

# El manejo lingüístico de la ontología textil

## Las equivalencias

Efrain Yujra Saire <sup>1</sup>

### Resumen

El artículo trata de explicar de manera muy resumida los pasos de la elaboración de una ontología de dominio, en este caso en particular, de una ontología textil. En la estructura de una ontología participan muchos componentes y uno de ellos es el aspecto del manejo léxico. Y ese es el punto central en el que nos detendremos en este pequeño trabajo que intenta compartir con la comunidad de investigadores un método para organizar la información de dominio. Entonces se describe el modo en que la información léxica se organiza empleando un elemento que denominamos *equivalencia* para recuperar los datos en la web semántica.

El lector verá que esta herramienta es bastante útil para sistematizar la información holísticamente por lo que seguramente la podrá reusar en su área de dominio.

**Palabras Clave:** Textil andino, ontología, equivalencia, sinónimo, traducción.

### Introducción

Se cree que el manejo lingüístico se limita a la enseñanza de lenguas, pero la lingüística también es muy útil cuando se quiere estructurar la información en un modelado conceptual de los componentes de un área de especialidad. Como no se conoce esto, no se contempla el conocimiento lingüístico para organizar la información. El buen uso de la terminología organizada en taxonomías y relaciones sirve para hacer una sistematización adecuada. Por ello, aprovechando la experiencia de la organización de la información en el Proyecto Comunidades Textiles, llevado a cabo por el ILCA y la Universidad de Birkbeck de Londres, se presentará la utilidad de las *equivalencias* al servicio de la difusión de la información en el entorno de la web.

Presentaremos, pues, una forma de solucionar el problema de la búsqueda de información precisa que se tiene en el mundo actual, para lo cual se combinó el conocimiento informático de ontologías con el conocimiento lingüístico de campos

---

<sup>1</sup> El autor es Lingüista e Investigador; Licenciado en Lingüística e Idiomas; participó en el Proyecto Comunidades Textiles dirigido por el ILCA apoyando el modelado ontológico y el manejo lingüístico. Ha publicado el texto Lenguaje e interlenguaje; le interesa el manejo lingüístico en diferentes áreas del conocimiento como en la educación, planificación y recuperación de lenguas, entre otros. Correo electrónico: efrainyujra@hotmail.com.

semánticos y taxonomías. En este entendido se debe comprender que las consultas en la web son procesadas por máquinas que contienen información rellena por humanos, pero que tienen la limitante de no poder relacionar los elementos componentes entre sí, ya que aquello sólo ocurre en la mente del ser humano. Si bien la interrelación de elementos es obvia para cualquier ser humano por su conocimiento holístico del mundo, es también compleja si no se tiene la capacidad de sistematización y abstracción cuando tiene que responder preguntas interrelacionadas.

En consecuencia, primero se tiene que alimentar a la máquina y enseñarle responder, inferir o diferenciar nuestras consultas. Se debe empezar entendiendo que cada ser humano tiene en su cabeza una gran ontología de información interrelacionada de acuerdo a su formación y a su experiencia, y que hay que transferir esa información ordenadamente a una máquina para que tenga la capacidad de responder de modo semejante al pensamiento humano.

Ahí es donde entra el manejo preciso de las *equivalencias* que ayudó al proyecto de la Ontología Textil a organizar la información dispersa en una sola estructura para recomponer el conocimiento de la Producción Textil y a la vez lograr que dicho conocimiento esté al alcance de los usuarios en la web.

Por tanto, en esta ponencia se expondrá el manejo lingüístico que se hizo en el proyecto de Comunidades de Práctica Textil.

Se aclarará lo que es una ontología y su aplicación desde lo lingüístico.

Se abordará los problemas más usuales de las equivalencias que se presentaron en el proyecto de Comunidades de Práctica Textil.

## **Manejo lingüístico en el proyecto de comunidades de práctica textil**

En el proyecto de Comunidades de Práctica Textil se propuso organizar la información referente al textil andino tomando como lineamiento la cadena productiva. Sin embargo, esa manera gradual de exponer el conocimiento fragmentaba las relaciones que estaban muy alejadas entre los momentos, actividades, procesos, materiales e instrumentos o bien se tendía a reduplicarla información en las diferentes etapas en que se requería presentar nuevamente la información. Es entonces que se recurre a la ontología que organiza los datos de modo más integral evitando la reduplicación de la información.

El otro problema era que se requería contar con datos para realizar pruebas de las búsquedas por lo que se tuvo que pensar en maneras de devolver la información de manera más precisa. Para esto coadyuvaban las palabras clave que en la ontología son todas las entradas relacionadas vertical u horizontalmente. Aquí se hallaron inconsistencias, ya que muchas palabras tendían a aparecer en varios lugares de una taxonomía o bien en taxonomías paralelas. Esto presionó al equipo a la generación de terminología adecuada. Si bien el peligro de esta generación artificial de terminología es que los usuarios se sientan confundidos, se ha querido ser lo más lógicos posibles para evitar falsas premisas. Lo siguiente es compartir con la comunidad esa nueva terminología y, a través de una

socialización adecuada, se logre convertirla en parte del conocimiento cotidiano de los estudiosos del textil.

Entonces, al encontrar que el vocabulario del textil se hacía cada vez más preciso, se tuvo que incorporar la participación del componente lingüístico por la complejidad de las búsquedas en el entorno de la web semántica. Esto quiere decir que se tuvo que imaginar las diferentes posibilidades de preguntas que el usuario podría hacer y lo que esperarías obtener al realizar dichas consultas. Se pudo generalizar cuatro tipos genéricos de preguntas que se podían responder sin mucho problema. El hecho era que las respuestas muy concretas no satisfacían al equipo de investigadores por lo que se tuvo que idear en maneras no textuales de responder a las consultas. Eso ya es parte de otro momento del proyecto.

Parecía que todo iba bien cuando se pensó en el usuario de la futura página especializada del textil. Por lo menos hallamos tres grupos de hablantes que podrían estar interesados en investigar sobre el textil: los hispanohablantes, los anglófonos y los hablantes de las lenguas originarias aymara y quechua. Esa amplia comunidad interesada en el textil iba a querer analizar desde su lengua la conceptualización y la pertinencia de la terminología empleada por lo que se tuvo que tener mucho cuidado en elaborar un vocabulario en la que se representaran las lenguas mencionadas.

## ¿Qué es una ontología?

Sería tedioso continuar hablando de la ontología sin saber qué es y en qué consiste. Por ello vamos a conceptualizar de manera muy clara lo que es.

La *ontología* es la organización de los elementos componentes y relacionados con una entidad. (Yujra, 2012). Esto quiere decir que una entidad se relaciona con sus elementos constitutivos y con otras entidades con las que tiene directa o indirecta relación.

Recordemos que la palabra “ontología” significa el estudio de una entidad, sea esta concreta o abstracta. En ese entendido, la entidad cultura que es abstracta, puede relacionarse con un producto textil que es concreto.

Hay que poner muy en claro que una entidad se compone tanto de la materialidad como de elementos abstractos como la carga cultural.

Una entidad se relaciona con otras entidades por su composición material y por su vida social, que comienza desde la concepción de la idea de elaborar el textil, ya que el textil es un documento que alberga información y, por supuesto, el emisor de esa información es quien elabora el textil mismo.

Para redondear la idea: El textil andino es una entidad. El textil se compone de fibras, tintes, técnicas y estructuras. Además se afilia a una determinada cultura por el estilo que presenta su iconografía. El textil se relaciona con los instrumentos con los que fue elaborado, con la comunidad en la cual se tejió y con la persona que lo usó.

Por esa particular complejidad que tiene el textil andino se lo tuvo que sistematizar en una ontología para poderlo representar de manera integral y no solamente desde un solo plano.

Pero cabe aclarar que existen dos componentes en una ontología: la parte conceptual, que es la que ya se ha descrito, y la parte informática que se encarga de colocar en un entorno especialmente desarrollado para la información ontológica desde una base de datos diferente a las convencionales y el empleo de buscadores semánticos que tratan de extraer información y presentarla de manera amigable al usuario.

## Elementos de una ontología

La ontología en cuanto a su estructura se compone de CLASES e INSTANCIAS, RELACIONES (JERÁRQUICAS y BINARIAS) y PROPIEDADES DE DATO o ATRIBUTOS.

Las **Clases** son los elementos concretos o abstractos que son conceptualizados teniendo un referente “real”<sup>2</sup>. La clase agrupa a subclases que pertenecen a su mismo grupo. Las instancias son los ejemplos que se pueden dar de las clases. Veamos los ejemplos:

**Clase:** Museo

**Subclase:** Museo de Bolivia

**Instancia:** Museo Nacional de Etnografía y Folklore

Esa agrupación de entidades se denomina taxonomía y es la primera relación que se conoce entre las clases. A esto denominamos relación jerárquica.

La relación binaria se da entre clases de otros grupos. Ejemplo:

Producto textil **\_se custodia en\_** Museo (relación entre clases)

Faja-bolsa ILCA\_MSF032 **\_se custodia en\_** MUSEF (relación entre instancias)

Véase que la **relación es binaria** porque intervienen dos **clases** diferentes, pero indirectamente se muestra la **relación jerárquica** donde la **instancia** es parte de una **clase**.

Finalmente, se tiene a las propiedades de dato o atributo que cada clase tiene como característica. Ejemplo:

Producto textil **\_tiene Tamaño\_** [...] (Propiedad de dato de una clase)

Faja-bolsa ILCA\_MSF032 **\_tiene Tamaño\_** Mediano (Propiedad de dato de una instancia).

En el modelado de clase se ha dejado Corcheas con puntos suspensivos que indican que se deberá llenar ese dato particular no siendo una clase que se relacione. En el modelado de instancia se puede apreciar el dato preciso con que se distingue la instancia.

2 “Real” no implica que la entidad referida sea solamente concreta, también puede ser abstracta como la entidad **cultura**.

## Problemas lingüísticos en el proyecto

como en todo proyecto, se fueron presentando problemas que requirieron de soluciones investigativas y algunos de los problemas fueron lingüísticos, esto quiere decir que se requería de soluciones en el manejo de la lengua. En este apartado sólo mencionaremos los problemas hallados y más adelante se abordará a detalle y ejemplificado los problemas y la respectiva solución que le dimos en el proyecto.

- **La denominación de los objetos:** Los productos textiles presentaron el problema de la ubicación en la taxonomía ya que algunos productos textiles incluían dentro de su grupo a otros por tener las mismas características y definiciones. En el caso de los instrumentos el problema se presentó en la característica de multiuso de un mismo instrumento lo cual obligaba a un elemento a aparecer en dos lugares de la taxonomía.
- **La denominación de los materiales:** los **materiales** tales como la fibra, los tintes, los mordientes, la madera, la arcilla, la piedra, entre otros, no contaban con definiciones adecuadas a la cadena productiva, sin embargo se los organizó relacionados a los **objetos e instrumentos**. Usualmente, se considera a la materialidad como atributo y no como clase, pero por una necesidad de filtro por materialidad, en el proyecto se lo manejó como clase. Cuando se hacen filtrados por clase se pueden detectar instancias de manera más precisa ya que como atributo o propiedad de dato no siempre tienen la misma redacción en el registro textil.
- **La denominación del color:** El color es una entidad muy compleja ya que su denominación se realiza por países. Es decir que un mismo color puede ser denominado de varias maneras en diferentes países, otra manera es emplear el listado de alguna institución dedicada al color como las empresas de tintes y finalmente por la codificación que de acuerdo a su nivel de claro oscuro podría tener diferentes gamas. A pesar de ello la codificación es el sistema de denominación más adecuado.
- **La denominación de las técnicas y estructuras:** Una de las novedades que propone el proyecto es la generación de denominaciones de técnicas y estructuras con concepciones divergentes a las tradicionales, ya que se parte desde la experiencia de la tejedora hacia la teorización. Lo difícil fue hacer coincidir dicha terminología con la terminología existente.

## Las equivalencias en el proyecto de comunidades de práctica textil

El empleo de toda esa terminología hizo que se elaboraran listados de palabras de cada una de las taxonomías y se las fue comparando para hallar redundancias. Esas tablas se fueron complicando ya que debían tener transversales que no se podían manejar en una base de datos convencional. Para esto se tuvo que recurrir a las equivalencias.

## Concepto de equivalencia

Previo a la conceptualización de ‘equivalencia’ es necesario conocer otros términos que le precedieron a los cuales reemplaza.

- **Sinónimo:** Palabras diferentes con el mismo referente o conceptualización. En palabras de Dubois (1979: 572) se dirá que “son sinónimas las palabras que poseen el mismo sentido, o aproximadamente el mismo, y formas diferentes”.
- **Término dialectal:** Son palabras que pueden diferenciarse de una región a otra con un mismo referente. Aunque Dubois (1979: 191) extiende el sentido de dialecto a las estructuras y a lo fonético en torno a una lengua.
- **Traducción:** palabra equivalente en dos o más lenguas. El diccionario lingüístico presenta el término traducir “que es enunciar en otra lengua (o lengua meta) lo que se enuncia en una lengua origen, conservando las equivalencias semánticas y estilísticas. (Dubois, 1979:606)

**Equivalencia:** Se refiere a palabras diferentes o con variación fonética o gráfica, dialectológicamente diversas, de idiomas distintos pero con un referente común. En el proyecto se ha querido centralizar con este término la igualdad o parecido conceptual respecto a un referente ya que apuntar desde varios grupos de palabras a un mismo referente o significado complica la llegada a las respuestas. Esta conceptualización no responde a la definición tradicional en lingüística ni a la concepción matemática que implica igualdad absoluta.

A continuación, es necesario saber para qué sirve una equivalencia y por qué se podría pensar que con las otras agrupaciones ya se podrían hacer las búsquedas. Además hay que aclarar el término búsqueda muy empleado en esta exposición y diferenciarla de una consulta semántica.

- ***Utilidad de la equivalencia***

La equivalencia simplifica la búsqueda de información ya que acumula todos los posibles términos asociados a un mismo concepto al cual se quiere llegar.

- ***¿Qué es una consulta o búsqueda?***

Es el ingreso de palabras clave en el buscador de las páginas web para acceder a la información.

- ***¿Qué es una consulta semántica?***

Es la búsqueda de información con cadenas de palabras interrelacionadas para acceder a la información con mayor precisión.

## Problemas que soluciona la equivalencia

Se debe partir con una premisa importante: para obtener una comparación que nos devuelva como resultado una diferencia o una igualdad tenemos que tener dos elementos que tengan conceptualmente igualdad o diferencia opuesta entre los dos términos en comparación. La información contextual y significativa previa de cada uno de los términos asegurará el máximo de precisión al momento de hacer un ejercicio básico de consulta. Pero no nos quedemos en sólo los dos elementos comparados; es importante hallar los niveles de comparación entre estos elementos. Aquí por niveles quiero decir las diferentes

posibilidades de comparación que se tiene entre dos elementos. De inicio se parte de un solo término, el segundo variará de acuerdo al nivel que se escoja.

Por ejemplo, el término textil según el nivel de comparación tendría varios pares de comparación. El primer nivel de comparación se da dentro del mismo idioma, lo que llamamos “sinónimo”, es decir una palabra que conceptualmente es semejante y que se podría alternar en el uso cotidiano de dicho término sin alterar el significado del enunciado.

Por ejemplo:

El textil andino es elaborado de manera manual.

El tejido andino es elaborado de manera manual.

Nótese que textil y tejido son manejados como equivalentes y el significado del enunciado no se ha alterado. Pero si se cambiara el contexto de uso de un término estaríamos frente a otro significado levemente diferente:

- 1) El textil es la industria más expandida por la China.
- 2) Para hacer un textil se requiere de fibra de camélidos.
- 3) Para hacer un tejido se requiere de fibra de camélidos

Entre las oraciones 1 y 2 la palabra textil tienen referentes distintos. En uno se refiere al conjunto de textiles producidos industrialmente y tal vez con alto grado de material sintético. Pero en 2 se refiere al conjunto de productos obtenidos de manera artesanal en el área andina. Lo interesante es que solamente se pueden alternar el uso del término ‘tejido’ entre 2 y 3, pero no en 1 ya que sonaría forzado. Esto implica que el conocimiento de uso de un término es determinante para la conceptualización y búsqueda de equivalentes dentro de un mismo idioma.

Otro ejemplo nos mostrará que un mismo término puede ser empleado con sentidos diferentes, lo cual dentro de una búsqueda podría hacer confundir al buscador. Consideremos a este como otro nivel dentro del mismo idioma. Veamos:

- 1) El tejido muscular está adormecido.
- 2) Un tejido requiere de tramas y urdidos.

Si bien la palabra tejido tiene cierto parecido en las dos oraciones los referentes son abismalmente diferentes. Por tanto, si se colocara en el buscador solamente la palabra tejido las respuestas serían infinitas de acuerdo al uso contextual que se le haya dado.

Por otro lado, existe alguna terminología en castellano que tienen referentes realmente diferentes: por ejemplo la palabra ‘lista’:

- 1) Las listas de estudiantes son muy largas.
- 2) Se debe usar listas para hacer un tejido.

Véase que aquí la palabra ‘lista’ tiene diferentes referentes en 1 y 2 y no tienen semejanza ni por comparación. Estos términos son de hecho los que más podrían confundir y desanimar al usuario cuando realiza una búsqueda.

El tercer nivel de comparación está fuera del idioma, es decir las palabras que se comparan están traducidas en otros idiomas. Es justamente desde aquí que se fue creando la tabla de equivalencias en el proyecto de la Ontología Textil.

Véamos los ejemplos:

ESPAÑOL	INGLÉS	AYMARA
Fibra	fiber	t'arwa

En el caso de la traducción las palabras equivalentes son obviamente diferentes, salvo algún término que haya sido transferido de un idioma a otro y que en ambos idiomas se escriban del mismo modo. Por ejemplo: *awayu* del aymara se pasó al español con la forma **ahuayo**, y sin variación alguna se transfirió al inglés como **ahuayo**. Obviamente, dentro de la literatura sobre textiles en inglés fue variada lo cual se presentó en la tabla de **equivalencias** como **variantes en inglés**: *mantle, carrying cloth, shawl*.

En este caso se puede apreciar que la palabra ahuayo de origen aymara fue adecuado al castellano de acuerdo al sonido y esta a su vez fue transferida al inglés sin ningún cambio.

Pero en cada idioma se podría hallar otro problema, el de la variación gráfica. Tenemos un ejemplo en castellano:

ESPAÑOL	VARIACIÓN GRÁFICA
ahuayo	aguayo

Veamos con otro ejemplo más completo la presencia de la mayoría de los componentes de la tabla de equivalencias.

ESPAÑOL	INGLÉS	AYMARA	QUECHUA	VARIANTES EN ESPAÑOL
Sobrecalza	Knitted legging legging	Sika winchuka	Pujllaypa	Guardapierna Polaina

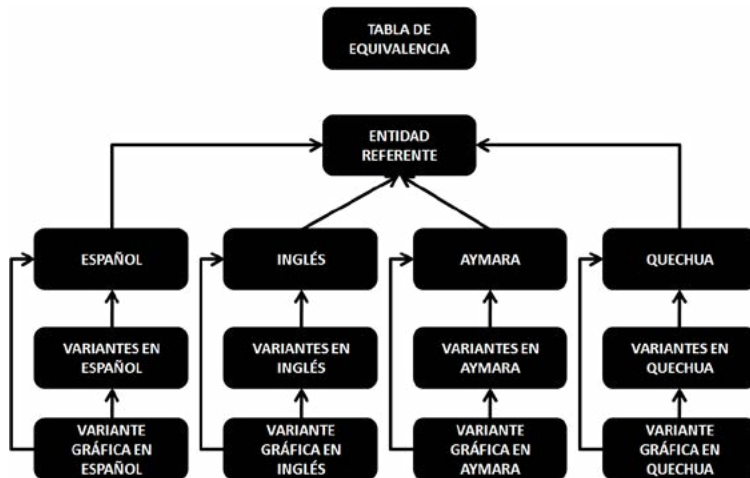
Por tanto, la precisión al buscar información parte desde el referente hacia la concepción fonético-gráfica que representa al objeto referido. Esto, que es tan obvio en la elaboración de vocabularios, glosarios y diccionarios, no lo sabe un buscador de información. En consecuencia, hay que alimentar una base de datos con información más detallada y precisa, en la medida de las posibilidades. Solamente así se puede conseguir que una página web responda satisfactoriamente a las consultas complejas.

Tomando nuestra tabla como ejemplo para una búsqueda podemos hacer la consulta escribiendo en el buscador ‘POLAINA’ que es la VARIANTE EN ESPAÑOL de



‘SOBRECALZA’. Sin duda, se recuperará la misma información que para la entrada principal porque las consideramos equivalentes. Y si un usuario de habla inglesa escribe ‘legging’, la información que se obtenga será la misma. En otras palabras, la tabla de equivalencias es una manera de generar palabras clave para acceder a la información almacenada en una base de datos. En el caso del textil andino, se pudo no sólo hacer una lista de palabras clave, sino que esta experiencia ayudó a sistematizar gran parte del modelado del conocimiento.

Para terminar esta exposición presento un gráfico donde se puede apreciar la estructura conceptual completa de una tabla de equivalencias.



Resumiendo, se puede decir que una entidad puede tener una entrada desde cualquier idioma, en el proyecto se trabajaron 4 idiomas, tomando como base el español. Cada idioma puede tener variantes o sinónimos mediante los cuales también poder llegar a la información. Además, tanto la entrada principal como las variantes pueden tener variantes gráficas con la cuales enriquecerían las posibilidades de acceder a la información.

## Conclusión

Compartir la experiencia que se tuvo en el proyecto Comunidades de Práctica Textil desde el campo lingüístico nos muestra que los elementos de cualquier área de especialidad puede ser sistematizada ontológicamente ya que cada elemento componente o relacionado forma parte de un conocimiento integral que se debe ir recuperando poco a poco.

La ontología y el uso de equivalencias presentan una nueva manera de acceder a la información. Pero exige un trabajo previo de los actores que generan la información, no así del componente informático, que si bien coadyuvan para hacer una página web, son sólo transcriptores de la información previamente sistematizada.

El textil andino puede ser estudiado desde diferentes facetas, pero si se la estudia integralmente se puede ir armando y respondiendo interrogantes que no se podían solucionar si se encerraba en una sola faceta. Es decir que al interactuar con la información de las otras facetas se puede obtener mejores resultados investigativos.

Finalmente, agradecer la posibilidad de compartir el conocimiento generado en el proyecto a la espera de que más entidades se atrevan a ingresar al mundo de las ontologías que amplía los horizontes a niveles desconocidos y justamente en lo desconocido está la respuesta que se está buscando.

## Bibliografía

ARANO, Silvia (2003) La ontología: una zona de interacción entre la Lingüística y la Documentación, [en línea]. "Hiptertext.net", núm. 2, 2003. <http://www.hiptertext.net>

CIDOC – CRM (1995) International Guidelines for Museum Object Information: The CIDOC Information Categories, Published by the International Committee for Documentation of the International Council of Museums.

CIDOC – CRM (2009) Definition of the CIDOC Conceptual Reference Model, Produced by the ICOM/CIDOC Documentation Standards Group, continued by the CIDOC CRM Special Interest Group. Editors: Nick Crofts, Martin Doerr, Tony Gill, Stephen Stead, Matthew Stiff.

DUBOIS, Jean y otros (1979). Diccionario de Lingüística. Madrid: Alianza editorial.

GETTY (2010) Top of the AAT hierarchies. En línea:

<http://www.getty.edu/vow/AATHierarchy?find=thesaurus&logic=AND&note=&subjectid=300000000>

GÓMEZ-Pérez, Asunción, FERNÁNDEZ-López, Mariano and CORCHO, Oscar (2004) Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web, Facultad de Informática, Universidad Politécnica de Madrid, España. ©Springer-Verlag London.

GÓMEZ Pérez, Asunción; SUÁREZ de Figueroa Baonza, Mari Carmen; VILLAZÓN, Boris (2008) Neon Methodology for Building Ontology Networks: Ontology Specification. En línea: <http://www.neon-project.org>

NOY, Natalya F. and MCGUINNESS, Deborah L. (2005) Desarrollo de Ontologías-101: Guía Para Crear Tu Primera Ontología. [noy@smi.stanford.edu](mailto:noy@smi.stanford.edu) and [dml@ksl.stanford.edu](mailto:dml@ksl.stanford.edu), Stanford University, Stanford, CA, 94305, Traducido del inglés por: Erick Antezana.

RAMOS, Esmeralda y NUÑEZ, Haydemar (2007) Ontologías: componentes, metodologías, lenguajes, herramientas y aplicaciones, Universidad Central de Venezuela, Facultad de Ciencias, Escuela de Computación. Lecturas en Ciencias de la Computación ISSN 1316-6239.

INSTITUTO DE LENGUA Y CULTURA AYMARA, (2009-2013) Materiales del proyecto, "Registro Textil, Equivalencias, Taxonomías", La Paz, ILCA, Inédito.

YUJRA Saire, Efrain (2012) Experiencias en el manejo de información del dominio del Textil Andino en el desarrollo de una ontología, ponencia dirigida a curadores de museos, Londres, ILCA, Inédito.