso de un conjunto completo de elementos, para describir conjuntos de datos de ciencias sociales tan completa y tan a fondo como sea posible. De esta manera proporciona al analista de datos, un conocimiento más amplio acerca de una determinada colección.
- Único documento, múltiples propósitos: La documentación contiene toda la información necesaria para producir diferentes salidas, esta puede ser reutilizada para diferentes necesidades y aplicaciones.
- Análisis on-line: Dado que DDI se extiende hasta el nivel de variable y proporciona una estructura uniforme estándar y el contenido de las variables, los documentos en DDI son fácilmente importados a sistemas de análisis en línea, haciendo que los conjuntos de datos sean más fáciles de utilizar para un público más amplio.
- Precisión en la búsqueda: Dado que en la documentación se etiqueta de manera particular, es fácil la búsqueda de campos específicos.

La especificación DDI permite describir todos los aspectos de un proyecto estadístico a detalle (antecedentes, metodología, áreas responsables, archivos de datos y variables), provee una lista comprensiva y estructurada de elementos y atributos para documentar la base de datos. No todos son necesarios para todos los proyectos estadísticos; por eso define campos obligatorios, recomendados y opcionales.

Hoy en día coexisten dos versiones del estándar DDI, una llamada DDI-CodeBook y otra DDI-Lifecycle. La primera es una versión menos cargada que tiene como principal objetivo documentar los datos de encuestas simples. Mientras que la segunda, abarca toda la especificación DDI-Codebook y la amplía. DDI-Lifecycle está diseñado para documentar y gestionar los datos a través de todo el ciclo de vida, desde la conceptualización a la publicación y análisis de datos y más allá.

3.1 DDI-CODEBOOK

Línea de desarrollo destinado a proporcionar la documentación completa necesaria para utilizar con eficacia y analizar los datos de las ciencias sociales, centrándose en encuestas simples y conjuntos de datos.

Originalmente creado para documentar conjuntos de datos que se archivan y para poder difundir entre los investigadores. DDI - Codebook incluye la misma información que un libro de códigos tradicional, que describe variables, preguntas, categorías y códigos usados como dominios de respuesta y valores de variables. Los datos se pueden mantener en formatos de paquetes estadísticos o en ASCII con suficiente información en el archivo de metadatos DDI para entender qué valores están asociados con las variables. Los metadatos están estructurados en un formato XML.

Codebook fue la primera versión de DDI, que fue publicada en el año 2000. La especificación del Codebook desde la versión 1 hasta la 2.1 está escrita en DTD (Data Type Documentation). Un documento de tipo DTD es una descripción de estructura y sintaxis de un documento XML. La última versión del Codebook (v2.5), la cual fue publicada a principios del 2014, está especificada en XML Schemas y facilita la migración de los documentos en ésta versión, a la versión DDI-Lifecycle.

XML Schemas son los sucesores de las DTD y tiene varias ventajas. La primera es que los esquemas soportan una amplia gama de tipos de datos, a diferencia de los DTDs que solamente admiten tipos de datos primitivos. Los esquemas incluyen el uso de namespaces y además están escritos en lenguaje XML y no en un lenguaje especializado, permitiendo el manejo y análisis como cualquier documento XML. [10]

A continuación se describen partes que se consideran importantes de la última versión del Codebook (v2.5) detallado en la página web de DDI Alliance. [11]

3.1.1 GRUPOS DE ATRIBUTOS

El siguiente es un fragmento del archivo codebook.xsd, donde se describe un grupo de atributos:

```
<xs:attributeGroup name="GLOBALS">
  <xs:attribute name="ID" type="xs:ID" use="optional"/>
  <xs:attribute name="xml:lang" type="xs:NMTOKEN" use="optional">
    <xs:annotation>
      <xs:documentation>DO NOT USE THIS ATTRIBUTE. Its inclusion is an error that was persisted to retain backward
compatibility. If this attribute has been used, transfer the content to xml:lang.
      </xs:documentation>
    </xs:annotation>
  </xs:attribute>
```

```

<xs:attribute ref="xml:lang" use="optional"/>
<xs:attribute name="source" default="producer">
  <xs:simpleType>
    <xs:restriction base="xs:NMTOKEN">
      <xs:enumeration value="archive"/>
      <xs:enumeration value="producer"/>
    </xs:restriction>
  </xs:simpleType>
</xs:attribute>
<xs:attribute name="elementVersion" type="xs:string" use="optional">
  <xs:annotation>
    <xs:documentation>Captures version of the element</xs:documentation>
  </xs:annotation>
</xs:attribute>
<xs:attribute name="elementVersionDate" type="dateSimpleType" use="optional">
  <xs:annotation>
    <xs:documentation>Indicates version date for the element. Use YYYY-MM-DD, YYYY-MM, or YYYY formats.
  </xs:documentation>
  </xs:annotation>
</xs:attribute>
<xs:attribute name="ddiLifecycleUrn" type="xs:anyURI" use="optional">
  <xs:annotation>
    <xs:documentation>Used to capture the DDI-Lifecycle type URN for the element. This may be captured during translation from
    DDI-Lifecycle to DDI-Codebook structure or in preparation for transferring to a DDI-Lifecycle structure.
  </xs:documentation>
  </xs:annotation>
</xs:attribute>
<xs:attribute name="ddiCodebookUrn" type="xs:anyURI" use="optional">
  <xs:annotation>
    <xs:documentation>Used to capture the DDI-Codebook type URN for the element. This is used to assign a DDI-Codebook specific
    URN to the element, according the format prescribed by the DDI-Codebook standard.</xs:documentation>
  </xs:annotation>
</xs:attribute>
</xs:attributeGroup>

```

Detalle de cada atributo:

- **ID** : identificador único del elemento
- **xml:lang** : este atributo está discontinuado, permanece para mantener la compatibilidad con versiones anteriores, el que se utiliza es el atributo `xml:lang`
- **xml:lang** : idioma del texto que está en el contenido del elemento

- **source** : enumerado que indica la fuente que proporcionó la información del elemento
- **elementVersion** : versión del elemento
- **elementVersionDate** : día de la versión del elemento
- **ddiLifecycleUrn** : es utilizado para la migración entre las dos versiones del estándar
- **ddiCodebookUrn** : especifica la URN (Uniform Resource Name) del elemento [12]

Todos los elementos del Codebook son una extensión de un tipo de elemento abstracto (**baseElementType**), el cual tiene este grupo de atributos (**GLOBALS**) definido. Por lo tanto, todos los elementos del Codebook van a tener la posibilidad de definir cualquier atributo del grupo descrito anteriormente.

3.1.2 DESCRIPCIÓN DEL DOCUMENTO <DOCDSCR>

En la primer parte de cualquier ejemplo de un Codebook aparece la información general del documento, como es la descripción, título, versión, etc. A continuación se muestra un fragmento de un ejemplo de Codebook, donde podemos ver los campos que contiene la descripción del documento:

```
<docDscr>
  <citation>
    <titlStmt>
      <titl>
        ENHA2006_02
      </titl>
      <IDNo>
        URY-INE-ENHA-2006-v02
      </IDNo>
    </titlStmt>
    <prodStmt>
      <producer affiliation="Asesoría Técnica - INE" role="Documentador">
        Enrico Benedetti
      </producer>
      <prodDate date="2009-03-18">
        2009-03-18
      </prodDate>
      <software version="1.2" date="2011-01-07">
        Metadata Editor
      </software>
    </prodStmt>
  </verStmt>
```

```

    <version type="02"/>
  </verStmt>
</citation>
</docDscr>

```

En esta sección se redacta la información bibliográfica del documento, es importante que esté completa, ya que se utiliza para identificar el archivo dentro de sistemas de búsqueda.

A continuación se detallan las grandes secciones que tiene la descripción del documento, algunas se pueden o no querer introducir en el Codebook. Como se puede ver en el ejemplo anterior, solamente se ingresó la información de la cita bibliográfica.

- **Cita bibliográfica <citation>**

Información bibliográfica que incluye información sobre el título, la mención de responsabilidad, información de producción y distribución, la serie y la información de la versión, el texto de una cita bibliográfica y notas (si las hay).

El autor dentro de esta sección, es la persona u organización directamente responsable por el contenido intelectual del documento, que puede no ser la misma que el responsable de la versión del documento que se toma como base, esta información se describe en otra sección (<docSrc>).

- **Guía del documento <guide>**

Bajo el tag <guide> se pueden definir una lista de términos y definiciones que se utilicen a lo largo del documento y se podría describir alguna ayuda para los usuarios.

- **Status del documento <docStatus>**

Se utiliza este campo para indicar si se está presentando o distribuyendo la documentación antes de que haya sido finalizado. Algunas organizaciones, emplean estrategias de procesamiento de datos que proporcionan una liberación de los datos y la documentación, en las distintas etapas de la elaboración.

- **Fuente del documento <docSrc>**

Esta sección contiene la información bibliográfica de la fuente del documento, incluye información sobre el título, la mención de responsabilidad, información de producción y distribución, la serie y la versión de la información, el texto de una cita bibliográfica y notas (si hay).

3.1.3 DESCRIPCIÓN DEL ESTUDIO

En esta sección (dentro del tag <stdyDscr>) se incluye información acerca de cómo se debería citar el estudio, quien recogió o compiló los datos, quien distribuye los datos, palabras clave sobre el contenido de los datos, el resumen (abstract) del contenido de los datos, los métodos y el procesamiento de recolección de datos, etc. Parte del

contenido del estudio de descripción de citas - por ejemplo, la Declaración de responsabilidad - puede ser idéntica a la de la cita de documentación. Este suele ser el caso cuando el productor de una colección de datos también produce la impresión o el libro de códigos electrónico para la recolección de datos.

A continuación se listan la información que se puede encontrar dentro de la descripción del estudio:

- **Cita bibliográfica <citation>**

Información bibliográfica que incluye información sobre el título, la mención de responsabilidad, información de producción y distribución, la serie y la información de la versión, el texto de una cita bibliográfica y notas (si las hay).

- **Alcance del estudio <studyInfo>**

Esta sección contiene información sobre el alcance de la recopilación de datos, incluyendo el contenido intelectual e información cronológica y geográfica.

Dentro de la información intelectual se definen palabras claves y clasificación de tópicos que cubren los datos.

Dentro del tag <abstract> de esta sección se resume el propósito, la naturaleza y el alcance de la recopilación de datos, las características especiales de su contenido, las principales áreas temáticas cubiertas, y qué preguntas se intentó responder cuando se realizó el estudio.

Luego se describe la información cronológica y geográfica. Se permite definir un periodo cubierto por los datos, fecha en la cual se recolectaron los datos, país o países en los cuales se recolectan los datos, información de cubrimiento geográfico, unidades de análisis, universos, tipos de datos (encuesta, censo, etc.).

- **Metodología y procesamiento <method>**

En esta sección se describe la metodología y el procesamiento empleados en la recopilación de datos. Se detalla la entidad que recolectó los datos, frecuencia con la cual se recolectaron, método de muestreo, tipo de instrumento utilizado, fuentes utilizadas.

- **Acceso a los datos <dataAccs>**

En esta sección se describen las condiciones de acceso y las condiciones de uso para la recolección de datos.

- **Condiciones de uso <otherStudyMat>**

En esta sección se puede agregar información de otro material asociado a la descripción del estudio, como apéndices, ejemplos de muestreos, detalles técnicos, etc.

Información sobre el o los archivos de datos que comprende una colección. Esta sección se puede repetir para colecciones con varios archivos.

Se puede describir el nombre del archivo, contenido, estructura, tamaño, tipo, formato de la información, versión, entre otras cosas.

3.1.5 DESCRIPCIÓN DE VARIABLES <DATADSCR>

Dentro de esta sección se encuentra la información de las variables. Existe la posibilidad de agrupar variables y nCubes.

- Grupo de variables <varGrp>

Grupo de variables que pueden compartir un tema común, surgir de la interpretación de una sola pregunta, o estar unidas por algún otro factor.

- Variable <var>

Dentro de esta sección se describen todas las variables del cuestionario. Cada variable tiene los siguientes posibles atributos:

- name: nombre de la variable, se podría limitar el largo de los caracteres
- wgt: indica si la variable es una ponderación
- wgt-var: referencia de variables que son la ponderación de esta
- qstn: ID de la pregunta para esta variable
- files: IDREF que identifica el o los archivos de donde pertenece la variable
- vendor: origen del formato puede ser SAS, SPSS, ANSI, and ISO
- dcml: cantidad de decimales de la variable
- intrvl: tipo de intervalo, puede ser continuo o discreto
- rectype: tipo de registro de la variable
- sdatrefs: referencias a resúmenes detallados en la sección de Descripción del estudio
- methrefs: referencias a metodologías y procesos detallados en la sección de Descripción del estudio
- pubrefs: link a publicación o citación
- access: ID de todos los elementos de la sección de acceso a datos que describen las condiciones de acceso para esta variable
- aggrMeth: tipo de método de cálculo de la variable (ej: suma, porcentaje)
- measUnit: unidad de medida de la variable (ej: kilómetros, metros)
- scale: escala de la variable (ej.: x1, x100)
- origin: punto de origen
- nature: naturaleza de la variable (nominal, ordinaria, intervalo, relación)

- additivity: tipo de adición, puede ser stock, flujo
- temporal: indica si la variable tiene información relacionada con el tiempo
- geog: indica si la variable tiene información geográfica
- geoVocab: esquema de codificación utilizado en la variable
- catQty: cantidad de categorías encontradas en la variable

Dentro de la variable existe una sección <qstn> donde se puede definir una pregunta, en caso de que la variable no sea una pregunta se deja la sección vacía. Cada pregunta tiene las siguientes secciones:

- preQTxt: Texto que describe condiciones bajo las cuales se debe realizar la pregunta
- qstnLit: Texto con la pregunta
- postQTxt: Texto que describe que va a ocurrir luego de realizar la pregunta
- forward: referencias a posibles preguntas posteriores
- backward: referencias a posibles preguntas anteriores
- ivuInstr: instrucciones para quien entrevista

3.1.6 OTRO MATERIAL RELACIONADO CON EL ESTUDIO <OTHERMAT>

Dentro de esta sección se permite incluir material relacionado al estudio como pueden ser manuales de usuario, glosarios, instrucciones de entrevistas, mapas, esquemas de bases de datos, etc.

3.2 DDI-LIFECYCLE

Línea de desarrollo destinado a documentar y gestionar datos en todo el ciclo de vida, desde la conceptualización hasta la publicación y más allá.

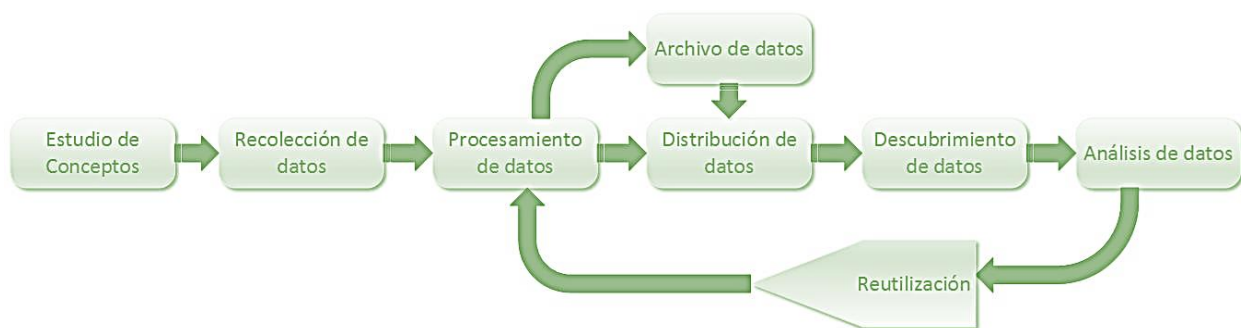


Fig. 1 - Etapas del ciclo de vida

En DDI-Lifecycle (a partir de la versión 3.0) el alcance del estándar ha aumentado, antiguamente DDI estaba focalizado en archivos de datos, actualmente soporta todo el ciclo de vida de los datos. Se puede expresar todo el proceso de evolución de la información. El diagrama anterior muestra las distintas etapas del ciclo de vida. Como se puede ver en la figura 1, es un método iterativo, se considera que la información puede ser reprocesada.

Este cambio de enfoque tiene varias repercusiones en el diseño del estándar. Las instancias van a ser más extensas, y los datos de los estudios se pueden encontrar en varios archivos con más flexibilidad. La relación de estudios e información de estudios base, puede ser detallada, así como la comparación de ellos. A pesar de que el enfoque de esta versión es muy ambicioso, se consideró el diseño para que la migración de versiones anteriores no sea muy costosa. El uso simple de DDI no tiene grandes diferencias entre las versiones 2.* y 3.*.

Se permite detallar información de la encuesta de una forma más rica. Muchos sistemas permiten la reutilización de preguntas, por lo que los metadatos de instrumentos son una parte importante del ciclo de vida. También se permite la reutilización de distintos elementos, como notas y citas, para que esto sea posible, se manejan referencias a materiales externos e internos.

A partir de la versión 3.0 de DDI se comenzó a escribir un modelo conceptual del estándar, además de los esquemas o DTD. Este documento puede ser de mucha utilidad para los usuarios y desarrolladores de aplicaciones que deben soportar DDI, ya que el vocabulario se vuelve más coherente y comprensible.

3.2.1 USO SIMPLE VS COMPLEJO

Existen dos caminos para la utilización de DDI. En casos que el estudio se deba repetir en distintos momentos y éstos se tengan que relacionar entre sí, se debe trabajar bajo un modelo complejo de DDI. En casos que el estudio sea único y no involucra grupos de estudios, se va a trabajar bajo un modelo simple. El modelo simple pretende representar el uso de DDI similar al de versiones anteriores.

A continuación se muestra un diagrama que determina si estamos bajo un modelo complejo o simple:

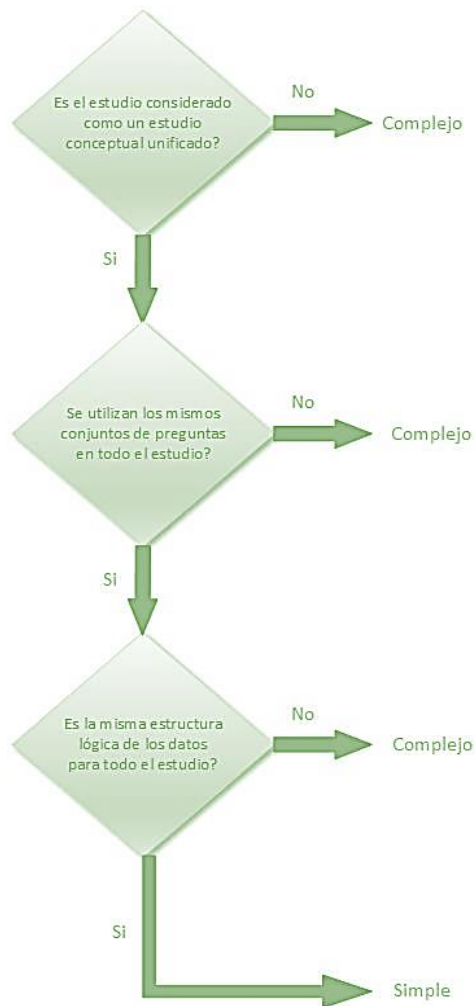


Fig. 2 - Determinación de Modelo Simple o Modelo Complejo

3.2.2 RELACIÓN CON OTROS ESTÁNDARES

DDI se construyó pensando para soportar varios estándares, así como las versiones anteriores. A continuación se listan los estándares de los cuales se mapearon relaciones en DDI 3:

- DDI-C
- Dublin Core (Basic Bibliographic Information)
- MARC (Bibliographic Information) [13]
- GSIM (Generic Statistical Information Model) [14]
- ISO/IES 11179 Data Registry [15]
- ISO 19118 (Geography) [16]

- SDMX (Aggregate data) [17]
- METS (Content Wrapper) [18]
- PREMIS (Preservation) [19]

3.2.3 ETAPAS DEL CICLO DE VIDA

El modelo GSBPM (Generic Statistical Business Process Model) [20] es una herramienta para describir y definir, en términos generales, procesos de negocio para elaborar estadísticas oficiales. Si bien GSBPM aplica a las organizaciones estadísticas oficiales y DDI no tiene porqué, los modelos de procesos tienen sus diferencias pero se asemejan bastante. Las diferencias entre los modelos son:

- GSBPM coloca la fase de archivo al final del proceso. Esta también puede ser la última etapa en DDI para una determinada organización, pero la diferencia es que en el modelo DDI no se limitan los procesos a una única organización. Por ejemplo las etapas de análisis y reutilización pueden llevarse a cabo en distintas organizaciones.
- El modelo DDI representa la fase de difusión con "Distribución de Datos", que tiene lugar antes de la fase de análisis. Esto refleja una diferencia de enfoque entre las comunidades de investigación y estadísticas oficiales, con este último poniendo un mayor énfasis en la difusión de datos, en lugar de la investigación basada en los datos difundidos por otros.
- DDI contiene el proceso de "Reutilización". Esto se trata en la primera fase de GSBPM (Especificar Necesidades), donde hay un sub-proceso para comprobar la disponibilidad de los datos existentes, y utilizarlos siempre que sea posible. También se refleja en el sub-proceso de integración de datos en la quinta fase (Proceso).
- El modelo DDI tiene fases separadas para el descubrimiento de datos y análisis de datos, mientras que estas funciones se combinan en la fase seis (Análisis) en GSBPM. En algunos casos, los elementos de la fase de análisis GSBPM pueden también ser cubiertos en la fase de "Procesamiento de datos" DDI, dependiendo de la magnitud del trabajo de análisis previo a la fase de "distribución de datos".



Fig. 3 - Modelo GSBPM

GSBPM	DDI
Especificar necesidades	Estudio de Conceptos Reutilización
Diseño	
Construcción	
Colección	Recolección de datos
Proceso	Procesamiento de datos Reutilización
Análisis	Descubrimiento de datos Análisis de datos Procesamiento de datos
Difusión	Distribución de datos
Archivo	Archivo de datos
Evaluación	

Fig. 4 - Modelo DDI vs Modelo GSBPM

En la Figura 4 se puede ver una analogía de ambos tipos de modelos.

A continuación se describen las actividades de cada etapa:

- **Especificar necesidades:**
 - Determinar las necesidades de la información
 - Consultar y confirmar necesidades
 - Establecer objetivos de salida
 - Identificar conceptos

- Chequear la disponibilidad de la información
- Preparar caso de negocio
- **Diseño**
 - Diseño de salidas
 - Diseño de descripción de variables
 - Diseño de recolección de datos
 - Marco de diseño y metodología de la muestra
 - Diseño de metodología de procesamiento estadístico
 - Diseño de sistema de producción y flujo de trabajo
- **Construcción**
 - Construcción del instrumento de recolección de datos
 - Construir o mejorar los componentes del proceso
 - Configurar flujos de trabajo
 - Testeo de sistemas de producción
 - Testeo de procesos de negocio estadísticos
 - Finalizar sistemas de producción
- **Colección**
 - Seleccionar muestras
 - Configurar la recolección
 - Ejecutar la recolección
 - Finalizar la recolección
- **Procesamiento**
 - Integrar datos
 - Clasificar y codificar
 - Revisar, validar y editar
 - Imputar
 - Derivar nuevas variables y unidades estadísticas
 - Calcular pesos
 - Calcular agregados
 - Finalizar archivos de datos

- **Análisis**
 - Preparar borradores
 - Validar las salidas
 - Examinar y explicar
 - Controlar divulgación
 - Finalizar salidas
- **Difusión**
 - Actualización de sistemas de salidas
 - Producir productos de difusión
 - Gestionar liberación de productos de difusión
 - Promover productos de difusión
 - Gestionar la asistencia al usuario
- **Archivo**
 - Definir reglas de archivado
 - Gestionar repositorio de archivos
 - Preservar los datos y los metadatos asociados
 - Disponer datos y metadatos asociados
- **Evaluación**
 - Reunir insumos de evaluación
 - Evaluar y generar un reporte
 - Acordar un plan de acción

3.2.4 IDENTIFICACIÓN, VERSIONADO, MANTENIMIENTO Y REFERENCIA

La estructura de identificación de los objetos en DDI-L es fundamental en el funcionamiento del estándar. Ésta se basa en la estructura de ISO/IEC 11179, la cual contiene tres partes:

- **Identificador de Agencia**: un identificador único para la agencia que maneja el objeto. Esta agencia es registrada en la web de DDI Alliance y puede tener sub-agencias.
- **Identificador único**: identificador único en el contexto de la agencia o en el objeto padre mantenible.
- **Número de versión**

No todos los elementos se pueden identificar individualmente, ya que algunos tienen sentido dentro de un contexto (ejemplos: descripciones, títulos, etc.). El resto de los objetos son identificables, versionables y mantenibles.

La estructura de la URN en DDI es:

“urn:ddi:agency[.sub-agency]:ID:Version”

3.2.5 INSTANCIAS DDI

Las instancias de DDI-L son publicadas e intercambiadas bajo uno de los dos formatos posibles, DDIInstance y FragmentInstance. El fragmento es uno de los aportes de la versión 3.2, ya que antes no existía el concepto. DDIInstance se puede utilizar para cuatro tipos de documentos: StudyUnit, Grupo, ResourcePackage y LocalHoldingPackage. Mientras que el FragmentInstance es un paquete utilizado para transferir objetos mantenibles o versionables incluyendo cualquier nota u otro material asociado. El manejo de fragmentos permite obtener un archivo XML válido, escrito bajo el estándar DDI, sin tener que incluir toda la información.

3.2.6 INSTRUMENTO

En la versión DDI3 existe el concepto de Instrumento, el cual se refiere a la definición de la estructura de un cuestionario. Para la definición del instrumento se pueden utilizar varias estructuras de control, como puede ser definir una cierta condición para hacer, o no, alguna pregunta. Dichas estructuras de control van definiendo el flujo del cuestionario.

A partir de la metadata generada en DDI del instrumento, es posible definir diferentes instancias del instrumento para la recolección de los datos, como cuestionarios listos para su impresión, cuestionarios online y también puede servir como entrada para otro sistema que permita la manipulación de metadatos y/o cuestionarios. [21]

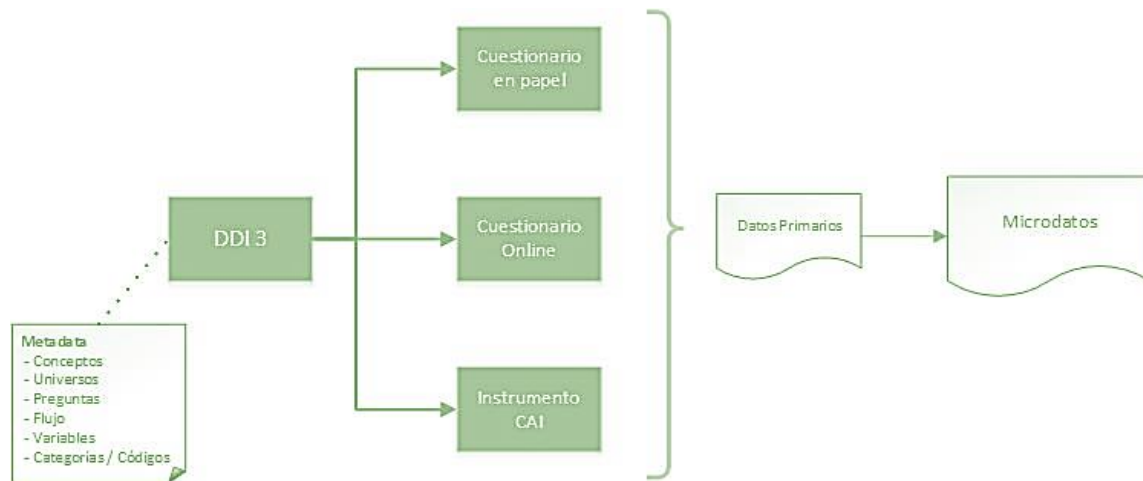


Fig. 5 - Instrumento

Una de las motivaciones que lleva a DDI a guardar metadata del instrumento es la posibilidad de que el instrumento pueda ser leído y procesado a través de sistemas computacionales. Para instrumentos complejos es fundamental la automatización del proceso.

En esta versión es posible agregar preguntas a un instrumento, detallar la estructura, el flujo y la lógica del instrumento y relacionar la información de instrumentos para el análisis de la información. Además se permite guardar otros metadatos, como la metodología y el proceso utilizado para administrar el instrumento.

En la figura 6 se muestra un diagrama de las componentes que tiene un instrumento [22]:

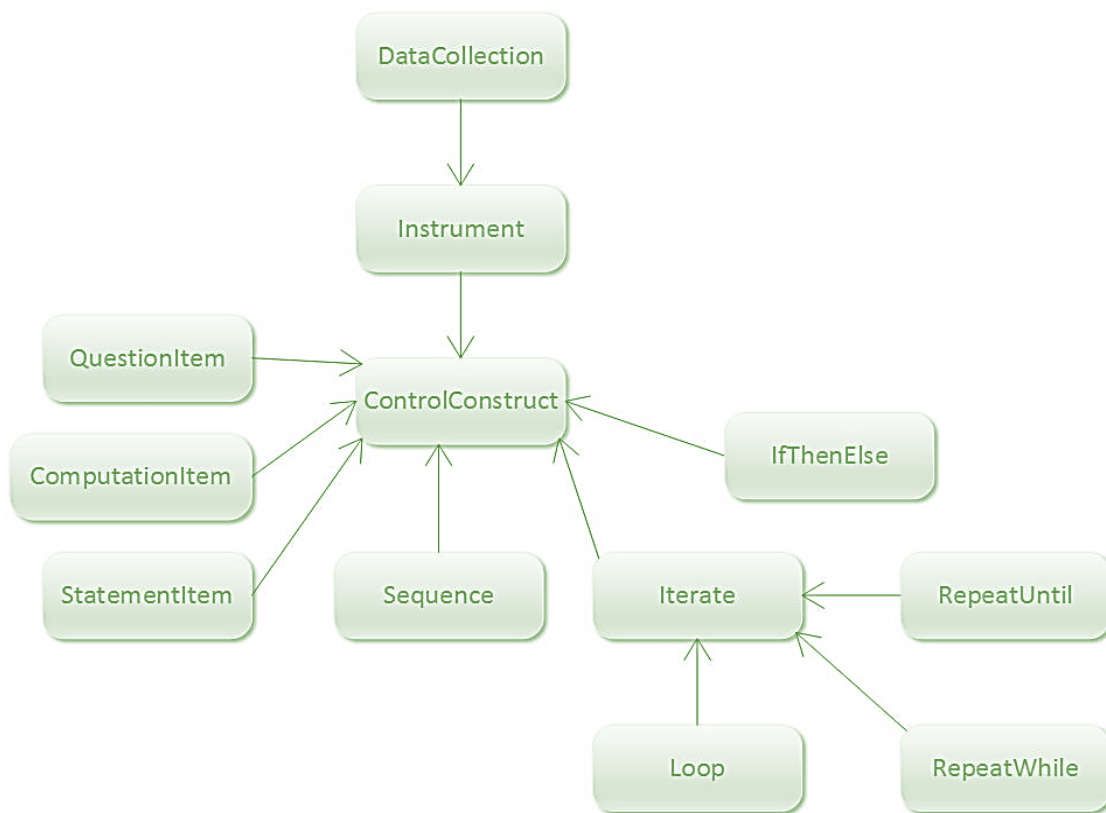


Fig. 6 - Componentes de un instrumento

El instrumento está definido por un conjunto de elementos ordenados. Las posibles estructuras de control y elementos que puede contener un instrumento, son:

- **Pregunta:** Se puede añadir al instrumento una referencia a una pregunta definida en el esquema de preguntas.
- **Secuencia:** Se puede añadir una referencia a una secuencia.
- **IfThenElse:** Define una estructura condicional. Las condiciones se pueden escribir en cualquier lenguaje, por ejemplo Java.
- **Iteración:** Existen tres tipos de iteraciones, RepeatUntil, RepeatWhile y Loop.

- **Ítem computacional:** estructura de control que proporciona un código y la asignación de una variable cuyo valor puede ser utilizado en el código definido.
- **Declaración:** Texto que se quiere mostrar en el instrumento.

Para cada una de las condiciones que se pueden ingresar en el instrumento, se puede definir un lenguaje en el cual va a estar escrita esta condición. Esto permite que el instrumento pueda ser interpretado por sistemas computacionales, y así poder utilizar los datos del instrumento escritos en DDI, como entrada para sistemas que quieran, por ejemplo realizar una encuesta WEB o realizar análisis estadísticos.

Una de las sintaxis utilizadas para escribir las condiciones puede ser SPSS. SPSS es un software de IBM muy utilizado para análisis estadístico de ciencias sociales, también existe la versión gratuita que es PSPP [23].

3.3 DDI-CODEBOOK VS DDI-LIFECYCLE

Existen diferencias muy significativas entre estas dos versiones del estándar DDI, ya que hubo varios avances en la última versión que hicieron que sean bastante diferentes entre sí.

Una de las diferencias más importantes, es la definición del instrumento en la versión ampliada (DDI-Lifecycle). Con la versión DDI-Codebook, muchas veces se califican los microdatos de mala calidad por no tener la información del flujo del cuestionario. En la nueva versión se pueden especificar, bajo la entidad instrumento, distintos elementos que van definiendo el flujo del formulario de manera de poder interpretar los microdatos correctamente.

El texto de las preguntas, en la versión DDI-Codebook, aparece sólo como parte de la descripción de una variable, para la versión ampliada se separan las entidades pregunta y variable, lo que permite compatibilidad con sistemas CAPI. Además la última versión soporta múltiples lenguajes para todos los textos definidos dentro del formulario, esto permite generar por ejemplo, formularios en distintos idiomas a partir del mismo XML.

Finalmente DDI-L tiene un diseño modular, el cual facilita la reutilización, versionado y mantenibilidad. Como muchas organizaciones están involucradas en la creación de metadata, existen reglas de mantenimiento, versionado e identificación que permiten la reutilización. Una agencia que mantiene objetos, es identificada con un código único basado en su dominio el cual va a aparecer en la URN de los objetos a mantener.

4 HERRAMIENTAS RELACIONADAS

A pesar que el estándar es relativamente nuevo en el tiempo se pueden encontrar varias herramientas que lo manejan ya sea Codebook o DDI-Lifecycle.

En este capítulo se presentan algunas de dichas herramientas que manejan el estándar DDI, junto con sus principales funcionalidades, sus ventajas y sus desventajas a nuestro criterio.

4.1 COLECTICA

Colectica es un conjunto de programas privativo para el manejo de estadísticas y encuestas estadísticas utilizando estándares abiertos.

La versión actual de Colectica 4.XX utiliza DDI 3.1. La nueva versión “Colectica 5.0” que actualmente se encuentra en Beta incorpora la última versión del estándar DDI 3.2.

Algunas de las principales funcionalidades que Colectica permite a sus usuarios (investigadores, archivólogos, programadores, etc.) son:

- Diseño de cuestionarios
- Análisis Estadísticos
- Documentar y gestionar metadatos.
- Desarrollo de Aplicaciones

Colectica está actualmente en uso por una variedad de centros de estudio universitarios y centros nacionales de estadística. [24]

Dentro del conjunto de programas que componen Colectica, nos interesa destacar uno de ellos: *Colectica Express* [25], el mismo es una herramienta que contiene funcionalidades interesantes respecto a la manipulación de metadata.

Sus principales funcionalidades son:

- Administración de metadatos estadísticos:
 - Gestión de Estudios Estadísticos.
 - Documentar y difundir datos.
 - Gestionar Clasificaciones.
 - Soporte multilinguaje.
- Diseño de encuestas y cuestionarios:
 - Administrar preguntas.

- Diseñar cuestionarios.
 - Analizar cuestionarios.
 - Generar cuestionarios en papel.
- Publicar Documentación:
 - Generar documentación de metadatos.
 - Generación de códigos estadísticos.
 - Importación:
 - Importar metadata de otros estándares.
 - Exportación:
 - Exportación a estándares de código abierto.
 - Compatibilidad con DDI3.

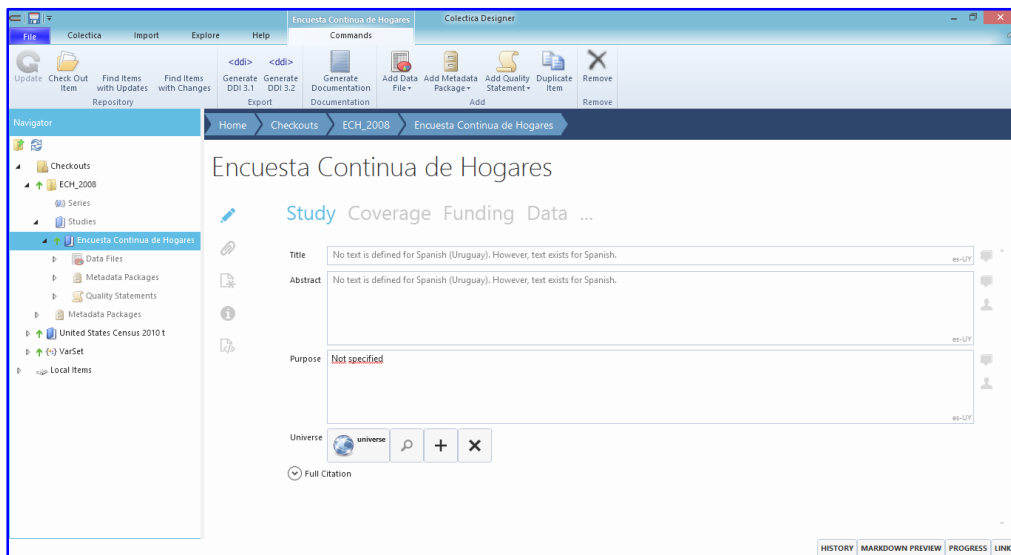


Fig. 7 - Colectica

Luego de ponernos en contactos con Colectica y conseguir una licencia de prueba, tanto de su versión actual como de la versión beta, podemos afirmar que resulta de gran utilidad la funcionalidad de conversión de metadatos de un estándar a otro.

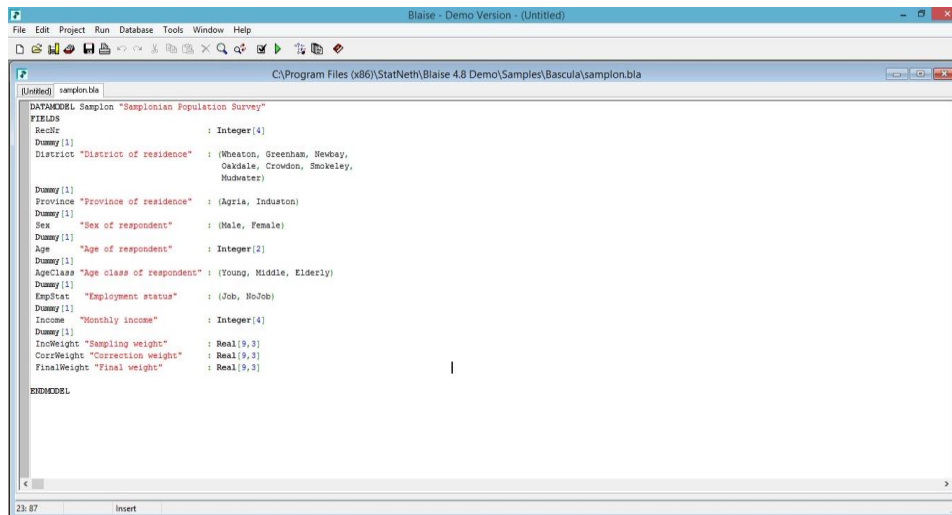
4.2 BLAISE

Se trata de un lenguaje privativo que permite construir cuestionarios, no es un estándar.

Blaise es adecuado para la creación de encuestas a través de un computador. Las preguntas pueden ser descritas en un modelo de datos. El modelo de datos describe una estructura fija de los datos de la encuesta y también puede

contener reglas, en las que, por ejemplo, puede especificar el orden de las preguntas y los condicionales. En muchos casos un modelo de datos Blaise describe un cuestionario Blaise. [26]

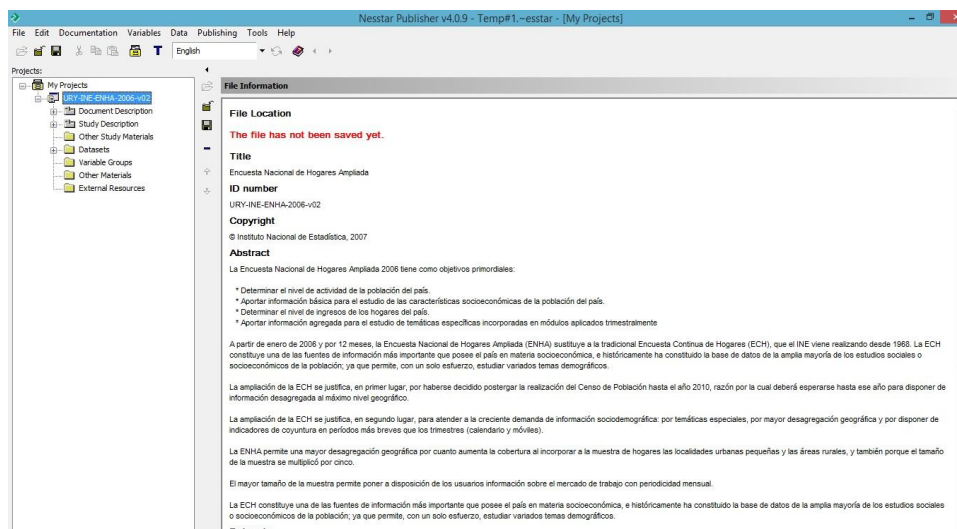
Es posible instanciar el cuestionario en distintos tipos (CAPI, CATI, CASI, CADE, CAWI).



4.3 NESSTAR

4.3.1 NESSTAR PUBLISHER

Ya fue presentado brevemente cómo una de las herramientas promovida por la IHSN. Es un editor para la preparación de metadatos y datos, como también para la publicación de los mismos en un catálogo. Maneja metadatos compatibles con la versión 2.x del estándar DDI y con el estándar Dublin Core. La aplicación fue desarrollada por Nesstar y está catalogada como un software libre. [27] [28]



4.3.2 NESSTAR SERVER

Ofrece todos los servicios habituales para la publicación de contenidos web, el servidor de Nesstar ofrece la capacidad de publicar información estadística que los usuarios podrán buscar, navegar, analizar y descargar. [29]

4.3.3 NESSTAR WEBVIEW

Es un sistema web para la difusión de datos. Puede ser utilizado para visualizar datos y metadatos que han sido publicados usando Nesstar Editorial.

Permite a los usuarios buscar, localizar, navegar, analizar, y descargar una gran variedad de datos estadísticos dentro de un navegador web.

No es necesario ningún software, Nesstar WebView permite a cualquier persona acceder a los datos contenidos en un Servidor Nesstar. [30]

Nesstar con su conjunto de aplicaciones, facilita la manipulación de metadatos y los propios datos.

5 ANÁLISIS

El objetivo del proyecto es construir una herramienta que permita la generación de formularios usando el estándar DDI en su última versión (DDI-Lifecycle 3.2). Además la herramienta debe permitir el manejo de formularios escritos en la versión DDI-Codebook.

La importancia de la herramienta está en la automatización del ingreso de formularios. Es fundamental tener información de los metadatos del formulario para poder definir reglas de validación que determinen la consistencia de los microdatos.

Como se vio en la sección [4.2 DDI-LifeCycle](#), la versión extendida del estándar describe un ciclo de vida de los datos. La herramienta a desarrollar está contenida en la etapa de “Estudio de Conceptos” del ciclo de vida, ya que en esta etapa es donde se realiza el diseño y la construcción. (Fig. 8)

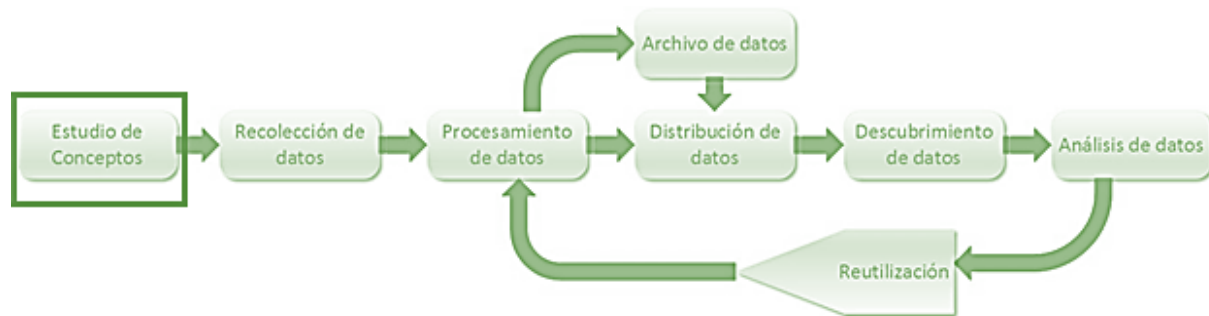


Fig. 8 - Herramienta en ciclo de vida

Más específicamente, la herramienta va a permitir al usuario realizar las siguientes actividades de las etapas que se vieron del modelo GSBPM:

- **Diseño**
 - Diseño de descripción de variables: De forma fácil e intuitiva se permite definir variables de diferentes tipos, para luego ser utilizadas en las preguntas que va a contener el instrumento.
 - Diseño de salidas: Se permite diseñar el instrumento que luego va a ser exportado en formato papel, para utilizar en la realización de una encuesta.
- **Construcción**
 - Construcción del instrumento de recolección de datos: La herramienta contiene una interfaz que permite la creación de instrumentos junto con sus metadatos asociados.

5.1 REQUERIMIENTOS

A continuación se describen los requerimientos del sistema:

RF1 - Importar de DDI 2 a DDI 3

El sistema debe permitir migrar un formulario escrito en DDI-Codebook a DDI-Lifecycle, utilizando las últimas versiones. Para DDI-Codebook la versión 2.3 y para DDI-Lifecycle la versión 3.2.

El cambio más significativo de las instancias de DDI en las versiones 3.*, es el uso explícito y requerido de namespaces XML. Cada namespace usa un prefijo que identifica el módulo de donde la descripción del elemento es extraída, a diferencia de DDI 2, donde no existen módulos definidos.

RF2 - ABM Preguntas

El sistema debe contar con un ABM de preguntas, para que estas puedan ser reutilizadas en distintos formularios. La idea es dar de alta las preguntas que se quieren agregar al formulario y luego crear el formulario con su flujo correspondiente.

RF3 - Diseño visual del formulario – Interfaz Gráfica.

Para construir el formulario es necesaria una interfaz gráfica, donde visualizar de forma intuitiva el flujo del mismo. En dicha interfaz debe ser posible la utilización de distintos elementos para la creación del formulario, como ser, preguntas, secuencias, estructuras de control, etc.

RF4 - Generación de formulario en papel (RTF)

El sistema debe contar con la generación del formulario en formato RTF, esto permite que el usuario pueda editar el documento final, ya sea para organizar el cuestionario, como para personalizarlo a su gusto.

RF5 - Exportar a DDI 3

Luego de creado el formulario se va a permitir generar la metadata en DDI v3.2.

RF6 - Importar DDI 3

Dado un archivo XML escrito bajo DDI 3, se va a permitir visualizar la información del mismo por medio de la interfaz gráfica y se va a permitir su modificación.

RF7 - Validar DDI 3

Dado un archivo XML, al momento de importarlo para trabajar con él, se validará que el mismo contenga una estructura válida acorde a la especificación del estándar en su versión 3.2

RF8 - Edición /Exportación de DDI 2

Se va a permitir modificar un Codebook, sin tener que exportar la información del Codebook a la versión DDI 3.

6 DISEÑO

6.1 FUNCIONALIDADES DE LA APLICACIÓN

La figura 9 muestra un diagrama que pretende describir las funcionalidades del sistema a desarrollar:

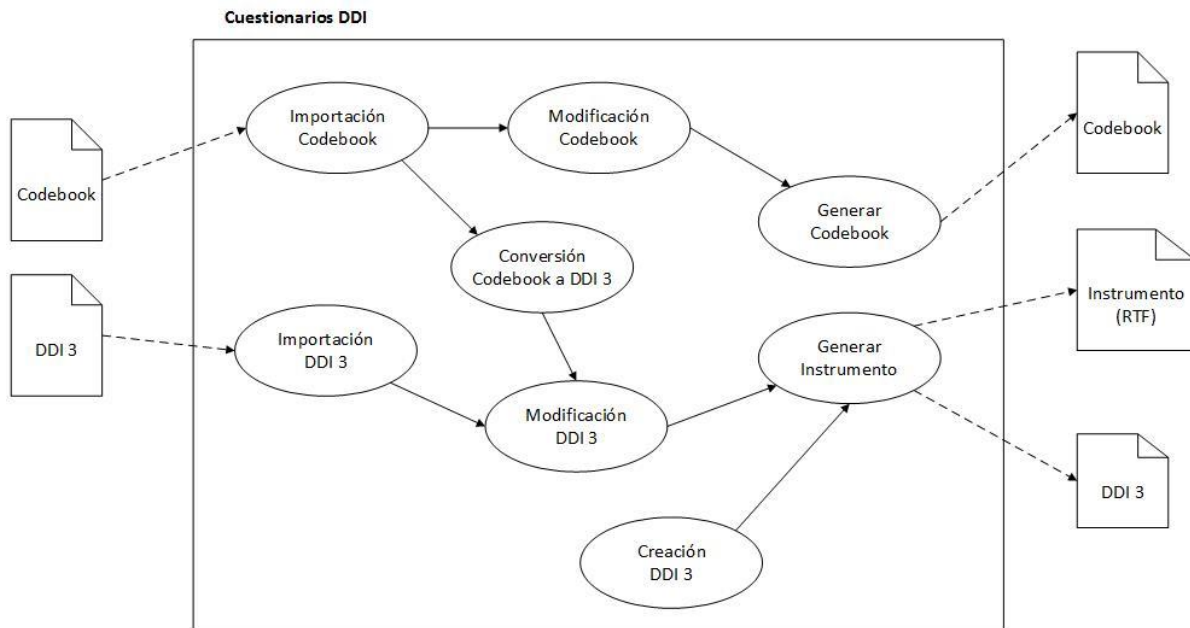


Fig. 9 – Funcionalidades

El sistema debe ser capaz de importar archivos escritos tanto en la versión DDI 3, como en la versión DDI 2. Los datos importados van a ser desplegados por la aplicación de manera clara e intuitiva para que el usuario pueda editar y exportar en cualquiera de las dos versiones.

Además es importante la funcionalidad de poder exportar el instrumento diseñado dentro de la aplicación, en formato RTF. Se seleccionó dicho formato para darle la libertad al usuario de poder editar el archivo exportado.

6.2 ESTRUCTURA DE LA APLICACIÓN

Para dejar la solución comprensible y mantenible vemos necesario dividirla en diferentes proyectos, ya que existen tres grandes e independientes secciones de la aplicación.

Se definen dos proyectos cuyo objetivo es el manejo de la importación y exportación de cada una de las versiones del estándar DDI: DDI 2 y DDI 3. Cada uno de estos proyectos va a contener todo lo necesario para poder leer un archivo escrito bajo el estándar correspondiente y escribir un archivo válido en la versión correspondiente.

Además se define un tercer proyecto que va a contener una capa de presentación y una segunda capa que va a ser la encargada de comunicarse con los proyectos anteriores y va a contener toda la lógica de negocio necesaria, para otorgarle la información a la capa de presentación.

En la figura 9 se muestra un diagrama de los tres proyectos con sus respectivas capas.

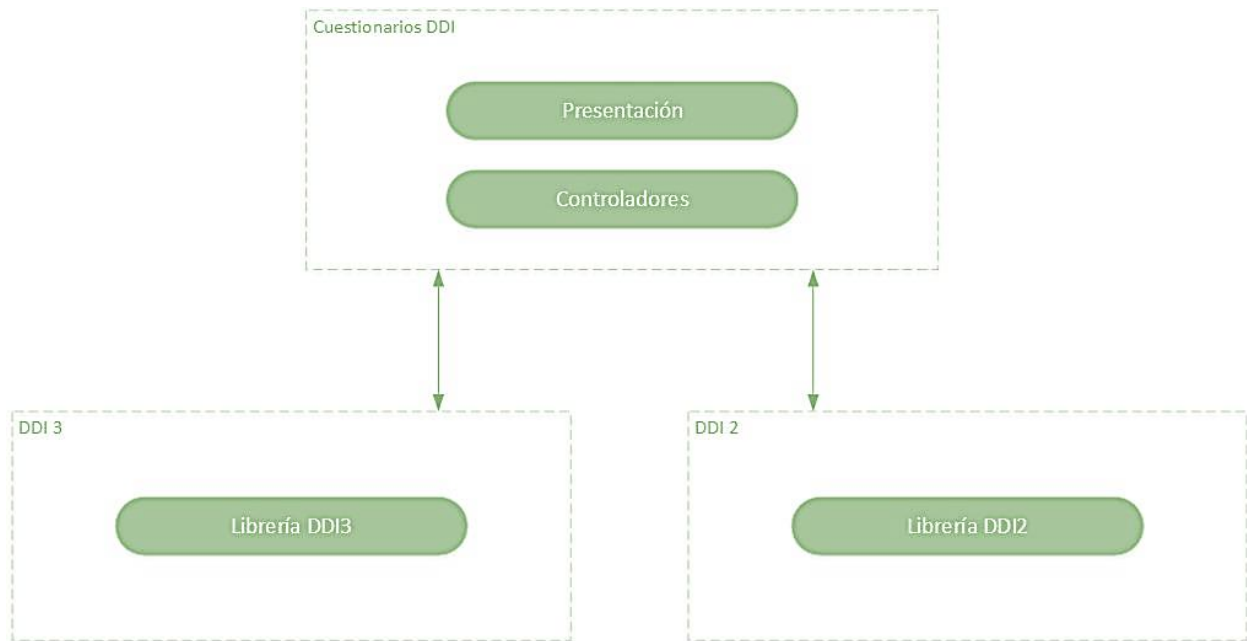
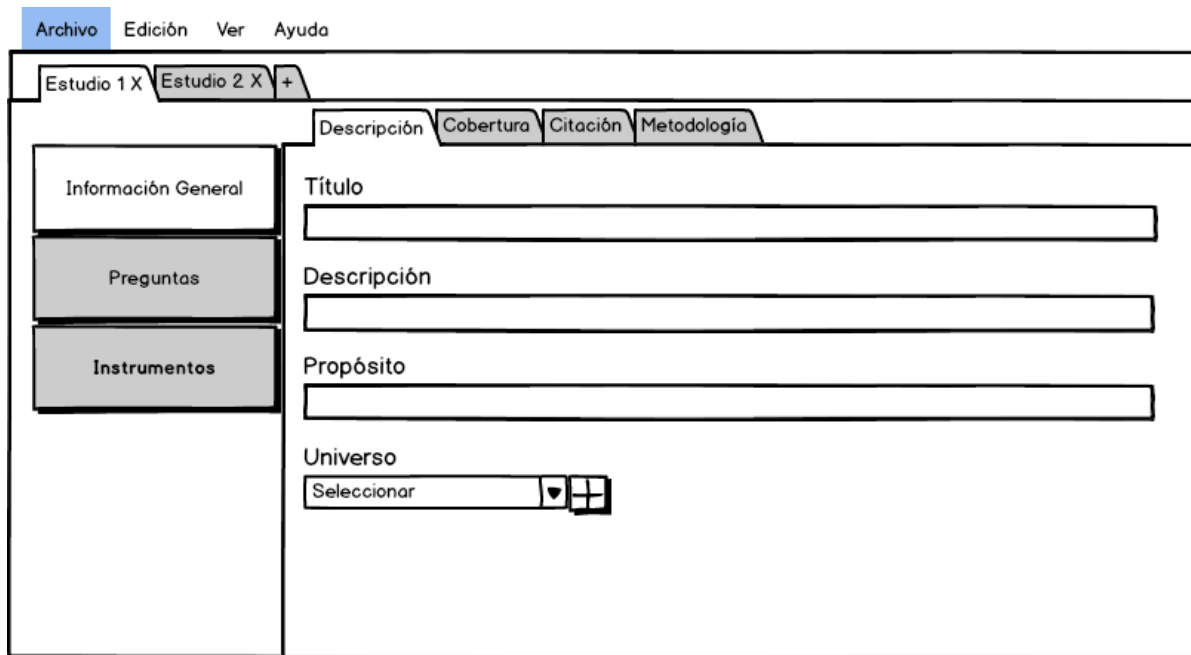


Fig. 10 - Capas

La lógica más compleja y crítica se encuentra en las capas que manipulan el estándar, además es necesaria una interfaz intuitiva y fácilmente manejable, por lo tanto la capa de presentación resulta también fundamental en la solución.

6.3 INTERFAZ

Se realizaron varios prototipos para el diseño de distintas secciones de la aplicación. Como se va a poder ver más adelante, el sistema desarrollado tiene más funcionalidades, pero en términos generales se mantuvo lo diseñado en los prototipos. A continuación se muestran las imágenes de ellos:



Archivo Edición Ver Ayuda

Estudio 1 X Estudio 2 X +

Descripción Cobertura Citación Metodología

Información General

Preguntas

Instrumentos

Título

Descripción

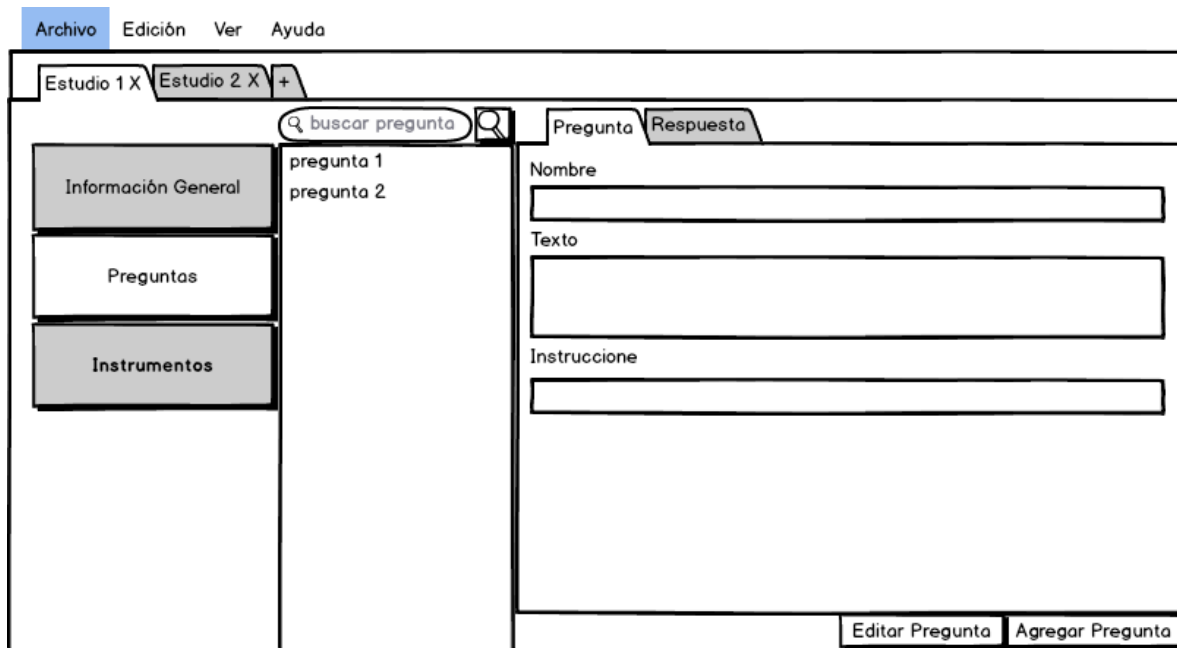
Propósito

Universo

Seleccionar ▼

Fig. 11 - Prototipo Información General

Como muestra la figura 11, la herramienta se divide en distintas pestañas, donde la primera va a contener información general del estudio. En el prototipo se muestran los datos más significativos, como el título del estudio, una descripción, el propósito del estudio y el universo al cual se le va a realizar el estudio.



Archivo Edición Ver Ayuda

Estudio 1 X Estudio 2 X +

buscar pregunta 🔍

Pregunta Respuesta

pregunta 1

pregunta 2

Nombre

Texto

Instruccion

Editar Pregunta Agregar Pregunta

Fig. 12 - Prototipo Preguntas

Archivo Edición Ver Ayuda

Estudio 1 X Estudio 2 X +

buscar pregunta

Pregunta Respuesta

Información General

Preguntas

Instrumentos

pregunta 1
pregunta 2

Opciones de Respuestas

Respuesta	Descripción

Tipo

Seleccionar ▼

Agregar Tipo

Respuesta	Tipo
Edad	N Numérico

Editar Pregunta Agregar Pregunta

Fig. 13 - Prototipo Respuestas de Preguntas

En las figuras 12 y 13 se muestran los prototipos diseñados para el alta, baja y modificación de preguntas. En este caso, el usuario podrá dar de alta una nueva pregunta ingresando Título, Texto e Instrucciones de la misma, como también la estructura de la respuesta. La pregunta creada quedará disponible para su posterior uso en un instrumento.

Archivo Edición Ver Ayuda

Estudio 1 X Estudio 2 X +

buscar pregunta

Información General

Preguntas

Instrumentos

pregunta 1
pregunta 2

Pregunta1-
Pregunta2-
 ▽ Si Respuesta2 < 10
 Pregunta 2.1-
 Pregunta 2.2-

Nueva Pregunta
Nueva Sentencia
Agregar Condicional
Agregar For Loop
Agregar While Loop
Agregar Until Loop

Texto pregunta: Pregunta 2.2

Variable: A8

Fig. 14 - Prototipo Instrumento

Por último, en la figura 14 se muestra el prototipo del instrumento. Se busca un diseño donde la creación del instrumento no resulte complejo para el usuario. En este caso se podrá seleccionar preguntas de una lista para incluir en el instrumento y agregar demás estructuras de control necesarias para determinar el flujo del mismo. A medida que se va armando el cuestionario, se podrá tener una visión esquemática del mismo.

6.4 CASOS DE USO

Fueron diseñados los principales casos de uso del sistema, éstos fueron los siguientes:

- Importación DDI
- Exportación DDI
- Exportación RTF
- Nueva Variable
- Nueva Pregunta
- Creación de instrumento

Dichos casos de uso se pueden ver en el [Anexo 1](#).

7 IMPLEMENTACIÓN

En este capítulo se describe como fue implementada la aplicación analizada y diseñada en los capítulos anteriores.

La solución es una aplicación de escritorio para Windows.

7.1 HERRAMIENTAS UTILIZADAS

Para el desarrollo se utilizaron las siguientes herramientas:

- Visual Studio 2012: entorno de desarrollo utilizado con c#
- Assembla: repositorio
- Tortoise: manejo del repositorio
- WPF: Framework para interfaz de usuario.
- MVVM Light Toolkit: Librería para patrón MVVM

7.2 PATRÓN MVVM

El patrón MVVM (Model-View-ViewModel) es un patrón de diseño para desacoplar todo código que tenga que ver con la interfaz de usuario, del resto del código. [31]

Cuando se utiliza dicho patrón, la aplicación queda dividida en las siguientes capas:

- **Modelo:** Capa que contiene todo lo relacionada a la lógica de negocio de la aplicación, es totalmente independiente de la interfaz de usuario.
- **Vista:** Capa que contiene todo el código de interfaz de usuario. Los datos y comandos que se obtienen del modelo de vista son declarados como enlaces.
- **Modelo de Vista:** Cada vista se corresponde con un modelo de vista. Este contiene todos los destinos de los enlaces declarados en su correspondiente vista.

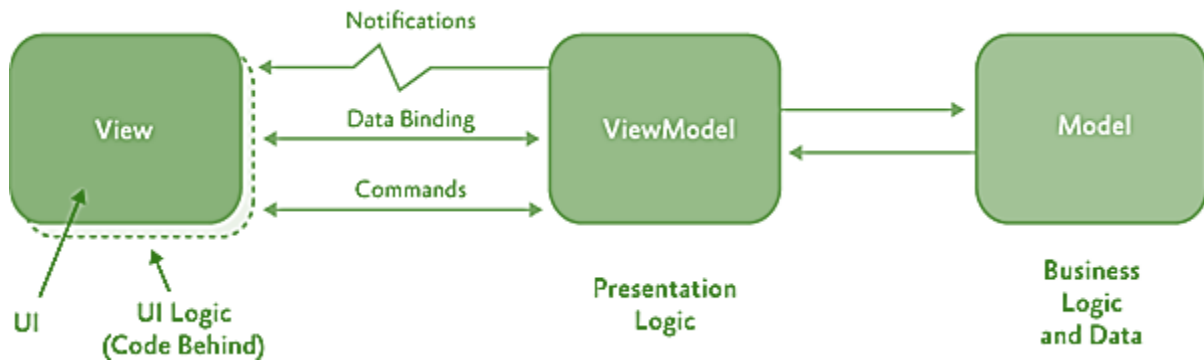


Fig. 15 – MVVM [32]

MVVM Light Toolkit es una librería, la cual fue utilizada en el desarrollo, la cual ayuda a la aplicación del patrón MVVM. [33]

7.3 MÓDULOS DE LA APLICACIÓN

La solución se divide en tres módulos. El primer módulo contiene una librería de clases para el manejo de los archivos XML para la versión DDI2. El segundo módulo es análogo al primero pero para la versión DDI3. Por último, el tercer módulo contiene las tres capas mencionadas en el patrón MVVM.

En la Fig. 16 se presenta un diagrama de los módulos y capas de la aplicación:

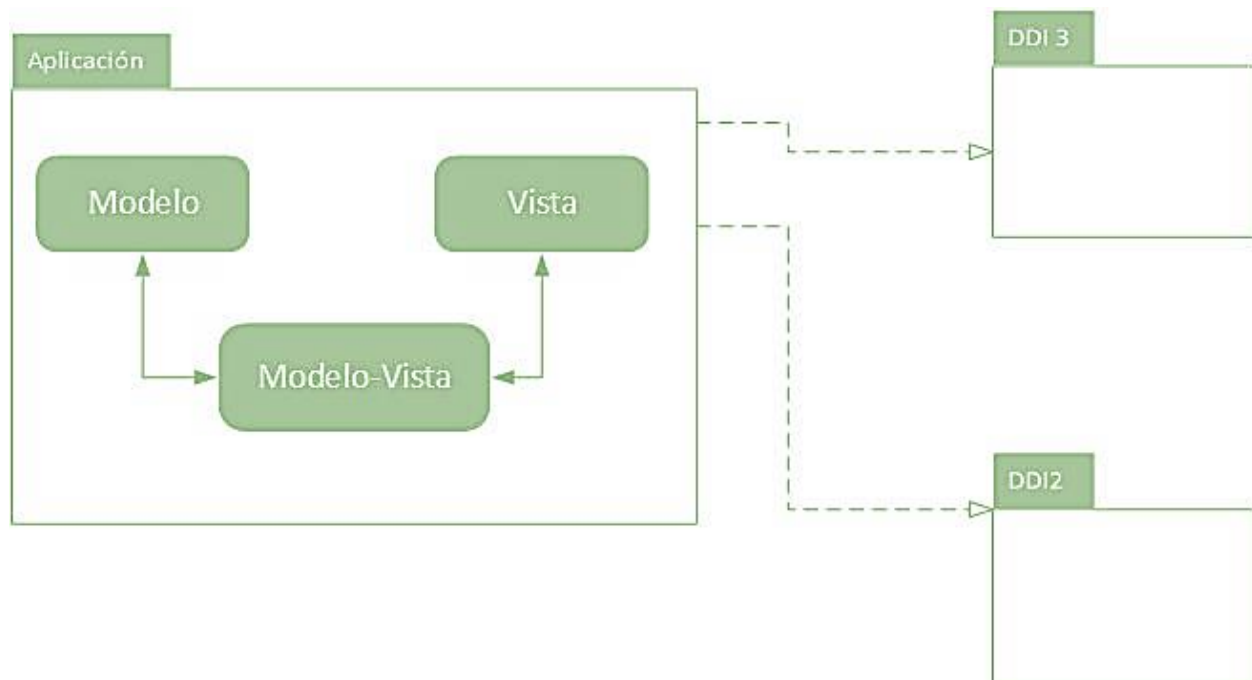


Fig. 16 – Módulos y Capas

7.4 PRINCIPALES FUNCIONALIDADES

7.4.1 IMPORTACIÓN/EXPORTACIÓN

Para la exportación e importación de las dos versiones de DDI se implementaron dos librerías, una para cada versión. La librería para la versión DDI2 contiene 176 clases las cuales se corresponden con la definición de las etiquetas definidas en la estructura del XML.

Como se vio en la definición del estándar, en la versión DDI3 se modularizó la estructura de los XML. Para cada uno de los esquemas se creó una carpeta, donde cada una contiene las clases correspondientes a las etiquetas definidas en el esquema XML.

En cada una de las librerías se encuentran las funciones que serializan y deserializan los archivos XML.

En la Fig. 17 se muestra a la izquierda la clase en c#, correspondiente con la instancia DDI y a la derecha la especificación de la instancia DDI en el esquema XML.

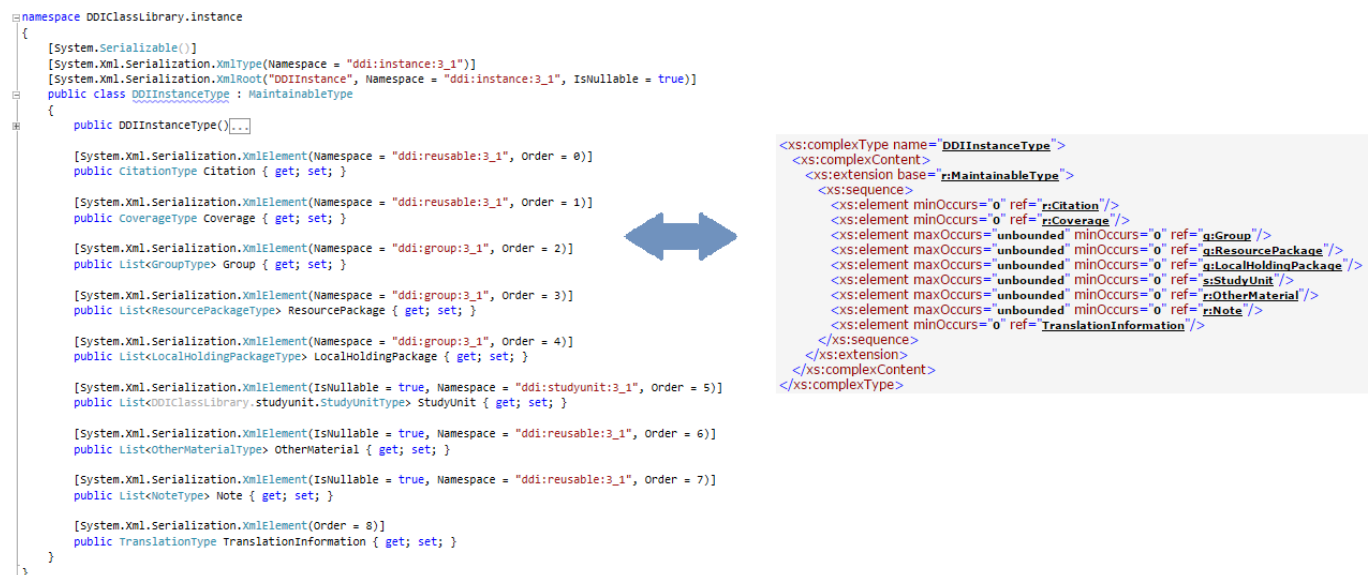


Fig. 17 - XML - Clases

7.4.2 CONVERSIÓN DE DDI2 A DDI3

Bajo la especificación de la versión DDI 3, en la página web del estándar, existe una sección donde se explica la correspondencia de las etiquetas XML entre las distintas versiones. [34]

Utilizando la información anterior se realizó un mapeo de cada uno de los campos. El mapeo de una versión a otra no es para nada intuitiva ni trivial, lo cual se requirió de mucho cuidado y esfuerzo para que la correspondencia tenga sentido.

7.4.3 CREACIÓN DE INSTRUMENTO

Para la creación del instrumento se implementó un editor con funcionalidades de “drag and drop” para agregar, mover y eliminar distintos tipos de elementos. En el editor se muestra, en una estructura de árbol, como va quedando el flujo del cuestionario.

A partir de la versión 3.2 de DDI se definió una instancia de tipo fragmento, la cual define una estructura XML para elementos particulares. Esto hace que sea posible escribir un archivo XML en versión DDI reducido, que contenga únicamente algunas secciones. Utilizando la estructura de fragmentos se implementó en la aplicación la posibilidad de exportar únicamente el instrumento con todos los elementos definidos dentro de él.

En la Fig. 18 se muestra un instrumento definido en la aplicación y el fragmento exportado:

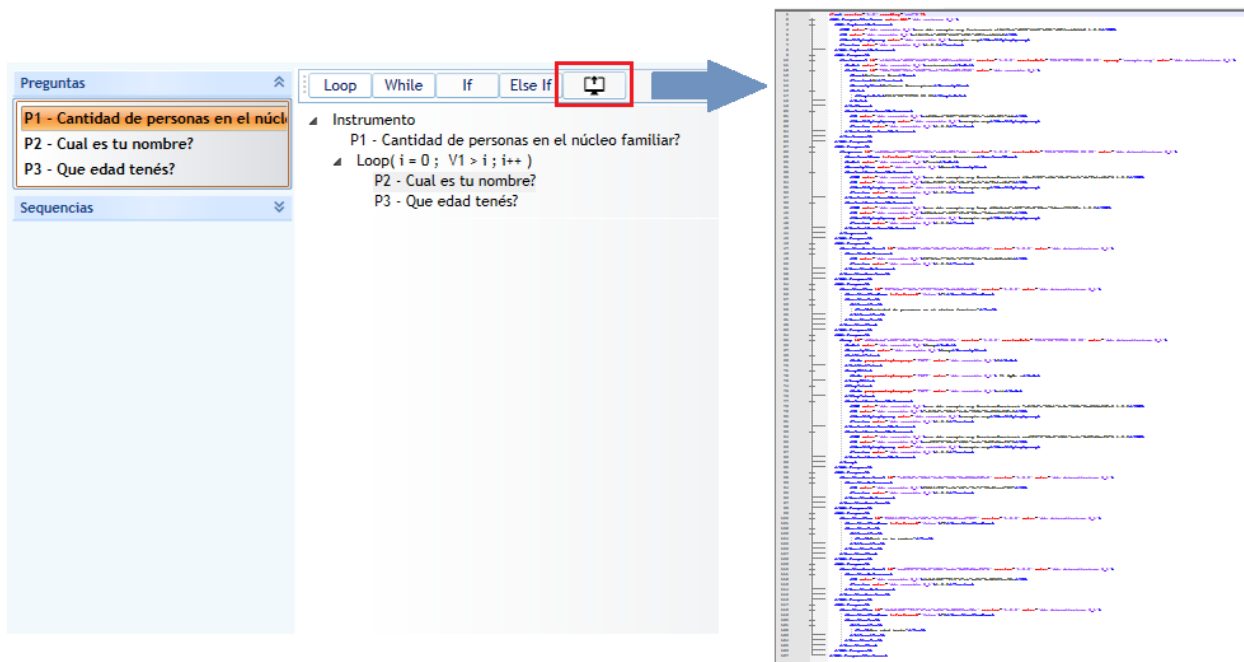


Fig. 18 - Exportación de fragmento

Se puede observar en la imagen anterior la complejidad que tiene escribir el archivo XML, si se tuviera que escribir a mano, teniendo en cuenta que el instrumento que se definió es muy simple.

7.4.4 EXPORTACIÓN A RTF

El instrumento luego de creado se puede exportar en formato RTF. El documento queda guardado en un formato estándar, para que el usuario luego lo pueda editar a su gusto.

A la derecha de la figura 19 se muestra el archivo en formato RTF del instrumento de ejemplo:

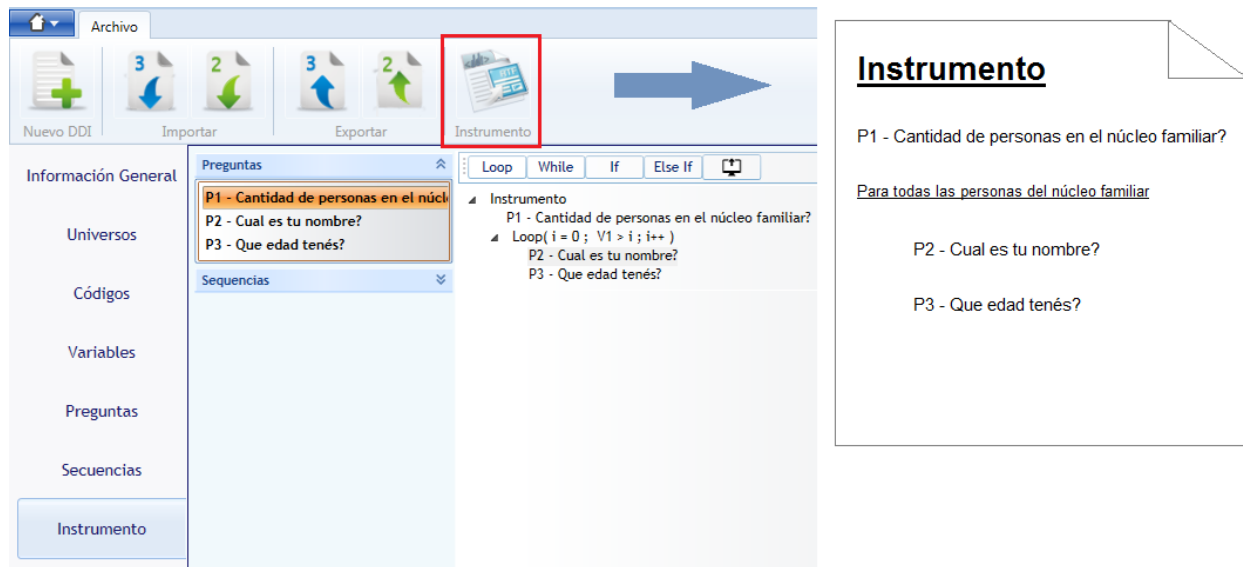


Fig. 19 - Exportación RTF

7.5 MANUAL DE USUARIO

También se escribió un manual de usuario que ayuda a comprender la herramienta, el cual se puede acceder desde el botón “Ayuda”. [Anexo 2](#)

8 EJEMPLOS

A fin de visualizar la utilidad de la herramienta se desean describir ciertos ejemplos donde se reflejan las principales funcionalidades. Los mismos fueron utilizados también para la validación del producto y serán mostrados en la presentación.

8.1 CODEBOOK

Ya que el Instituto Nacional de Estadística del Uruguay utiliza el estándar DDI en su versión DDI 2, se obtuvo la Encuesta Continua de Hogares (ECH) de Uruguay, del año 2010. Dicha encuesta es la última encuesta extendida, cuyos metadatos están publicados por el INE.

La ECH es una encuesta que ofrece información anualmente sobre las características demográficas básicas de la población y de los hogares que componen (tipología y tamaño) así como de las viviendas que habitan.

La herramienta desarrollada permite importar el Codebook que contiene la ECH del 2010. Se pueden apreciar dentro de la aplicación y de una forma más amigable todos los datos que contienen el archivo, información general, universos, códigos, variables y preguntas.

Dado que la versión DDI 2 no tiene definido los conceptos de secuencia e instrumento, a priori estas secciones en la aplicación aparecerán vacías. Sin embargo utilizando los datos importados del Codebook se podrán construir distintas secuencias y el instrumento para luego exportarlo en formato RTF.

Finalmente, este Codebook se podrá exportar a formato DDI3.

Funcionarios del INE nos han expresado que para ellos es muy útil tener una herramienta donde pueda reutilizar información de encuestas anteriores, dado que hoy en día la herramienta que utilizan no se los permite.

8.2 DDI 3

Para la versión DDI 3 se trabajó con el censo 2010 de los Estados Unidos, es uno de los cuestionarios de investigación más reconocidos de dicho país. La Oficina del Censo de los Estados Unidos (United States Census Bureau) forma parte del Departamento de Comercio de los Estados Unidos, es el organismo gubernamental que se encarga del censo en dicho país. Su objetivo es dar una fuente de datos de calidad sobre la población del país y su economía.

[35]

El censo contiene únicamente diez preguntas, de acuerdo a US Census Bureau, cada pregunta ayuda a determinar cómo más de 400 millones de personas están distribuidas a lo largo de todo el país. [36]

Con únicamente diez preguntas, el XML de metadatos contiene más de 2500 líneas.

Se podrá importar este archivo con la aplicación y visualizar/modificar en el editor el instrumento que contiene dicho archivo.

Como vimos anteriormente, la versión DDI 3.2 tiene definido una estructura XML, la cual permite escribir una instancia válida de DDI 3.2 incluyendo solamente algún fragmento, sin tener que exportar la información completa. Utilizando esta idea, la herramienta permite exportar un fragmento que contiene únicamente el instrumento, sin la necesidad de exportar el DDI entero.

9 VALIDACIÓN INSTITUTO NACIONAL DE ESTADÍSTICA

El Instituto Nacional de Estadística es el principal impulsor del estándar DDI del país, no sólo lo aplica para las encuestas que el Instituto realiza, sino que también brinda asistencia a otros entes que desean comenzar a trabajar con el estándar.

El INE utiliza el estándar DDI en el marco del Programa Acelerado de Datos (PAD), la principal motivación para su utilización radica en que mediante un estándar se hace posible la comparación de datos a nivel mundial. Además, al ser un estándar impulsado por el Banco Mundial y por la OCDE, lo utilizan las principales organizaciones de estadística del mundo. Desde que se comenzó a usar, en el año 2009, hasta la fecha se utiliza la versión 2.x (Codebook).

Para la creación de los metadatos, el INE utiliza Nesstar Publisher (herramienta impulsada por la IHSN), para la publicación utiliza el Catálogo ANDa de acuerdo a las recomendaciones realizadas por el IHSN para la estandarización de las publicaciones.

Catálogo de metadatos y microdatos

El catálogo de metadatos y microdatos ANDa (Archivo Nacional de Datos) es un sistema de catalogación de información acerca de las operaciones estadísticas que produce el Instituto Nacional de Estadística (INE).

Su interfaz web facilita a los usuarios la navegación, búsqueda y descarga de microdatos y metadatos de censos, encuestas y registros administrativos.

La información de metadatos y documentos de referencia se presenta en un formato estandarizado ampliamente utilizado internacionalmente. El estándar de metadatos DDI (Data Documentation Initiative) permite que la información sea presentada de la misma forma para los distintos proyectos, ordena los documentos metodológicos y conceptuales, y facilita la búsqueda y comparación de variables de las diferentes operaciones estadísticas, entre otras posibilidades.

El Catálogo ANDa permite:

- buscar los metadatos de censos, encuestas y registros administrativos según un período en particular, para un determinado tema o por palabras clave, a través de los requerimientos de cada usuario
- visualizar de forma amigable la información de metadatos de manera estandarizada, con un mismo diseño, estilo y estructura del contenido de las diversas operaciones estadísticas
- buscar variables de diferentes proyectos estadísticos de acuerdo a criterios establecidos por los usuarios
- comparar las variables encontradas en los distintos proyectos por medio de una interfaz interactiva que permite al usuario una lectura más ágil
- simplificar el uso de los microdatos de las operaciones estadísticas, a través de información completa de cada una de las variables que los componen, facilitando a los usuarios su explotación
- comprender cabalmente los proyectos estadísticos, desde su planificación, objetivos, universo de estudio, metodología, captura y procesamiento de datos, hasta las definiciones de cada variable o indicador publicado
- evaluar la calidad de los datos y conocer las restricciones en cuanto al uso y las inferencias que se pueden hacer sobre los mismos

En www.ine.gub.uy/anda se puede visitar el catálogo de metadatos y microdatos ANDa.

Fig. 19 – Catálogo ANDa [37]

Tuvimos la oportunidad de tener dos entrevistas con Lucía Pérez de la Unidad de Gestión de la Calidad del INE.

El objetivo del primer encuentro fue validar el producto a construir, su diseño y funcionalidades. Además se pretendía obtener información de cómo el Instituto Nacional de Estadística se maneja con el estándar, qué herramientas utiliza y cuáles son sus procedimientos en la creación/publicación de metadatos. El primer encuentro

fue realizado antes de comenzar la etapa de implementación. En esta instancia fueron presentadas las funcionalidades del sistema con sus respectivos prototipos y los mismos fueron validados satisfactoriamente.

Para el segundo encuentro el principal objetivo fue la validación del producto terminado. El mismo fue realizado al terminar la etapa de implementación, donde se presentó la aplicación terminada con un caso de estudio real del INE, Encuesta Continua de Hogares del 2006.

Fue destacada la sencillez de la aplicación en contraposición con la complejidad de la actual herramienta utilizada, Nesstar. La funcionalidad más importante a mencionar es la capacidad de comenzar un estudio a partir de un estudio anterior a través de la importación de un XML anterior. Actualmente el INE debe digitar encuesta a encuesta todos los datos nuevamente, a pesar que generalidades del estudio, preguntas y/o variables se repitan entre estudios diferentes. Por tal motivo la documentación de las encuestas no resulta un trabajo sencillo ni tampoco ágil. Esto ocurre por ejemplo con la ECH.

Ante las contras del aplicativo Nesstar, consultamos por qué no se utiliza alguna otra herramienta para la manipulación de metadatos. El motivo es que Nesstar, además de ser la aplicación recomendada por la IHSN, es gratuita, mientras que otras herramientas como Colectica tienen licencias de muy alto costo.

10 DESARROLLO DEL PROYECTO

10.1 CRONOGRAMAS

En la figura 20 se muestra el cronograma del proyecto, se pueden observar los tiempos de las diferentes etapas, los tiempos generales del proyecto y algunos hitos importantes del mismo.

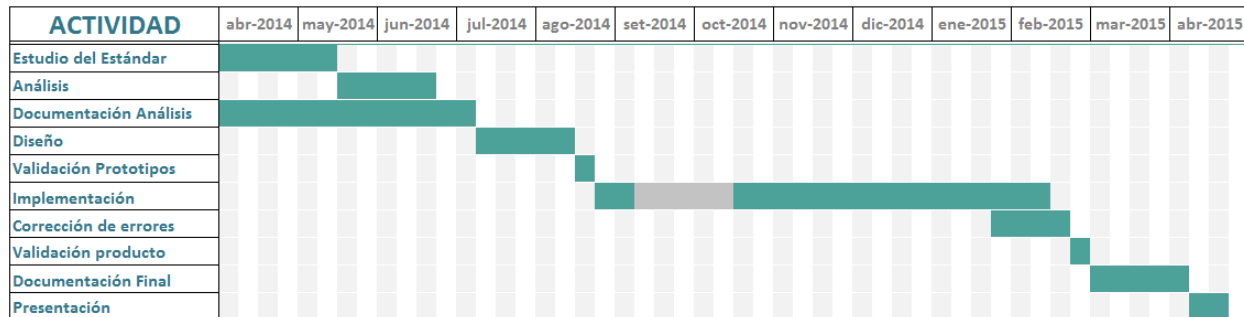


Fig. 20 - Cronograma

10.2 ETAPAS

El proyecto fue dividido en seis etapas:

Estudio del estándar

Dada la complejidad del estándar, fue necesario dedicar al comienzo del proyecto un buen tiempo para su entendimiento.

Algunas de las preguntas planteadas para comenzar la investigación fueron:

- ¿Qué es? ¿Para qué sirve?
- ¿Para qué se usa? ¿Dónde se usa?
- ¿Quién lo creo? ¿Cuál es la motivación que lleva a la creación?
- ¿Qué más hay?
- ¿Cuál es la estructura?
- ¿Qué se puede lograr?

Se analizaron las diferentes versiones del estándar y se realizó un estudio comparativo de las mismas. Además de una breve investigación para determinar qué versiones son utilizadas por las organizaciones.

Surge la necesidad de analizar otros estándares, diferencias y similitudes con DDI, aplicación de los mismos, etc.

Análisis

Luego de conocer más en profundidad el estándar, comenzamos a centrarnos en los objetivos del proyecto.

En esta etapa se realizó un estudio detallado de XML, para de esta forma poder conocer la información con la que contamos y de qué forma el mismo está estructurado. También se realizó un relevamiento de las diferentes herramientas existentes actualmente que manejan el estándar, cómo lo usan y quien las utiliza.

En esta etapa se realizó la definición y documentación de requerimientos, la creación de casos de uso y los diagramas de secuencia.

Diseño

Se realizó el diseño de la aplicación y su arquitectura. La principal interrogante de esta etapa fue si realmente era necesario una base de datos. Luego del análisis de varios factores, se decidió no utilizar una base de datos para la persistencia de los XML. Para tomar la decisión se tomó en cuenta que la aplicación ya cuenta con la exportación del XML en edición y con la futura importación del mismo para continuar el trabajo, en este caso la persistencia en base de datos sólo agrega complejidad al tratamiento, sin agregar ninguna utilidad.

Se finalizó la etapa de diseño con la creación de prototipos y la validación de los mismos por parte del INE.

Implementación

Cómo se puede observar en el cronograma (Fig.20) fue la etapa más extensa del proyecto, teniendo una interrupción de cinco semanas por cuestiones laborales.

Podemos decir que la etapa de desarrollo fue una etapa iterativa incremental de 5 pasos.

En primera instancia se atacaron los requerimientos RF1, RF6 y RF7 (Importación). Dada la complejidad de los XML de ambas versiones del estándar resultaron ser estos los requerimientos de mayor riesgo y complejidad.

En una segunda etapa, se implementó el requerimiento RF5 (Exportación), para cerrar de esta forma un ciclo semi completo. Luego de esta etapa el sistema permitía la importación de un archivo DDI3 y la exportación del mismo.

En una tercera iteración, fueron desarrollados los requerimientos relacionados a la creación de nuevos datos RF2 (Nuevas preguntas) y RF3 (Creación de formularios).

Luego de haber pasado un buen tiempo familiarizándonos con los XML de ambas versiones del estándar, nos encontrábamos en condiciones de implementar la conversión de DDI2 a DDI3. Es un requerimiento crítico e importante, pero es necesario una formación previa para su desarrollo, por ese motivo no fue implementado en etapas más tempranas del desarrollo a pesar de estar dentro de las funcionalidades más importantes. En esta misma etapa fue implementada la exportación de DDI2. (RF8)

En la etapa final del desarrollo fue implementada la generación de cuestionarios en RTF. (RF4)

Verificación y validación

En la sección 9, fueron presentados algunos ejemplos, los cuales fueron también utilizados para las pruebas de la aplicación y para verificar el correcto desempeño de la misma.

Luego de pruebas varias y validación interna, se realizó la validación del producto con el INE.

Documentación

Esta etapa fue realizada a lo largo de todo el proyecto.

11 CONCLUSIONES

Al comenzar con la etapa de investigación del proyecto notamos que había mucha información relacionada con el estándar DDI y en gran mayoría importante para la investigación. Existen muchos temas relacionados con la estandarización de los metadatos para encuestas, de los cuales hay mucha información de relevancia para este proyecto. Sin embargo, específicamente del estándar DDI, no existen grandes investigaciones ni documentación por fuera de la definición del mismo. Además se trabajó con la última versión del estándar, la cual fue liberada en Marzo de 2014, la reciente liberación de la última versión contribuye también a la escases de trabajos e investigaciones sobre la misma.

A pesar de que los estándares de datos, el manejo de metadatos y microdatos para encuestas, son temas en auge en las redes e investigaciones, además de existir mucha información desperdigada en grandes organizaciones, no existen actualmente muchos sistemas que manejen DDI en todas sus formas y/o versiones y mucho menos sistemas gratuitos.

Una de las motivaciones del proyecto fue que al comenzar la investigación del estándar DDI, no existían herramientas libres para poder definir un instrumento, ni herramientas que manejaran las dos versiones de DDI. Creemos que esto ocurre dado que la nueva versión del estándar recién se está afianzando en algunas organizaciones y resulta realmente compleja la migración de una versión a otra. Las organizaciones han comenzado a dar pequeños pasos en pro de este cambio.

Consideramos que la herramienta desarrollada puede ser de mucha utilidad para organizaciones como el INE de Uruguay, complementando las herramientas ya utilizadas. Una de las grandes mejoras que puede aportar es la reutilización de datos ingresados en estudios anteriores. Además, los archivos XML escritos bajo el estándar DDI son muy extensos y complejos de entender, ésta es otra de las razones por las cuales resultaría útil la aplicación. Es importante resaltar que se ratificó por medio de entrevistas con el INE que este tipo de herramientas serían de utilidad para el instituto.

Dado que las organizaciones se encuentran dando algunos pasos para la estandarización de los metadatos para las encuestas, creemos que se está abriendo actualmente una interesante línea de investigación en este tema.

12 TRABAJOS FUTUROS

Dado que el tema del proyecto es muy amplio y se puede extender mucho más, se tuvo que limitar el alcance del mismo, dejando de lado algunas funcionalidades que son muy interesantes para la aplicación.

A continuación se describen varios trabajos a futuro que podrían realizarse sobre el sistema, que consideramos que serían de mucha utilidad.

12.1 REPOSITORIO

Dado que la versión DDI 3 se basa mucho en la reutilización, versionado y mantenibilidad de objetos definidos dentro del estándar, existe la posibilidad de desarrollar un repositorio donde se puedan administrar “objetos” escritos en DDI. El repositorio podría ser un sistema de almacenamiento centralizado para la gestión de los recursos de datos de DDI, lo que permitirá flujos de trabajo basados en DDI colaborativos y podría proporcionar control de versión automática.

12.2 EXPORTACIÓN A HTML

El sistema desarrollado, exporta el instrumento en formato RTF, sería interesante y de mucha utilidad realizar la exportación en formato HTML. De esta manera otro sistema podría consumir el archivo exportado y realizar la encuesta vía WEB.

12.3 APLICACIÓN WEB

Dado que la aplicación en un principio fue pensada para uso local, se realizó una aplicación de escritorio para disminuir las dificultades. Perfectamente podría ser una aplicación WEB donde se permita a cada usuario mantener sus propios trabajos y con la ayuda del repositorio poder compartir datos. Esto facilita el acceso a los usuarios de la aplicación y mejora el trabajo colaborativo entre usuarios. Para este desarrollo es imprescindible trabajar sobre la persistencia de los datos de cada uno de los usuarios.

12.4 VALIDACIÓN MICRODATOS

Actualmente tenemos los microdatos por un lado y por otro los metadatos. Resulta de interés contar con algún tipo de software que nos permita realizar la validación de integridad de los microdatos de acuerdo a los metadatos.

12.5 PUBLICACIÓN DE API

Por último, existe la posibilidad de exponer los servicios desarrollados para la aplicación, de manera que otros sistemas puedan consumirlos. Esta API debería especificar cada servicio expuesto, de manera que el usuario externo

pueda consumirlo correctamente. Un ejemplo, podría ser la funcionalidad de validar un archivo escrito en versión DDI 3 o la conversión de la versión DDI 3 a la versión DDI 2.

13 BIBLIOGRAFÍA

- [1] "IPUMS-I: FAQ." [Online]. Available: <https://international.ipums.org/international-action/faq#ques7>. [Accessed: 22-Jul-2014].
- [2] "Microdata (statistics)," *Wikipedia, the free encyclopedia*. 16-Jul-2014.
- [3] "Metadata." [Online]. Available: <http://www.sedic.es/autoformacion/metadatos/tema1.htm>.
- [4] "Paris 21." [Online]. Available: <http://www.paris21.org/>.
- [5] "IHSN NADA." [Online]. Available: <http://www.ihsn.org/home/software/nada>.
- [6] "IHSN Metadata editor." [Online]. Available: <http://www.ihsn.org/home/software/ddi-metadata-editor>.
- [7] "PAD." [Online]. Available: <http://adp.ihsn.org/node/204>.
- [8] "INE Uruguay." [Online]. Available: <http://www.ine.gub.uy/>.
- [9] "Estudio CAPI vs PAPI." [Online]. Available: <http://www.jos.nu/Articles/abstract.asp?article=262233>.
- [10] "Capítulo 4: XML Schema." [Online]. Available: <http://bibing.us.es/proyectos/abreproy/11387/direccion/VOLUMEN+1%252F>.
- [11] "DDI-Codebook 2.5." [Online]. Available: <http://www.ddialliance.org/Specification/DDI-Codebook/2.5/>.
- [12] "URN Definition." [Online]. Available: <https://www.ietf.org/rfc/rfc3406.txt>.
- [13] "MARC Standard." [Online]. Available: <http://www.loc.gov/marc/>.
- [14] "GSMI." [Online]. Available: <http://www1.unece.org/stat/platform/display/gsim/Generic+Statistical+Information+Model>.
- [15] "iso-iec 11179." [Online]. Available: <https://www.iso.org/obp/ui/#iso:std:iso-iec:11179:-1:ed-2:v1:en>.
- [16] "ISO 19118." [Online]. Available: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44212.
- [17] "SDMX – Statistical Data and Metadata Exchange." .
- [18] "METS." [Online]. Available: <http://www.loc.gov/standards/mets/METSOverview.v2.html>.
- [19] "PREMIS." [Online]. Available: <http://www.loc.gov/standards/premis/>.
- [20] UNECE Secretariat, "Generic Statistical Business Process Model," Abril-2009. [Online]. Available: <http://www1.unece.org/stat/platform/download/attachments/8683538/GSBPM+Final.pdf>.
- [21] W. Thomas, "Course on DDI 3: Putting DDI to Work for You," 03-Dec-2009. [Online]. Available: http://www.iza.org/conference_files/eddi09/ppt/thomas_wendy_course.pdf.
- [22] K. Dinkelmann, "DDI Version 3 and Instrument Documentation," 26-May-2006. .
- [23] "IBM - Statistical analysis software package - SPSS Statistics," 11-Apr-2015. [Online]. Available: <http://www-03.ibm.com/software/products/en/spss-statistics>. [Accessed: 11-Apr-2015].
- [24] "Colectica | CrunchBase." [Online]. Available: <http://www.crunchbase.com/organization/colectica>. [Accessed: 26-Jul-2014].
- [25] "Colectica Express - DDI Metadata and Survey Design Software." [Online]. Available: <http://www.colectica.com/software/express>. [Accessed: 26-Jul-2014].
- [26] "Blaise | Survey Software for Professionals." [Online]. Available: <http://www.blaise.com/>. [Accessed: 04-Aug-2014].
- [27] "Nesstar - Publish Data on the Web." [Online]. Available: <http://www.nesstar.com/software/publisher.html>. [Accessed: 30-Jul-2014].
- [28] "IHSN Software." [Online]. Available: <http://www.ihsn.org/home/software/ddi-metadata-editor>.
- [29] "Nesstar - Publish Data on the Web." [Online]. Available: <http://www.nesstar.com/software/server.html>. [Accessed: 30-Jul-2014].
- [30] "Nesstar - Publish Data on the Web." [Online]. Available: <http://www.nesstar.com/software/webview.html>. [Accessed: 30-Jul-2014].
- [31] "MVVM." [Online]. Available: <https://msdn.microsoft.com/en-us/magazine/dd419663.aspx>.
- [32] "Patrón MVVM." [Online]. Available: <https://msdn.microsoft.com/en-us/library/gg405484%28v=pandp.40%29.aspx>.
- [33] "MVVM Light Toolkit," *CodePlex*. [Online]. Available: <https://mvvmlight.codeplex.com/Wikipage?ProjectName=mvvmlight>. [Accessed: 24-Mar-2015].

- [34] "DDI2 to DDI3." [Online]. Available: <http://www.ddialliance.org/Specification/DDI-Lifecycle/3.0/#3.0convert>.
- [35] C. for N. M. and Promotions and U. C. B. Public Information Staff, "US Census Bureau 2010 Census." [Online]. Available: <http://www.census.gov/2010census/about/interactive-form.php>. [Accessed: 08-Apr-2015].
- [36] "Census USA." [Online]. Available: <http://www.census.gov/2010census/about/interactive-form.php>.
- [37] "Catalogo ANDa." [Online]. Available: <http://www.ine.gub.uy/vari0s/noticias%202014/Folleto%20SEN%20Metadatos.pdf>.

ANEXO 1 – CASOS DE USO

Nombre: Importación DDI

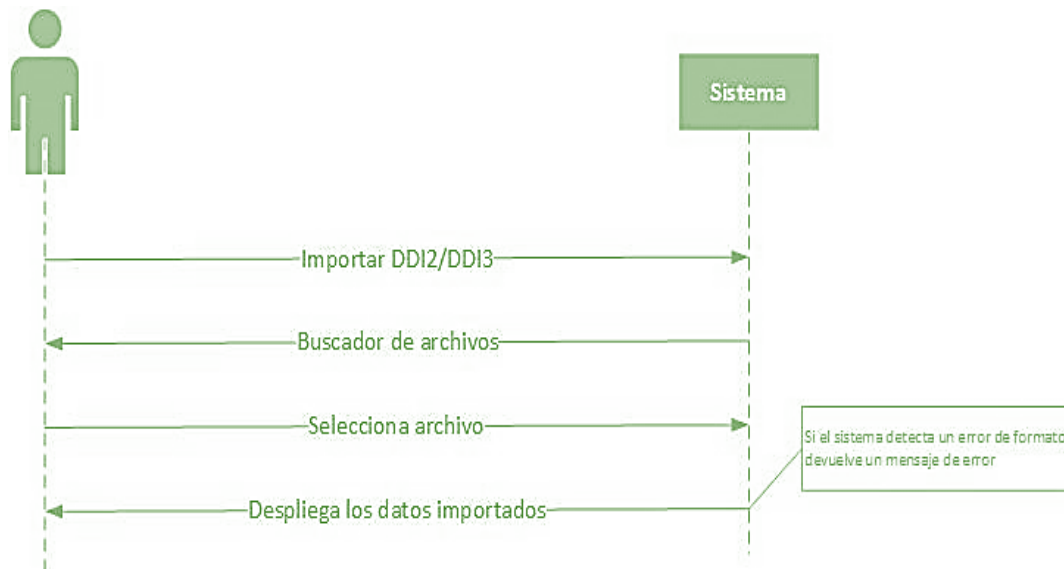
Objetivo: Importar un archivo en formato XML escrito bajo el estándar DDI.

Actores: Usuario del sistema

Precondiciones: Tener un archivo XML para importar.

Descripción: Este caso de uso comienza cuando el usuario desea importar un archivo XML al sistema. El usuario selecciona el archivo y el sistema despliega los datos importados.

Diagrama:



Flujo Principal:

1. El usuario le indica al sistema que desea importar un archivo en formato XML escrito bajo la versión 3 del estándar DDI
2. El sistema despliega un buscador para que el usuario seleccione el archivo
3. El usuario busca el archivo escrito bajo la versión DDI3 y lo selecciona para ser importado
4. El sistema valida el archivo y muestra los datos importados en las distintas secciones
5. Fin del Caso de Uso

Flujos Alternativos:

3. A. **El usuario desea importar un archivo XML escrito en la versión DDI 2.**

3. A.1. El usuario selecciona desde el menú la opción de importar la versión DDI 2

3. A.2. Vuelve al punto 2

4. A. **El archivo no tiene el formato adecuado**

4. A.1 El sistema detecta que el archivo XML no tiene un formato adecuado y devuelve el siguiente mensaje de error: “El archivo tiene un formato incorrecto.”

4. A.2 Fin del Caso de Uso

Nombre: Exportación DDI

Objetivo: Exportar un archivo en formato XML escrito bajo el estándar DDI.

Actores: Usuario del sistema

Precondiciones: El usuario ya ingresó en la aplicación e ingresó la información deseada.

Descripción: Este caso de uso comienza cuando el usuario desea exportar la información ingresada en el sistema a un archivo XML escrito bajo el estándar DDI.

Diagrama:



Flujo Principal:

1. El usuario le indica al sistema que desea exportar un archivo en formato XML escrito bajo la versión 3 del estándar DDI
2. El sistema despliega un buscador para que el usuario seleccione la ruta del archivo

3. El usuario selecciona la ruta donde desea dejar el archivo
4. El sistema muestra un mensaje de éxito, indicando que la exportación se realizó correctamente
5. Fin del Caso de Uso

Flujos Alternativos:

1. A. El usuario desea exportar la información ingresada en un archivo XML escrito en la versión DDI 2.

1. A.1. El usuario selecciona desde el menú la opción de exportar la versión DDI 2
1. A.2. Vuelve al punto 2

4. A. Errores en la información ingresada

4. A.1 El sistema detecta algún error en la información ingresada y despliega un mensaje de error indicando el mismo.
 4. A.2 Fin del Caso de Uso
-

Nombre: Exportación RTF

Objetivo: Exportar un archivo en formato RTF que contiene la información del instrumento.

Actores: Usuario del sistema

Precondiciones: El usuario ya ingresó en la aplicación y creó un instrumento.

Descripción: Este caso de uso comienza cuando el usuario desea exportar la información del instrumento en un archivo en formato RTF.

Diagrama:



Flujo Principal:

1. El usuario le indica al sistema que desea exportar en formato RTF el instrumento ingresado
2. El sistema despliega un buscador para que el usuario seleccione la ruta del archivo
3. El usuario selecciona la ruta donde desea dejar el archivo
4. El sistema muestra un mensaje de éxito, indicando que la exportación se realizó correctamente
5. Fin del Caso de Uso

Flujos Alternativos:

4. A. Errores en la información ingresada

4. A.1 El sistema detecta algún error en la información ingresada y despliega un mensaje de error indicando el mismo.
4. A.2 Fin del Caso de Uso

Nombre: Nueva Variable

Objetivo: Crear una nueva variable.

Actores: Usuario del sistema

Precondiciones: Existen códigos ingresados en el sistema.

Descripción: Este caso de uso comienza cuando el usuario desea crear una nueva variable en el sistema.

Diagrama:



Flujo Principal:

1. El usuario le indica al sistema que desea ingresar una nueva variable
2. El sistema deja editable los campos de las variables
3. El usuario ingresa Nombre, Texto, Descripción, Universo y Tipo de Variable “Numérico”
4. El sistema muestra tres nuevos campos, Tipo de Número, Valor Máximo y Valor Mínimo
5. El usuario indica los datos correspondientes a la nueva variable
6. Fin del Caso de Uso

Flujos Alternativos:

3. A. Tipo de Variable Texto

3. A.1 El usuario ingresa Nombre, Texto, Descripción, Universo y Tipo de Variable “Texto”
3. A.2 El sistema muestra tres nuevos campos, Largo Mínimo, Largo Máximo y Expresión Regular.
3. A.3 Vuelve al punto 5

3. B. Tipo de Variable Fecha

3. B.1 El usuario ingresa Nombre, Texto, Descripción, Universo y Tipo de Variable “Fecha”
3. B.2 El sistema muestra un campo nuevo, Tipo de Fecha.
3. B.3 Vuelve al punto 5

3. C. Tipo de Variable Código

- 3. C.1 El usuario ingresa Nombre, Texto, Descripción, Universo y Tipo de Variable “Código”
 - 3. C.2 El sistema muestra una lista de códigos para que el usuario pueda seleccionar entre los códigos que existen en el sistema.
 - 3. C.3 Vuelve al punto 5
-

Nombre: Nueva Pregunta

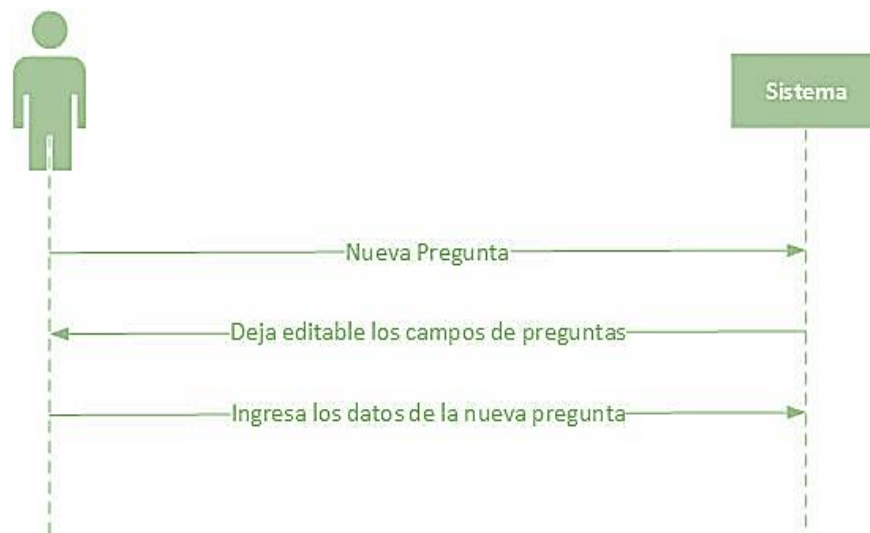
Objetivo: Crear una nueva pregunta.

Actores: Usuario del sistema

Precondiciones: Existen variables en el sistema.

Descripción: Este caso de uso comienza cuando el usuario desea crear una nueva pregunta en el sistema.

Diagrama:



Flujo Principal:

1. El usuario le indica al sistema que desea ingresar una nueva pregunta
 2. El sistema deja editable los campos de las preguntas
 3. El usuario ingresa Título, Texto, Instrucciones, y selecciona la variable a la cual quiere asociar con la pregunta
 4. Fin del Caso de Uso
-

Nombre: Creación de instrumento

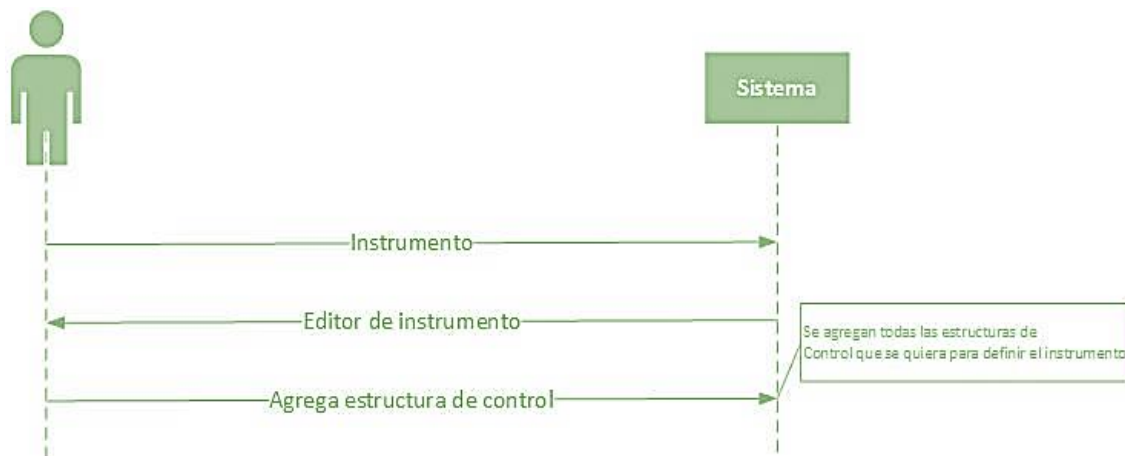
Objetivo: Crear un instrumento.

Actores: Usuario del sistema

Precondiciones: Existen preguntas y secuencias cargadas en el sistema.

Descripción: Este caso de uso comienza cuando el usuario desea crear un instrumento en el sistema, utilizando información ya ingresada.

Diagrama:



Flujo Principal:

1. El usuario le indica al sistema que desea definir el instrumento
2. El sistema muestra el editor del instrumento y dos listas, una que contiene las preguntas que existen en el sistema y otra que contiene las secuencias creadas.
3. El usuario agrega al instrumento un elemento de tipo “Pregunta”, arrastrando dicha pregunta desde la lista de preguntas.
4. El sistema despliega la pregunta en el editor del instrumento
5. Fin del Caso de Uso

Flujos Alternativos:

3. A. **Elemento “Secuencia”**

3. A.1 El usuario agrega al instrumento el elemento de tipo “Secuencia”, arrastrando dicha secuencia desde la lista de secuencias
3. A.2 El sistema despliega la secuencia en el editor del instrumento

- 3. A.3 Vuelve al punto 5
- 3. B. **Estructura de control “Loop”**
 - 3. B.1 El usuario agrega al instrumento el elemento de tipo “Loop”
 - 3. B.2 El sistema despliega la estructura de control en el editor del instrumento
 - 3. B.3 El usuario le indica al sistema que quiere edita la condición de dicha estructura de control
 - 3. B.4 El sistema despliega los campos “Etiqueta”, “Descripción”, “Condición”, “Valor Inicial” y “Paso”.
 - 3. B.5 El usuario indica los datos de la condición y presiona “Guardar”
 - 3. B.6 Vuelve al punto 5
- 3. C. **Estructura de control “While”**
 - 3. C.1 El usuario agrega al instrumento el elemento de tipo “While”
 - 3. C.2 El sistema despliega la estructura de control en el editor del instrumento
 - 3. C.3 El usuario le indica al sistema que quiere edita la condición de dicha estructura de control
 - 3. C.4 El sistema despliega los campos “Etiqueta”, “Descripción” y “Condición”.
 - 3. C.5 El usuario indica los datos de la condición y presiona “Guardar”
 - 3. C.6 Vuelve al punto 5
- 3. D. **Estructura de control “If”**
 - 3. D.1 El usuario agrega al instrumento el elemento de tipo “If”
 - 3. D.2 El sistema despliega la estructura de control en el editor del instrumento
 - 3. D.3 El usuario le indica al sistema que quiere edita la condición de dicha estructura de control
 - 3. D.4 El sistema despliega los campos “Etiqueta”, “Descripción” y “Condición”.
 - 3. D.5 El usuario indica los datos de la condición y presiona “Guardar”
 - 3. D.6 Vuelve al punto 5
- 3. E. **Estructura de control “Else If”**
 - 3. E.1 El usuario agrega al instrumento el elemento de tipo “Else If”
 - 3. E.2 El sistema despliega la estructura de control en el editor del instrumento
 - 3. E.3 El usuario le indica al sistema que quiere edita la condición de dicha estructura de control

- 3. E.4 El sistema despliega los campos “Etiqueta”, “Descripción” y “Condición”.
- 3. E.5 El usuario indica los datos de la condición y presiona “Guardar”
- 3. E.6 Vuelve al punto 5
- 5. A. **Sigue agregando estructuras de control**
 - 5. A.1 El usuario desea agregar otra estructura de control al instrumento
 - 5. A.2 Vuelve al punto 3

ANEXO 2 – MANUAL DE USUARIO

1. INTRODUCCIÓN

Esta aplicación tiene como objetivo el manejo y publicación de metadatos, utilizando el estándar DDI en sus distintas versiones.

El objetivo de este manual es ayudar y guiar al usuario a utilizar la herramienta, suponiendo que se tiene un conocimiento básico de las versiones de DDI, DDI3 – LifeCycle y DDI2 – Codebook.

2. INSTALACIÓN

Cuestionarios DDI es una aplicación de escritorio para ejecutar sobre el sistema operativo Windows. No se requiere de ningún requisito en particular.

3. SECCIONES DEL CUESTIONARIO

Cuando el usuario selecciona “Nuevo DDI” o importa un archivo XML a través del menú principal (el cual se ve en detalle más adelante), se muestra la información para ser ingresada o modificada. La aplicación está dividida en 7 secciones a través de tabuladores que se describen a continuación.

3.1 INFORMACIÓN GENERAL

A continuación se muestra una imagen de la aplicación cuando se selecciona “Información General”.

The screenshot shows the 'Información General' form. On the left, there is a sidebar with the following tabs: Información General (selected), Universos, Códigos, Variables, Preguntas, Secuencias, and Instrumento. The main content area contains the following fields:

- Título:** A text input field.
- Propósito:** A text input field.
- Descripción:** A large text input field.
- Creador:** A text input field.
- Fecha de Publicación:** A date picker with the text 'Select a date' and a calendar icon.
- Universo:** A dropdown menu.
- Cobertura Temporal:** A date picker with the text 'Select a date' and a calendar icon.
- Cobertura Geográfica:** A text input field.
- Cobertura Temática:** A text input field.

Dentro de “Información General” se encuentran los siguientes campos:

- Título: Título del estudio que se quiere llevar a cabo.
- Propósito: Se ingresa una descripción del propósito del estudio que se va a realizar.
- Descripción: Descripción del estudio, se pueden ingresar los objetivos del estudio.
- Creador: Nombre de la persona o empresa que lleva a cabo el estudio.
- Fecha de publicación: Se selecciona la fecha en la cual se publica el estudio.
- Universo: Combo donde aparecen todos los universos ingresados en la sección de universos, la cual se describe a continuación.
- Cobertura Temporal: Se ingresa una fecha que indica cuando se realiza el estudio.
- Cobertura geográfica: Se ingresa el nombre del lugar donde se va a llevar a cabo el estudio, puede ser un país, una ciudad, etc.
- Cobertura Temática: Se ingresa una descripción de los temas que abarca el estudio.

3.2 UNIVERSOS

La siguiente imagen muestra la sección de “Universos”.

Nombre	Descripción
census2010	People Living in the United States on April 1, 2010
census	People Living in the United States

En esta sección se despliega una lista de los universos que tiene el estudio, los cuales van a aparecer como opciones en los combos de selección de universos que se encuentran tanto en “Información General” como en “Variables”, que vamos a ver más adelante.

Para agregar un nuevo universo a la lista, se debe ingresar el *Nombre*, una *Descripción* y finalmente presionar el botón de añadir: 

En caso de que no se ingrese el *Nombre* y la *Descripción* del universo y se presione el botón de añadir, los campos se indicarán con error (Como se muestra en la siguiente imagen), deshabilitando al usuario a ingresar el universo.

Nombre

Descripción

Para eliminar un universo de la lista, se debe seleccionar dicho universo y luego presionar el botón de eliminar: 

3.3 CÓDIGOS

La siguiente imagen muestra la sección de “Códigos”.

Información General

Universos

Códigos

Variables

Preguntas

Secuencias

Instrumento

Buscar 

Género

Ciudades

Nombre

Género

Valores

Valor Descripción 

Lista de Valores:

Valor	Etiqueta
1	Hombre
2	Mujer



En la lista de la izquierda que se muestra en la imagen anterior, se despliegan todas las listas de códigos que se encuentran en el estudio.

Para buscar una lista específica se ingresa el nombre en la caja de texto “Buscar”:

Buscar

El Sistema va filtrando automáticamente las listas que contienen en el *Nombre* las letras que se ingresan en el campo “Buscar”.

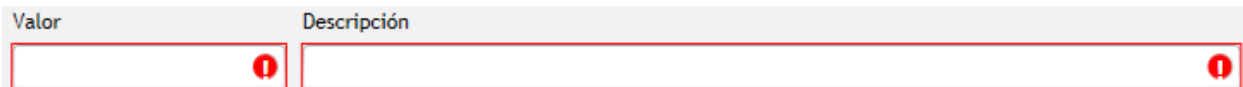
Para ingresar una nueva lista de códigos, se debe presionar el botón añadir: , el cual despliega en la lista una nueva lista con el nombre “Nueva Lista de Códigos”.

Para editar una lista de códigos, se debe seleccionar. Al seleccionar se despliegan los siguientes campos:

- **Nombre:** El cual no puede ser vacío.
- **Valores:** Grilla en la cual aparecen los códigos con dichos valores y descripciones.

Para agregar un nuevo valor a la lista, se debe ingresar el *Valor*, una *Descripción* y finalmente presionar el botón de añadir que se encuentra al lado del campo *Descripción*: 

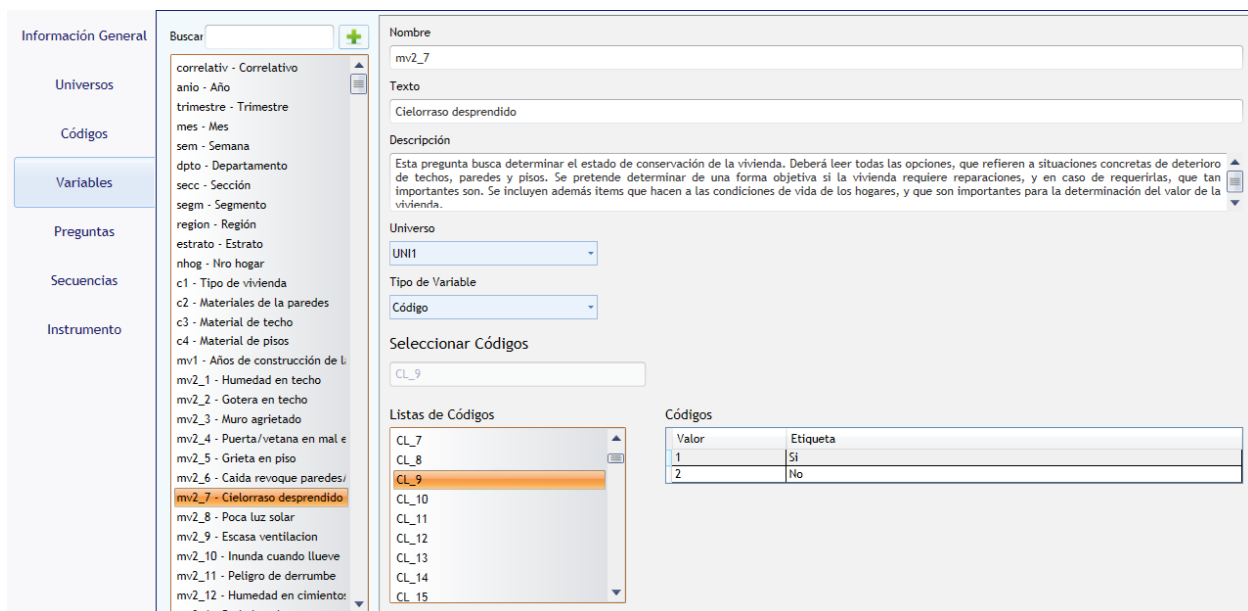
En caso de que no se ingrese el *Valor* y la *Descripción* del código y se presione el botón de añadir, los campos se indicarán con un error (Como se muestra en la siguiente imagen), deshabilitando al usuario a ingresar el código.



Para eliminar un código de la lista, se debe seleccionar dicho código y luego presionar el botón de eliminar: 

3.4 VARIABLES

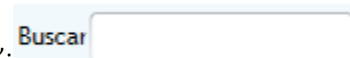
A continuación se muestra la sección de “Variables”.



Valor	Etiqueta
1	Si
2	No

En la lista de la izquierda que se muestra en la imagen anterior, se despliegan todas las variables que se encuentran en el estudio.

Para buscar una variable específica se ingresa el nombre en la caja de texto “Buscar”:

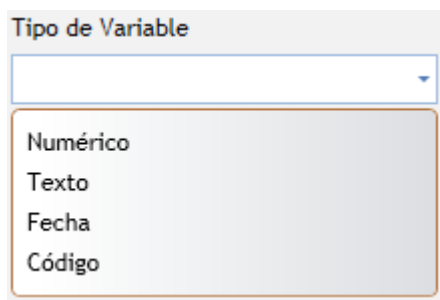


El sistema va filtrando automáticamente las variables que contienen en el *Nombre* o *Texto* las letras que se ingresan en el campo “Buscar”.

Para ingresar una nueva variable, se debe presionar el botón añadir: , el cual despliega en la lista una nueva variable con el nombre “Nueva Variable”.

Para editar una variable, se debe seleccionar. Al seleccionar se despliegan los siguientes campos:

- **Nombre:** El cual no puede ser vacío.
- **Texto:** Se ingresa una descripción corta de lo que significa dicha variable. Campo que puede quedar vacío.
- **Descripción:** Se ingresa una descripción de la variable, campo que puede quedar vacío.
- **Universos:** Combo de selección con los universos cargados en la sección de “Universos”
- **Tipo de Variable:** Combo que indica el tipo de dato de la variable. La siguiente imagen muestra las distintas opciones del combo:

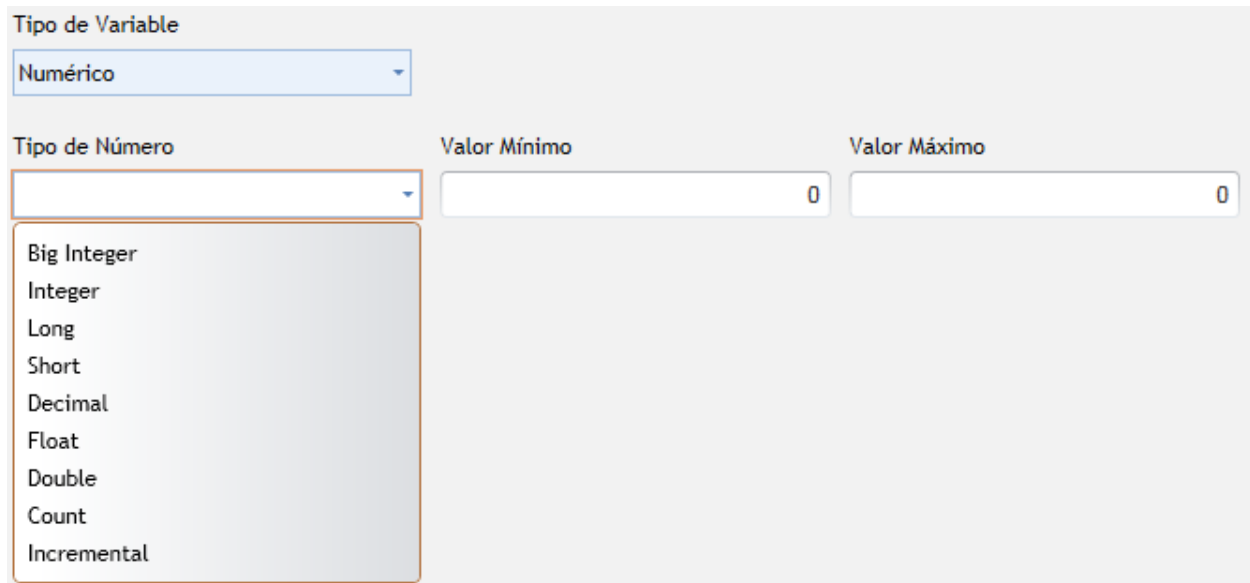


Formulario de selección de Tipo de Variable. El menú desplegado muestra las siguientes opciones:

- Numérico
- Texto
- Fecha
- Código

Para las distintas opciones de “Tipo de Variable” el sistema despliega distintos campos que pueden ser ingresados:

- Tipo de Variable: Numérico



Formulario de configuración para Tipo de Variable: Numérico. El menú desplegado para Tipo de Número muestra las siguientes opciones:

- Big Integer
- Integer
- Long
- Short
- Decimal
- Float
- Double
- Count
- Incremental

Además, se muestran los campos Valor Mínimo y Valor Máximo, ambos con el valor 0.

Para las variables de tipo numéricas se puede especificar:

- **Tipo de Número:** en la imagen anterior se pueden ver las distintas opciones.
- **Valor Mínimo:** valor mínimo que puede tomar la variable.

- **Valor Máximo:** valor máximo que puede tomar la variable.

- Tipo de Variable: Texto

Formulario de configuración para variables de tipo Texto. Incluye un menú desplegable 'Tipo de Variable' con 'Texto' seleccionado, y tres campos de entrada: 'Largo Mínimo' (valor 0), 'Largo Máximo' (valor 0) y 'Expresión Regular'.

Para las variables de tipo texto se puede especificar:

- **Largo Mínimo:** valor mínimo de caracteres que puede tener la variable.
- **Largo Máximo:** valor máximo de caracteres que puede tener la variable.
- **Expresión Regular:** Se puede ingresar un texto que indique la expresión regular que puede tener la variable.

- Tipo de Variable: Fecha

Formulario de configuración para variables de tipo Fecha. Incluye un menú desplegable 'Tipo de Variable' con 'Fecha' seleccionado, y un submenú 'Tipo de Fecha' que muestra las siguientes opciones: Fecha Hora, Fecha, Hora, Año, Mes, Día, Día Mes, Año Mes y Duración.

Para las variables de tipo fecha se puede indicar el tipo de fecha, cuyas opciones se muestran en la imagen anterior.

- Tipo de Variable: Código

Tipo de Variable
Código

Seleccionar Códigos
CL_9

Listas de Códigos

- CL_7
- CL_8
- CL_9
- CL_10
- CL_11
- CL_12
- CL_13
- CL_14
- CL_15

Códigos

Valor	Etiqueta
1	Si
2	No

Cuando se selecciona “Código” para el tipo de la variable se despliega una sección donde se listan todas las listas de códigos que se ingresaron en la sección anterior. Al seleccionar una lista, se despliegan los códigos de dicha lista a la derecha. Para relacionar la lista de códigos con la variable se debe hacer doble click sobre ella. Al relacionar una lista de códigos con una variable, queda el nombre de la lista en el campo de texto, como se muestra en la imagen anterior.

3.5 PREGUNTAS

A continuación se muestra la sección de “Preguntas”:

Información General

Universos

Códigos

Variables

Preguntas

Secuencias

Instrumento

Buscar

Q_1 - ¿Es hombre o mujer? Homi
Q_2 - ¿Cuántos años cumplidos t
Q_3 - ¿Tiene cédula de identida
Q_4 - PARA MAYORES DE 6 AÑOS
Q_5 - ¿Cree tener ascendencia...
Q_6 - ¿Cuál es su relación de pa
Q_7 - ¿La madre vive en este ho
Q_8 - ¿El padre vive en este hog
Q_9 - ¿Tiene cónyuge o pareja e
Q_10 - ¿Quién es?
Q_11 - ¿Cuál es el tipo de unión
Q_12 - ¿Actualmente está...
Q_13 - A lo largo de su vida, ¿ha
Q_14 - ¿Cuántos?
Q_15 - ¿cuántos de ellos viven e
Q_16 - ¿cuántos de ellos viven e
Q_17 - ¿Cuántos de ellos viven e
Q_18 - ¿Cuántos fallecieron?
Q_19 - ¿En que mes y año nació
Q_20 - ¿En que mes y año nació
Q_21 - ¿Tiene derechos vigentes
Q_22 - Con respecto al cuidado
Q_23 - Solo para quienes contes
Q_24 - ¿Quién paga la cuota?
Q_25 - Solo para quienes contes
Q_26 - ¿Cuál? Anote el N° de per
Q_27 - ¿Está afiliado a algún sist
Q_28 - ¿Sabe leer y escribir?
Q_29 - ¿Tiene algún hijo o hija?

Título
Q_15

Texto
¿cuántos de ellos viven en este hogar?

Instrucciones
A continuación y sólo para las mujeres se realizarán las siguientes preguntas sobre la fecundidad, de modo que en el caso de los hombres luego de preguntar el número total de hijos nacidos vivos se pasa a la pregunta 42 (E.2). En el caso de las mujeres se preguntará adicionalmente ¿cuántos de ellos viven en este hogar?

Variable

Variable asociada:
e39_2 - Hijos viven en el hogar Numérico

Buscar variable

Nombre	Texto	Tipo de Variable
Nombre	Etiqueta	Tipo

En la lista de la izquierda que se muestra en la imagen anterior, se despliegan todas las preguntas que se encuentran en el estudio.

Para buscar una pregunta específica se ingresa el nombre en la caja de texto “Buscar”:

Buscar

El sistema va filtrando automáticamente las preguntas que contienen en el *Título* o *Texto* las letras que se ingresan en el campo “Buscar”.

Para ingresar una nueva pregunta, se debe presionar el botón añadir: , el cual despliega en la lista, una nueva pregunta con el título “Nueva Pregunta”.

Para editar una pregunta, se debe seleccionar. Al seleccionar se despliegan los siguientes campos:

- **Título:** El cual no puede ser vacío.
- **Texto:** Se ingresa el texto de la pregunta.
- **Instrucciones:** Se ingresan las instrucciones para realizar la pregunta, campo que puede quedar vacío.
- **Sección de Selección de Variable:** Para poder asociar una pregunta a una variable se despliega la siguiente sección:

Variable

Variable asociada:

f101 - Disponibilidad para trabajar 

Buscar variable

Nombre Texto Tipo de Variable 

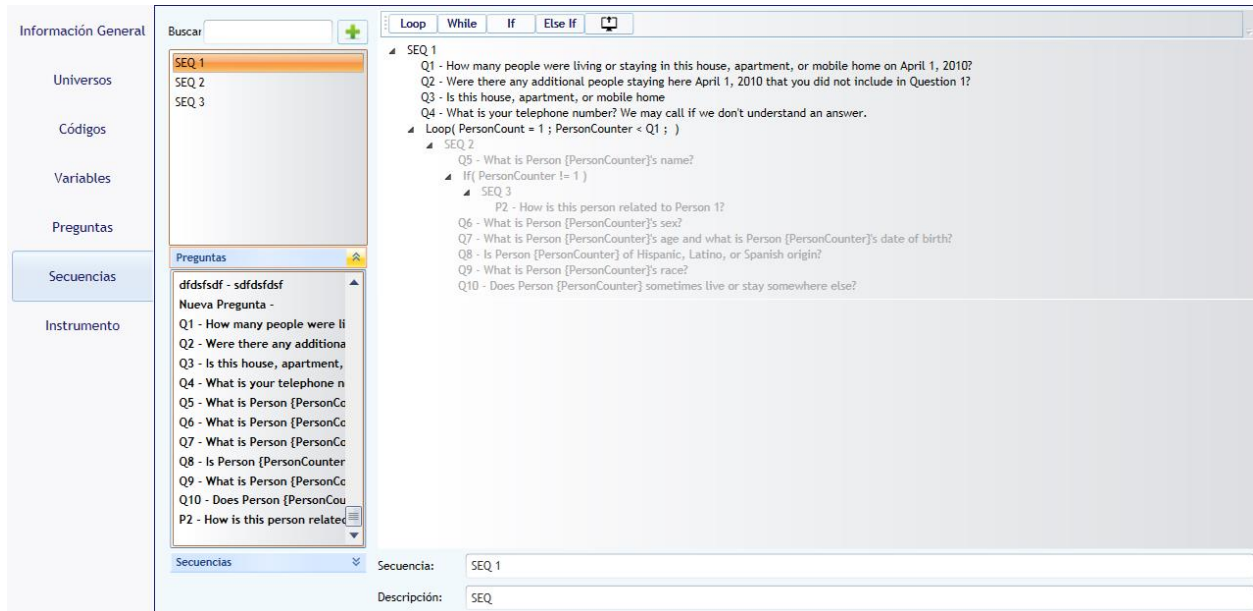
Nombre	Etiqueta	Tipo
correlativ	Correlativo	Número
año	Año	Texto
trimestre	Trimestre	Número
mes	Mes	Código
sem	Semana	Texto
dpto	Departamento	Código
secc	Sección	Texto
...	...	...

Se permite hacer una búsqueda de la variable que se desea asociar a través de los campos Nombre, Texto y Tipo de Variable, presionando el botón de Búsqueda . Las variables que se listan son las ingresadas en la sección anterior.

Para asociar la variable se debe hacer doble click desde la grilla, quedando la información en el campo “Variable asociada”. Para desasociar la variable de la pregunta, se presiona el botón de eliminar , dejando los campos en blanco.

3.6 SECUENCIAS

A continuación se muestra la sección de “Secuencias”



Dentro de esta sección se pueden construir secuencias para utilizar en el Instrumento o dentro de otras secuencias.

En la lista de más arriba de la izquierda, que se muestra en la imagen anterior, se despliegan todas las secuencias que se encuentran en el estudio.

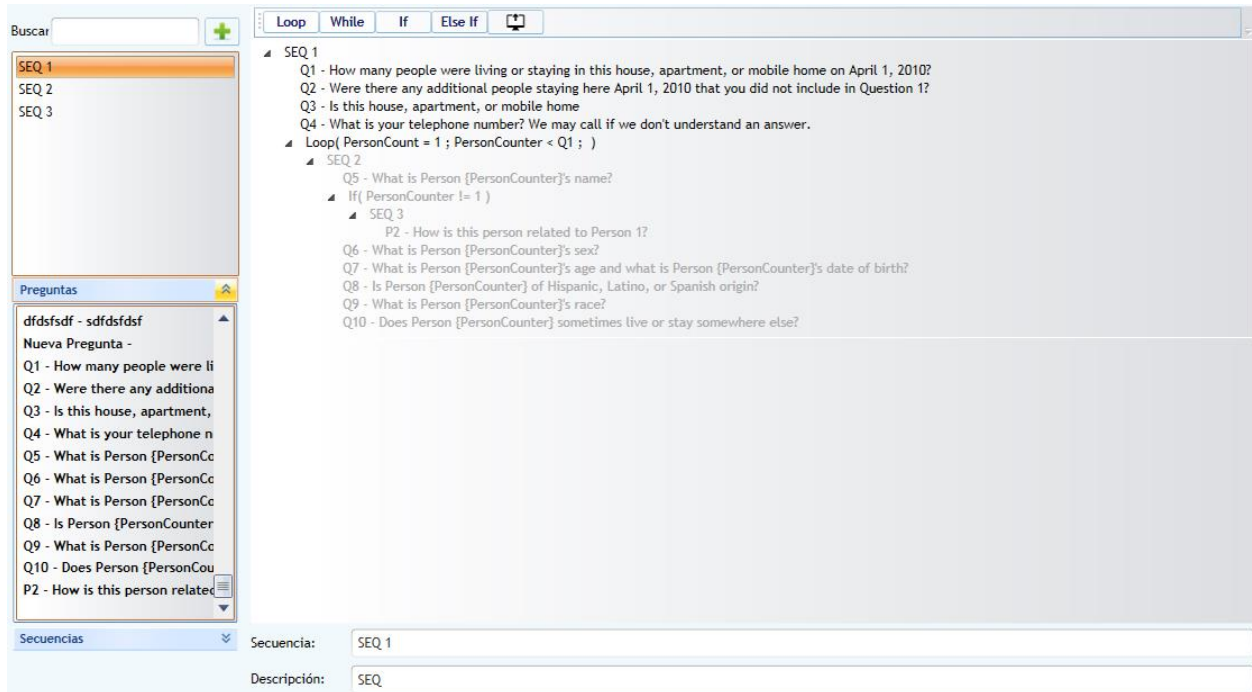
Para buscar una secuencia específica se ingresa el nombre en la caja de texto “Buscar”:



El Sistema va filtrando automáticamente las secuencias que contienen en el *Nombre* las letras que se ingresan en el campo “Buscar”.

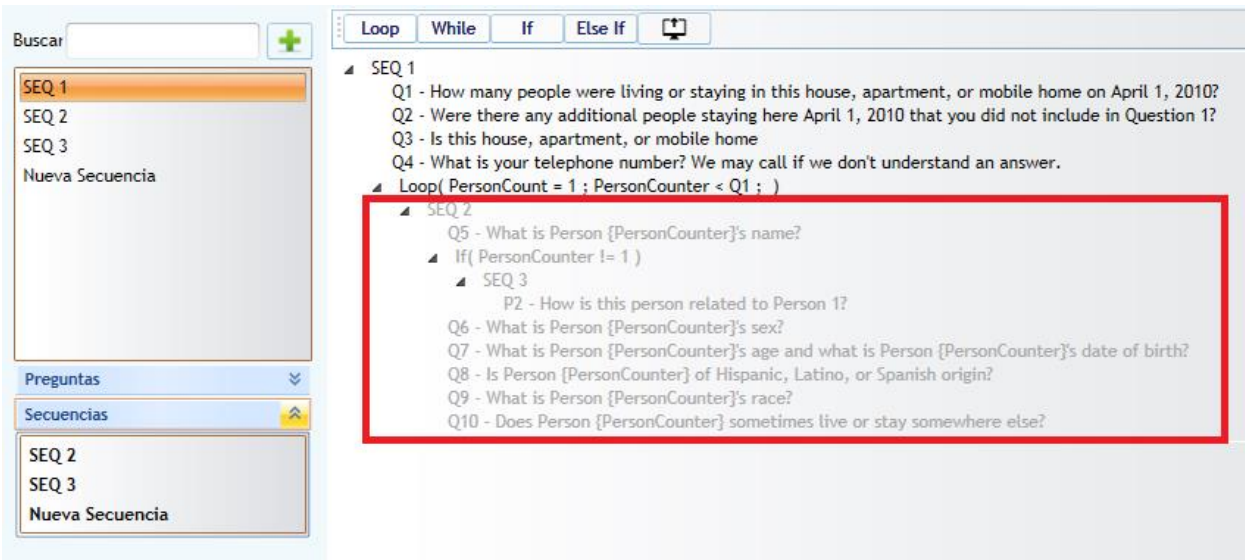
Para ingresar una nueva secuencia, se debe presionar el botón añadir: , el cual despliega en la lista, una nueva secuencia con el nombre “Nueva Secuencia”.

A continuación se muestra el cuadro de edición de secuencias:



Para la edición de una secuencia se debe seleccionar y luego editar. Las acciones que se permiten hacer son las siguientes:

- **Agregar Pregunta:** Desde la lista de preguntas se puede arrastrar una pregunta para ingresar en la secuencia.
- **Agregar Secuencia:** Desde la lista de secuencias se puede arrastrar una secuencia para ingresar en la secuencia que se está editando. Se listan todas menos la que esta seleccionada para editar. La secuencia que se agrega queda con un color gris claro y no se permite editar, como muestra la siguiente imagen:



Desde la barra de herramienta superior, se pueden realizar las siguientes acciones:

- **Agregar Loop:** Con el botón “Loop” se permite ingresar un loop a la secuencia y haciendo doble click aparece el siguiente “Pop Up”:

La lista de la izquierda muestra todas las variables del estudio, las cuales se pueden filtrar ingresando el Nombre en la caja de texto “Buscar”. Para utilizar la variable dentro de la condición se hace doble click sobre la variable, lo cual deja ingresado en la condición dicho nombre.

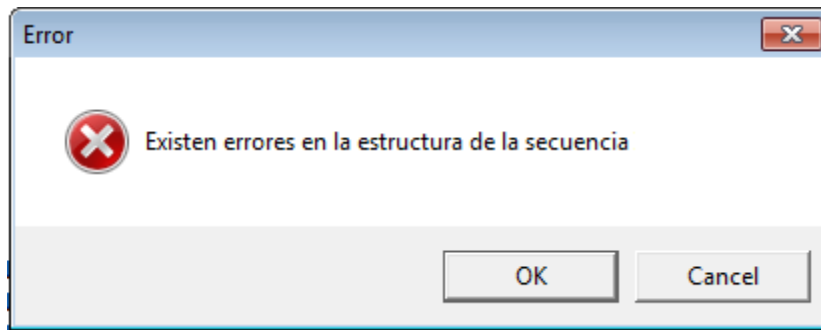
Para la condición del loop existen los siguientes campos:

- o **Etiqueta:** Identificador de la condición
 - o **Descripción:** Descripción de la condición del loop.
 - o **Condición:** Condición escrita en lenguaje PSPP.
 - o **Valor inicial:** Valor inicial de la variable que se encuentran en la condición.
 - o **Paso:** Indica el paso de la variable que se encuentra en la condición.
- **Agregar While:** Con el botón “While” se permite ingresar un while a la secuencia y haciendo doble click aparece el siguiente “Pop Up”:

La lista de la izquierda muestra todas las variables del estudio, las cuales se pueden filtrar ingresando el Nombre en la caja de texto “Buscar”. Para utilizar la variable dentro de la condición se hace doble click sobre la variable, lo cual deja ingresado en la condición dicho nombre.

Para la condición del while existen los siguientes campos:

- o **Etiqueta:** Identificador de la condición
 - o **Descripción:** Descripción de la condición del while.
 - o **Condición:** Condición escrita en lenguaje PSPP.
- **Agregar If:** Con el botón “If” se permite ingresar un if a la secuencia y haciendo doble click aparece el mismo “Pop Up” que para el “While”, permitiendo ingresar la condición.
 - **Agregar Elself:** Con el botón “Elself” se permite ingresar un else if a la secuencia y haciendo doble click aparece el mismo “Pop Up” que para el “While”, permitiendo ingresar la condición.
 - **Exportar Fragmento:** Por último con el botón  se permite exportar en formato fragmento de la versión DDI 3 la secuencia seleccionada. En caso de que la estructura de la secuencia no sea correcta, por ejemplo que haya un IfElse sin un If antes, el sistema muestra el siguiente mensaje de error:

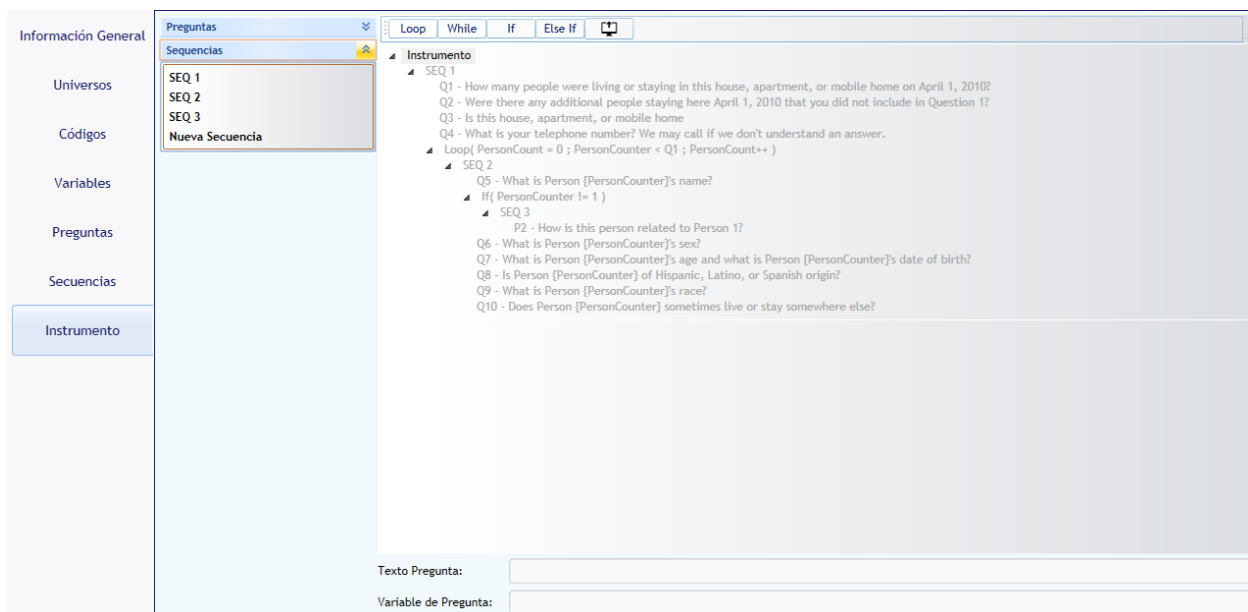


Debajo del cuadro de edición de secuencia se encuentran los siguientes campos editables con el nombre de la secuencia y una descripción:

Secuencia:	SEQ 1
Descripción:	SEQ

3.7 INSTRUMENTO

A continuación se muestra la sección de “Instrumento”



Dentro de esta sección se pueden construir el instrumento del estudio.

Las acciones que se permiten hacer sobre el instrumento son las siguientes:

- **Agregar Pregunta:** Desde la lista de preguntas se puede arrastrar una pregunta para ingresar en el instrumento.

- **Agregar Secuencia:** Desde la lista de secuencias se puede arrastrar una secuencia. La secuencia queda con un color gris claro y no se permite editar desde el instrumento, solamente se editan desde la sección de “Secuencias” como se vio en la sección anterior.

Desde la barra de herramienta superior, se pueden realizar las siguientes acciones:

- **Agregar Loop:** Con el botón “Loop” se permite ingresar un loop al instrumento y haciendo doble click aparece el siguiente “Pop Up”:

Buscar

correlativ
anio
trimestre
mes
sem
dpto
secc
segm
region
estrato
nhog
c1
c2
c3

Edición Condición

Etiqueta :

Descripción :

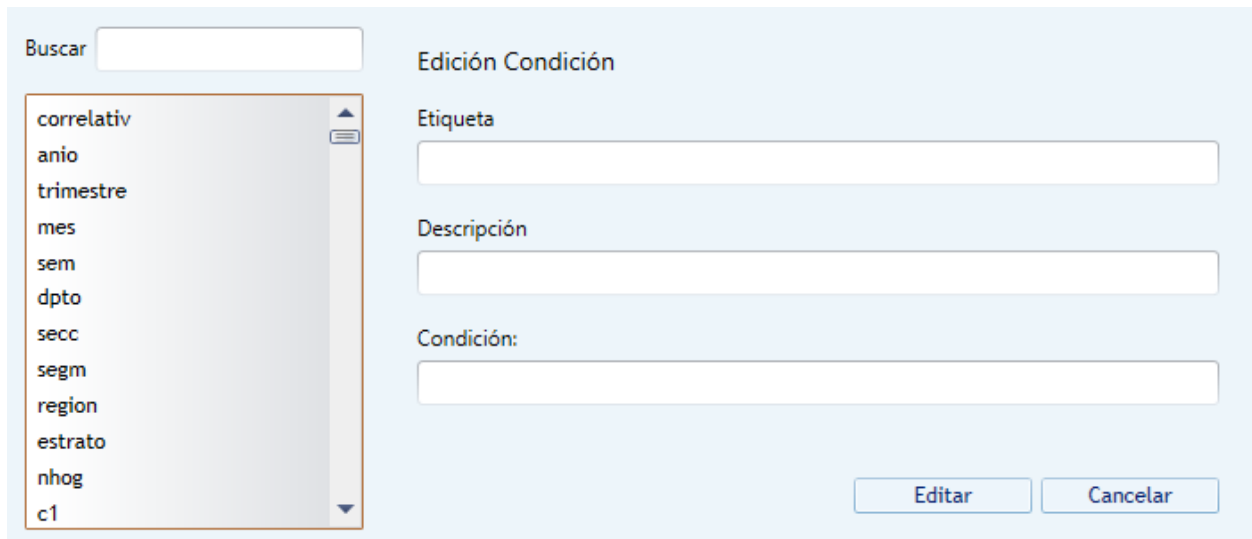
Condición:

Valor inicial: Paso:

La lista de la izquierda muestra todas las variables del estudio, las cuales se pueden filtrar ingresando el Nombre en la caja de texto “Buscar”. Para utilizar la variable dentro de la condición se hace doble click sobre la variable, lo cual deja ingresado en la condición dicho nombre.

Para la condición del loop existen los siguientes campos:

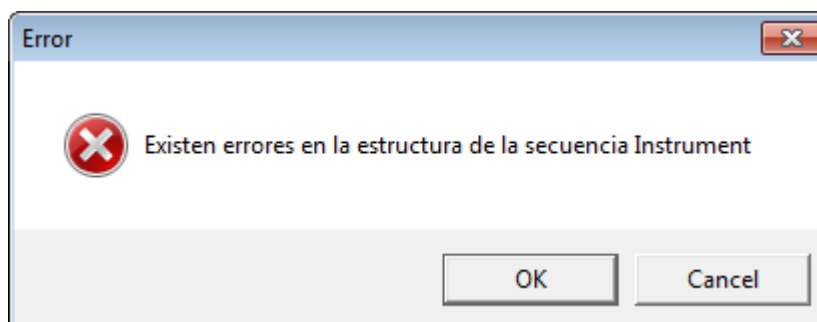
- o **Etiqueta:** Identificador de la condición
 - o **Descripción:** Descripción de la condición del loop.
 - o **Condición:** Condición escrita en lenguaje PSPP.
 - o **Valor inicial:** Valor inicial de la variable que se encuentran en la condición.
 - o **Paso:** Indica el paso de la variable que se encuentra en la condición.
- **Agregar While:** Con el botón “While” se permite ingresar un while al instrumento y haciendo doble click aparece el siguiente “Pop Up”:



La lista de la izquierda muestra todas las variables del estudio, las cuales se pueden filtrar ingresando el Nombre en la caja de texto “Buscar”. Para utilizar la variable dentro de la condición se hace doble click sobre la variable, lo cual deja ingresado en la condición dicho nombre.

Para la condición del while existen los siguientes campos:

- o **Etiqueta:** Identificador de la condición
 - o **Descripción:** Descripción de la condición del while.
 - o **Condición:** Condición escrita en lenguaje PSPP.
- **Agregar If:** Con el botón “If” se permite ingresar un if al instrumento y haciendo doble click aparece el mismo “Pop Up” que para el “While”, permitiendo ingresar la condición.
 - **Agregar Elself:** Con el botón “Elself” se permite ingresar un else if al instrumento y haciendo doble click aparece el mismo “Pop Up” que para el “While”, permitiendo ingresar la condición.
 - **Exportar Fragmento:** Por último con el botón  se permite exportar en formato fragmento de la versión DDI 3 el instrumento. En caso de que la estructura del instrumento no sea correcta, por ejemplo que haya un IfElse sin un If antes, el sistema muestra el siguiente mensaje de error:

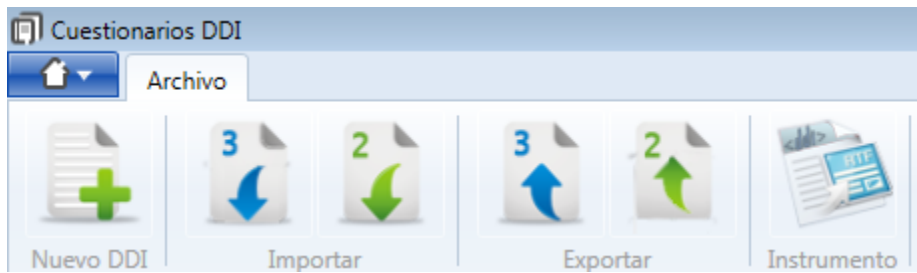


Debajo del cuadro de edición del instrumento se encuentran los siguientes campos que indican el texto de la pregunta seleccionada y la variable asociada a dicha pregunta. Estos campos no se pueden editar desde esta sección.

Texto Pregunta:	
Variable de Pregunta:	

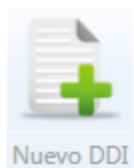
4. MENÚ PRINCIPAL

A continuación se muestra el menú principal y se describe cada una de dichas funcionalidades:



4.1 NUEVO DDI

Al presionar el botón “Nuevo DDI” se despliegan todos los datos del cuestionario en blanco para permitir al usuario ingresar un nuevo cuestionario.



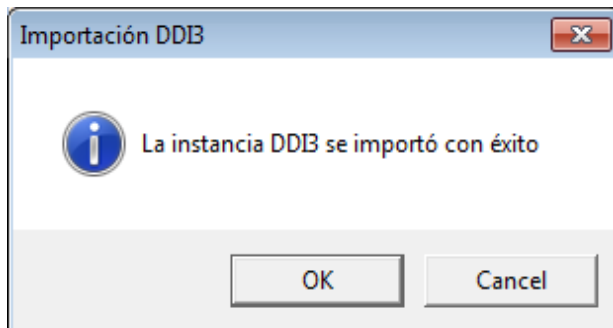
4.2 IMPORTACIÓN

Para permitir al usuario ingresar los datos de una instancia DDI existente, existen los siguientes botones:



La primera opción permite importar un archivo de tipo XML con información de una instancia o fragmento en la versión DDI3. La segunda opción permite importar un archivo XML conteniendo información de una instancia escrita en versión DDI2.

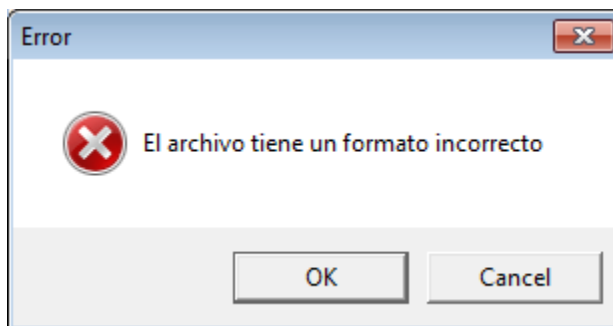
En caso que el archivo DDI3 se pueda importar con éxito, el sistema muestra el siguiente mensaje:



En caso que el archivo DDI2 se pueda importar con éxito, el sistema muestra el siguiente mensaje:



En caso de que el archivo no tenga una estructura válida, el sistema muestra el siguiente mensaje:

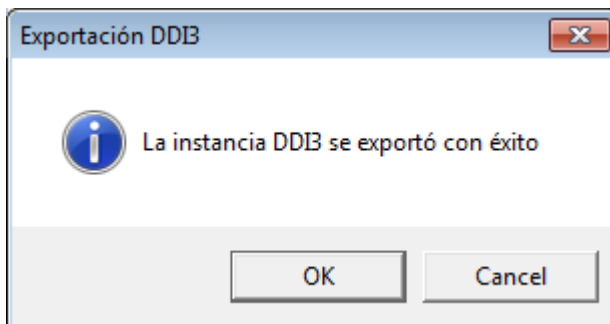


4.3 EXPORTACIÓN

Así como la importación, también existe la posibilidad de exportar en ambas versiones la información ingresada por el usuario a través de los siguientes botones:



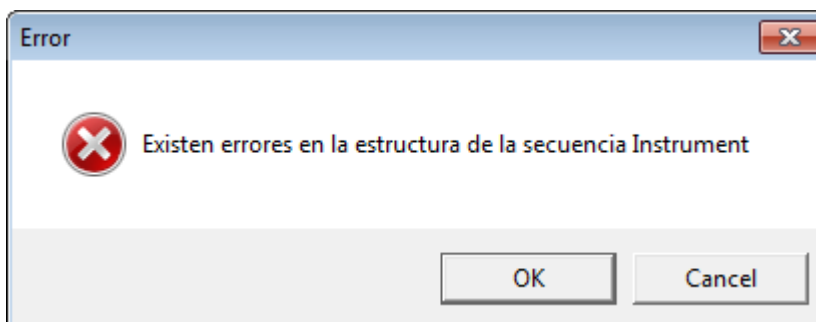
En caso que el archivo DDI3 se pueda exportar con éxito, el sistema muestra el siguiente mensaje:



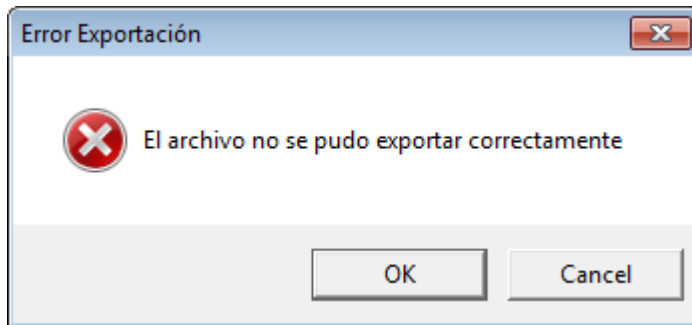
En caso que el archivo DDI2 se pueda exportar con éxito, el sistema muestra el siguiente mensaje:



En caso de que el instrumento no tenga una estructura válida y se quiera exportar en DDI 3, el sistema muestra el siguiente mensaje:



En caso de que la aplicación no pueda exportar el archivo por algún error genérico, se muestra el siguiente mensaje de error:



NOTA: Ya que la versión DDI2 no tiene manejo de instrumento y secuencias en su estructura, al momento de exportar en la versión DDI2, si existen definidas secuencias e instrumentos, estos no van a aparecer en el archivo XML exportado.

4.4 RTF

Por último en el menú principal se encuentra el siguiente botón cuya funcionalidad es exportar en formato RTF el instrumento creado:

