

UNIVERSIDAD TÉCNICA DE AMBATO



CENTRO DE POSGRADOS

PROGRAMA DE MAESTRÍA EN BIBLIOTECOLOGÍA Y ARCHIVOLOGÍA

COHORTE 2021

TEMA: LENGUAJE CONTROLADO Y LENGUAJE NATURAL EN LA
RECUPERACIÓN DE LA INFORMACIÓN EN LAS BIBLIOTECAS
UNIVERSITARIAS

Trabajo de Titulación, previo a la obtención del Grado Académico de Magíster en
Bibliotecología y Archivología

Modalidad del Trabajo de Titulación: Proyecto de Titulación con Componente de
Investigación Aplicada

Autora: Licenciada Nube del Rocío Salto Morquecho

Directora: Licenciada Cristina Alexandra Mantilla Sandoval Magister.

Ambato – Ecuador

2023

A la Unidad Académica de Titulación del Centro de Posgrados

El Tribunal receptor del Trabajo de Titulación, presidido por el Ingeniero Héctor Fernando Gómez Alvarado. PhD, e integrado por los señores: Ingeniera María José Bravo Ramos Magíster, y Licenciada María del Carmen Gaybor Orbea Magíster, designados por la Unidad Académica de Titulación del Centro de Posgrados de la Universidad Técnica de Ambato, para receptor el Trabajo de Titulación con el tema: “*Lenguaje controlado y lenguaje natural en la recuperación de la información en las bibliotecas universitarias*”, elaborado y presentado por la señora Licenciada Nube del Rocío Salto Morquecho, para optar por el Título de Cuarto Nivel de Magíster en Bibliotecología y Archivología; una vez escuchada la defensa oral del Trabajo de Titulación, el Tribunal aprueba y remite el trabajo para uso y custodia en las bibliotecas de la UTA.

Ing. Héctor Fernando Gómez Alvarado. PhD.
Presidente y Miembro del Tribunal

Ing. María José Bravo Ramos Mgs.
Miembro del Tribunal

Ing. María del Carmen Gaybor Orbea Mgs.
Miembro del Tribunal

AUTORÍA DEL TRABAJO DE TITULACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el Trabajo de Titulación presentado con el tema: LENGUAJE CONTROLADO Y LENGUAJE NATURAL EN LA RECUPERACIÓN DE LA INFORMACIÓN EN LAS BIBLIOTECAS UNIVERSITARIAS, le corresponde exclusivamente a: la Licenciada Nube del Rocío Salto Morquecho, Autora bajo la Dirección de la Licenciada Cristina Alexandra Mantilla Sandoval Magister Directora del Trabajo de Titulación, y el patrimonio intelectual a la Universidad Técnica de Ambato.

Licenciada Nube del Rocío Salto Morquecho
c.c.: 0301770665

AUTORA

Licenciada Cristina Alexandra Mantilla Sandoval Magister
c.c.:1802039030

DIRECTORA

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que el Trabajo de Titulación, sirva como un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los Derechos de mi trabajo, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad.

Licenciada Nube del Rocío Salto Morquecho
c.c.: 0301770665

INDICE GENERAL DE CONTENIDOS

Portada.....	i
A la Unidad Académica de Titulación del Centro de Posgrados	ii
AUTORÍA DEL TRABAJO DE TITULACIÓN.....	iii
DERECHOS DE AUTOR	iv
INDICE GENERAL DE CONTENIDOS	v
ÍNDICE DE TABLAS.....	vii
ÍNDICE DE FIGURAS.....	viii
AGRADECIMIENTO	ix
DEDICATORIA.....	x
RESUMEN EJECUTIVO	xi
CAPÍTULO I.....	1
EL PROBLEMA DE INVESTIGACIÓN	1
1.1. Introducción	1
1.2. Justificación	2
1.3 Objetivos	3
1.3.1. General.....	3
1.3.2. Específicos	3
CAPITULO II.....	4
2.1. ANTECEDENTES INVESTIGATIVOS.....	4
2.2. Tipos de lenguaje de Indización	8
2.2.1. Según el control.....	10
2.2.2. Según la coordinación	12
2.2.3. Según la estructura.....	12
2.2.4. Según su especialización.....	13
2.3. Clasificación del lenguaje documental.....	14
2.3.1. Lenguaje natural.....	14
2.3.2. Lenguaje controlado	18
2.3.3. Tesoros	21
2.4. Recuperación de la información	23
2.4.1. Etapas del proceso de búsqueda de información	24
2.4.2. Técnicas de búsqueda	27
2.4.3. Los metadatos	31

CAPITULO III.....	34
MARCO METODOLÓGICO.....	34
3.1. Ubicación.....	34
3.2. Equipos y materiales	35
3.3. Tipo de investigación	36
3.4. Idea a defender.....	37
3.5. Población o muestra	37
3.6. Recolección de información	37
3.6.1. Validación de los instrumentos.....	38
3.7. Procesamiento de la información y análisis estadístico.....	38
3.8. Variables respuesta o resultados alcanzados.....	38
CAPITULO IV.....	39
RESULTADOS Y DISCUSIÓN.....	39
4.1. Análisis de los resultados	39
4.2. Interpretación de datos	39
CAPÍTULO V.....	61
CONCLUSIONES, RECOMENDACIONES, BIBLIOGRAFÍA Y ANEXOS	61
5.1. Conclusiones	61
3.2. Recomendaciones	63
3.3. BIBLIOGRAFÍA	64
3.4. ANEXOS	67

ÍNDICE DE TABLAS

Tabla 1 Lenguajes documentales según Slype	8
Tabla 2 Lenguajes documentales según Iglesias	9
Tabla 3 Lenguajes documentales según Gil Urdiciain	9
Tabla 4 Lenguajes documentales según Chaumir	9
Tabla 5 Lenguajes documentales según Sánchez Luna	10
Tabla 6 Ventajas y desventajas de los lenguajes controlado y natural	15
Tabla 7 Población a ser encuestada	37
Tabla 8 Sexo, edad e instrucción académica	39
Tabla 9 Biblioteca en la que labora	40
Tabla 10 Años de experiencia	41
Tabla 11 Denominación del puesto	43
Tabla 12 Plataforma de gestión bibliográfica	44
Tabla 13 Tipos de registros bibliográficos	46
Tabla 14 Actividades de Indización.....	47
Tabla 15 Uso del lenguaje documental	48
Tabla 16 Importancia de la indización	49
Tabla 17 Instrumentos de indización	50
Tabla 18 Dificultades de indización	52
Tabla 19 Frecuencia de capacitaciones	53
Tabla 20 Interoperabilidad	54
Tabla 21 Lenguaje documental recomendado	55
Tabla 22 Unificación de lenguaje documental	56
Tabla 23 Instrumento de indización.....	57
Tabla 24 Recuperación de la información.....	58
Tabla 25 Cosecha de información	59

ÍNDICE DE FIGURAS

Figura 1 Proceso de búsqueda de información	24
Figura 2 Técnicas de búsquedas	28
Figura 3 Operadores booleanos	30
Figura 4 Ubicación Universidad de Cuenca	34
Figura 5 Biblioteca en la que labora	41
Figura 6 Años de experiencia	42
Figura 7 Denominación del puesto	43
Figura 8 Plataforma de gestión bibliotecaria	45
Figura 9 Tipos de registros bibliográficos.....	46
Figura 10 Actividades de Indización	48
Figura 11 Uso del lenguaje documental	49
Figura 12 Importancia de la indización.....	50
Figura 13 Instrumentos de indización.....	51
Figura 14 Dificultades de indización	52
Figura 15 Frecuencia de capacitación.....	53
Figura 16 Interoperabilidad.....	54
Figura 17 Lenguaje documental recomendado	55
Figura 18 Unificación de lenguaje documental	56
Figura 19 Instrumento de indización	57
Figura 20 Recuperación de la información	58
Figura 21 Cosecha de información	59

AGRADECIMIENTO

A Dios por prestarme la vida y salud, para poder llegar a cumplir mi sueño tan anhelado de especializarme en la rama que siempre me he desempeñado. A mi familia que siempre son mi razón de ser, el motivo para mi superación, por todo el apoyo incondicional que siempre me brindan, a mi madre Carmen, mi esposo Eduardo y mis tres bellos hijos Angy, Jeremy y Amy. De igual forma agradezco a las universidades: Universidad Técnica de Ambato por brindarme la formación académica de manera especial a la licenciada Cristina Mantilla mi directora y a la Universidad de Cuenca a través del Centro de Documentación Regional Juan Bautista Vázquez, quienes me permitieron realizar mi trabajo de investigación.

DEDICATORIA

El presente Trabajo de Titulación, lo dedico a Dios por haberme dado la vida y nunca haberme dejado sola en mis instantes de dificultad. A mis colegas bibliotecarios que aman la profesión.

UNIVERSIDAD TÉCNICA DE AMBATO
CENTRO DE POSGRADOS
MAESTRÍA EN BIBLIOTECOLOGÍA Y ARCHIVOLOGÍA
COHORTE 2021

TEMA:

LENGUAJE CONTROLADO Y LENGUAJE NATURAL EN LA RECUPERACIÓN DE LA INFORMACIÓN EN LAS BIBLIOTECAS UNIVERSITARIAS

MODALIDAD DE TITULACIÓN: *Proyecto de Titulación con Componente de Investigación Aplicada*

AUTOR: *Licenciada Nube del Rocío Salto Morquecho*

DIRECTOR: *Licenciada Cristina Alexandra Mantilla Sandoval, MSc.*

FECHA: *Ocho de marzo de dos mil veinte y tres*

RESUMEN EJECUTIVO

La presente investigación tiene por objetivo analizar los diferentes tipos de lenguajes documentales, para facilitar la recuperación de la información en las bibliotecas universitarias. Sabiendo que los lenguajes documentales son el puente directo que establece la comunicación entre el documento original y los usuarios ya que relaciona la información contenida en el documento y la información requerida por el investigador. Es por ello que en esta investigación se aborda los temas como lenguaje natural, lenguaje controlado, su clasificación, las tipologías, también se habla sobre los sistemas de recuperación de la información.

Con la finalidad de analizar y conceptualizar los diferentes tipos de lenguaje documental se revisó diversas fuentes bibliográficas como libros, trabajos de titulación, artículos académicos, páginas web de Bibliotecas extranjeras, entre otros sitios de autores reconocidos y expertos en la temática como son Naumis Peña (2006), Gil Urdiciain (1996), Martín Gavilán (2011) Slype (1991), estos profesionales en sus documentos tienen una estrecha relación en cuanto al análisis de los lenguajes documentales y en sí la indización documental.

El presente trabajo tiene un enfoque de investigación mixta, cualitativa y cuantitativa debido a que para el desarrollo del capítulo dos se implicaron un proceso de análisis

de las visiones y concepciones en torno a los lenguajes documentales y la recuperación de la información. Para la consecución de los objetivos se planteó un abordaje desde la investigación documental y la investigación aplicada. Además, se utilizó la técnica de encuesta aplicando un instrumento validado por expertos en la temática, a los profesionales que laboran en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca. Entre las recomendaciones más evidentes se puede anotar la capacitación permanente al personal con una visión a crear un tesoro propio para el Centro, para normalizar la terminología que se utiliza en el CDR.JBV y facilitar la recuperación de la información

DESCRIPTORES: *BIBLIOTECAS UNIVERSITARIAS, LENGUAJE CONTROLADO, INDIZACIÓN, LENGUAJE DOCUMENTAL, LENGUAJE NATURAL, TESAURO*

CAPÍTULO I

EL PROBLEMA DE INVESTIGACIÓN

1.1. Introducción

El lenguaje controlado y el lenguaje natural son herramientas de la indización, esta forma parte del análisis documental ya que es una de las actividades fundamentales que se realiza en cualquier unidad de información, ya que consiste en la representación exacta de los documentos originales, cuya finalidad es que el sistema de información proporcione solamente aquellos documentos que sean relevantes de acuerdo con las necesidades informativas del usuario y de esta manera, hacerlos más fácilmente accesibles

La presente investigación titulada “Lenguaje controlado y Lenguaje natural para la recuperación de la información en las bibliotecas universitarias”, tiene por objetivo general analizar los diferentes tipos de lenguajes documentales, para la recuperación de la información en las bibliotecas universitarias. Para lo cual se recupera bibliografía, analiza y conceptualiza la tipología del lenguaje documental, como también los diversos tipos de sistemas de recuperación de la información que se utiliza en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca.

Esta investigación está estructurada por cinco capítulos de la siguiente manera. En el primer capítulo consta el planteamiento del problema, con su debida justificación y los objetivos planteados. El segundo capítulo aborda definiciones, conceptos y clasificaciones tanto de lenguaje documental como de la recuperación de la información. En el tercer capítulo se aborda la metodología empleada para la realización de la investigación. En el cuarto capítulo se analiza y discute los resultados obtenidos de las encuestas realizadas a los profesionales que laboran en las bibliotecas universitarias del CDR.JBV, por último, en el quinto capítulo aborda las conclusiones y recomendaciones que surgen de la investigación realizada.

El CDR.JBV de la Universidad de Cuenca cuenta con tres bibliotecas especializadas y una general por lo que es necesario conocer si la información procesada en el Centro es recuperada de manera oportuna ya que en la biblioteca del campus Central que es una biblioteca general se encuentra centralizado el Área de Normalización y Procesos Técnicos, pero en cada Campus hay personal que también realiza la indización del material bibliográfico.

1.2. Justificación

El presente trabajo de investigación se basa en la necesidad de conocer el uso del lenguaje documental para la indización de la información y por ende la interoperabilidad y recuperación de la información disponible en la web como es el caso de las bibliotecas universitarias que gestionan y a su vez brindan acceso a la información ya sea de las bibliotecas digitales o híbridas, catálogos, repositorios institucionales, bases digitales entre otros, se encuentran indizadas con distintos sistemas de organización del conocimiento también conocido como lenguaje documental, entre ellos tenemos listas de encabezamientos de materia, tesauros, sistemas de clasificación o en su defecto el lenguaje libre, esto hace que los usuarios realicen las búsquedas en cada sistema de información sin tener resultados favorables, pérdida de tiempo y esfuerzo.

Conociendo esta problemática se ve la necesidad de analizar el lenguaje documental que se utiliza en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, ya que al momento de indizar se usa únicamente el Tesauro de la Unesco, al igual que en las bibliotecas de los campus a pesar de que son bibliotecas especializadas en Agronomía, Veterinaria, Artes y Ciencias de la hospitalidad, en la biblioteca de Medicina utilizan los descriptores en Ciencias de la Salud - DeCS, por esta razón existe la necesidad analizar y definir los tipos de lenguajes documentales, los listados de descriptores libres, lenguaje natural y lenguaje controlado y dar a conocer a los profesionales de la información cuales son las ventajas y desventajas de usar correctamente el lenguaje documental y el beneficio al momento de la recuperación de la información.

1.3 Objetivos

1.3.1. General

Analizar los diferentes tipos de lenguajes documentales, para la recuperación de la información en las bibliotecas universitarias.

1.3.2. Específicos

- Definir teóricamente los diferentes tipos de lenguajes documentales, que facilita la recuperación de la información.
- Analizar las herramientas que utiliza el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, para la recuperación de la información.
- Determinar la relación que existe en el manejo del lenguaje controlado y natural que facilite la recuperación de la información en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca.

CAPITULO II

2.1. ANTECEDENTES INVESTIGATIVOS

El lenguaje documental es un producto de construcción compleja con un alto grado de subjetividad, en la actualidad tanto bibliotecas como archivos utilizan los metalenguajes para indizar y describir la información de cada uno de los materiales bibliográficos para establecer la comunicación entre el documento original y el usuario, por lo tanto, en este trabajo de investigación se pretende definir teóricamente, analizar y determinar la relación existente entre el lenguaje natural, el lenguaje controlado y la recuperación de la información por lo que a continuación citaré algunos artículos en los que mencionan sobre temas similares como también la terminología necesaria para el desarrollo de esta investigación.

Revisando literatura sobre el tema, Soler Monreal (2009) en su tesis doctoral sobre evaluación de vocabularios controlados en la indización de documentos, cuyo objetivo fue comprobar el comportamiento de varios vocabularios controlados en la indización de documentos mediante índices de consistencia entre indizadores, y en segundo lugar, comparar los tesauros con nuevas herramientas que en los últimos años están emergiendo con gran fuerza como las ontologías, las folksonomías y las taxonomías, para cumplir con los objetivos tuvo que construir tres vocabularios controlados sobre Biblioteconomía y Documentación.

Una vez construidos los vocabularios controlados fueron sometidos a distintas pruebas con diferentes indizadores, tanto estudiantes como profesionales a fin de comprobar la eficacia de dichos vocabularios a través de índices de consistencia, llegando a tener los siguientes resultados tanto para estudiantes como para los profesionales, el listado de descriptores obtiene mejores resultados de consistencia en la indización, seguido del tesoro aumentado y el tesoro casi en la misma proporción.

En un artículo de Mochón-Bezares *et al.* (2017) con el tema: Etiquetado social y *blog-scraping* como alternativa para la actualización de vocabularios controlados: aplicación práctica a un tesoro de Biblioteconomía y Documentación, cuyo objetivo

de este artículo es comparar las etiquetas en lenguaje libre, frente al lenguaje controlado no estructurado de las listas de palabras clave, con el fin de comprobar cuál de estos dos es una mejor fuente de nueva terminología para el Tesoro de Biblioteconomía y Documentación. Luego de analizar las etiquetas llega a la siguiente conclusión que las etiquetas en lenguaje libre pueden ser una mejor y más rápida vía de aporte de nueva terminología a los vocabularios controlados que los listados de lenguaje controlado no estructurado.

García-Jiménez *et al.* (2019) en su estudio sobre la indización /etiquetado y los lenguajes documentales en cinco diarios españoles menciona que en esta investigación parte de la preocupación sobre el modo en el que se organiza el conocimiento en los diarios españoles. Tiene un doble objetivo, por un lado, determinar cómo se desarrolla el análisis, indización y etiquetado de los contenidos periodísticos y por otro es la delimitación de la existencia, características y aplicación operativa de lenguajes documentales. Se observa divergencias en las tareas de indización y etiquetado, la conexión con técnicas SEO (Search Engine Optimization), así como la continuación en el empleo de lenguajes documentales, si bien no en todos los casos, llegando a la conclusión general, se puede afirmar que los servicios de documentación que permanecen en los periódicos analizados están centrados, fundamentalmente, en tareas de recuperación de información.

Luego de analizar la literatura antes citada puedo manifestar que es muy importante usar el lenguaje documental para la recuperación de la información, puesto que el lenguaje natural puede minimizar la cantidad de documentos recuperados ya que el uso incorrecto de términos en la indización de la documentación produce el llamado silencio de información, para profundizar más en los temas a investigar a continuación se conceptualiza la tipología de los lenguajes documentales como también los sistemas de recuperación de información.

2.1. Lenguajes documentales

Los lenguajes documentales son la base fundamental para la indización de documentos en las bibliotecas universitarias, por lo tanto, bibliotecarios, documentalistas y profesionales de la información deben conocer de su existencia, funcionalidad y aplicar en sus tareas diarias, para entender mejor el concepto de lenguaje documental a continuación citaré algunos conceptos de autores que hablan sobre lenguaje documental tanto en bibliotecología, archivología y las ciencias de la información.

Martín Gavilán (2009) aporta lo siguiente: “Los lenguajes documentales son, pues, los que establecen la comunicación entre el documento original y el usuario, haciendo de puente o intermediario entre las informaciones contenidas en los documentos y las informaciones solicitadas por los usuarios”. En este contexto se puede mencionar que en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, el uso de lenguaje documental se realiza de una manera híbrida puesto que se apoya en el tesoro de la Unesco, como también en los términos del vocabulario controlado de la Frascati, pero la mayor parte de términos son de lenguaje natural.

De nuestra lengua natural escrita con todas sus reglas, hacemos cortes o segmentos por áreas de interés a los cuales hacemos un análisis para saber su contenido, de qué tratan, y entonces es cuando surgen los lenguajes documentales, en este sentido comparto la siguiente aportación de Slype (1991) “Un lenguaje documental es un sistema de signos que permite representar el contenido de los documentos con el fin de recuperar documentos pertinentes en respuesta a consultas que tratan sobre ese contenido” (p. 21)

Otra forma de conceptualizar los lenguajes documentales comenta y aclara el mismo Slype, diciendo que para algunos autores un lenguaje documental es un lenguaje artificial. Para este estudio considero el concepto de Slype que toma a los lenguajes documentales como parte del lenguaje natural y que Naumis está de acuerdo y lo expresa en los siguientes párrafos.

Naumis (2007) como se citó en Martín Gavilán (2011) expresa lo siguiente:

Los lenguajes documentales se construyen en diferentes lenguas, para interactuar en el proceso documental, al permitir indizar y recuperar información con ciertas unidades léxicas aisladas de un lenguaje general que representan un cuerpo de conocimientos. Otro de sus enunciados es los lenguajes documentales se crean generalmente sobre los términos de comunicación especializada, empleados en los diferentes campos del conocimiento.

Otros autores como Martínez y Murillo (2009) definen al lenguaje documental como se cita a continuación:

El lenguaje documental es un lenguaje intermediario, en la medida en que sirve de puente entre las informaciones contenidas en los documentos y las informaciones solicitadas por los usuarios. El tratamiento uniforme de la información permite generar la representación formalizada del contenido de los documentos y, así, facilitar la recuperación, manual o automática, de esta información. La elaboración de un lenguaje documental apropiado es clave para la consecución del flujo informativo (p. 61).

En conclusión, puedo mencionar que el lenguaje documental es el lenguaje controlado y el lenguaje natural son aquellas palabras que usamos sin control o normalización para indizar y esta presenta problemas como las polisemias, las homonimias que dificultan la recuperación de la información, cabe mencionar que el lenguaje natural si se utiliza en algunas ocasiones y por lo tanto el lenguaje documental surge para restar inconvenientes al momento de indizar, se debe también indicar que no es lo mismo lenguaje documental que lenguaje de indización, en este caso se aclara que todo los lenguajes documentales son lenguajes de indización como las listas de palabras claves y descriptores libres, pero no todos los lenguajes de indización son lenguajes documentales, por lo tanto es importante analizar el lenguaje documental resaltando como es en este caso en esta investigación el lenguaje controlado frente al lenguaje natural para ello se analizará la siguiente tipología.

2.2. Tipos de lenguaje de Indización

Existen varias formas de clasificar el lenguaje de indización, su función es la correcta indización para recuperar la información, autores como Naumis Peña (2006), Gil Urdiciain (1996), Martín Gavilán (2011) Slype (1991) utilizan términos como lenguajes de indización, lenguajes terminológicos, lenguajes controlados, lenguajes normalizados, lenguaje documentario, y lenguaje de búsqueda de información, además de lenguajes documentales. En esta investigación se utilizará el término lenguajes documentales porque es la designación más frecuente en la literatura bibliotecológica por las autoridades en la materia.

A continuación, se describe la clasificación tipológica del lenguaje de indización de acuerdo a diversos autores:

Slype (1991) clasifica el lenguaje documental de la siguiente manera:

Tabla 1

Lenguajes documentales según Slype

	Lenguajes de indización	Lenguajes libres
Lenguajes documentales	Lenguajes de clasificación	Clasificación Decimal Dewey
		Clasificación Decimal Universal
		Clasificación de la Biblioteca del Congreso
		Clasificación Facetada de Ranganathan
		Clasificación Internacional de Patentes

Iglesias Maturana (1997) considera la siguiente clasificación:

Tabla 2

Lenguajes documentales según Iglesias

Indización automática	Concordancias
	Índices permutados
	Índice KWIC
	Índice KWOC
	Indización por citas
Lenguajes de indización	Esquemas de clasificación
	Tesauros
	Listas de encabezamientos de materia
	Listados de palabras claves

Gil Urdiciain (1996) clasifica de la siguiente manera:

Tabla 3

Lenguajes documentales según Gil Urdiciain

Lenguajes documentales	Lenguajes libres	Listas de descriptores libres
	Lenguajes controlados	Clasificaciones
		Tesauros

Chaumier (1986) clasifica al lenguaje documental de la siguiente manera:

Tabla 4

Lenguajes documentales según Chaumier

Lenguaje documental	Lenguajes evolucionados	Transforma los terminos del texto
	Lenguajes de descripción de datos	
	Lenguajes controlados	Lista de “vedettes” Tesauros
	Clasificaciones	Clasificaciones de tipo decimal Clasificaciones de facetar

Sánchez Luna (2012) clasifica al lenguaje documental de acuerdo a la siguiente tabla:

Tabla 5
Lenguajes documentales según Sánchez Luna

Control	Libres	Listas de descriptores libres
	Controlados	Clasificaciones, tesauros, catálogos de autoridad
Coordinación	Precoordinados	Clasificaciones Listas de encabezamientos de materia
	Poscoordinados	Listas de descriptores libres Listas de palabras clave - Tesauros
Estructura	Jerárquica	Clasificaciones jerárquicas
	Combinatoria	Léxicos documentales - tesauros
	Sintáctica	De gramática simple De gramática elaborada

Luego de haber colocado la clasificación tipológica del lenguaje documental a continuación se explica cada uno de ellos.

2.2.1. Según el control

De acuerdo al control se puede mencionar los siguientes:

- **Libres**

El lenguaje libre denominado también lenguaje natural no es un lenguaje propiamente denominado lenguaje documental por carecer de vocabulario controlado, a estos corresponden las listas de descriptores libres y las listas de palabras claves, en este contexto coincide con Linares donde explica el uso de lenguaje natural al momento de indizar y los problemas que ocasiona.

Rubio Liniers (2020) menciona lo siguiente con respecto al lenguaje libre:

Está cargado de metáforas, sinonimias y homonimias y necesita ser organizado y normalizado para permitir una recuperación eficaz del documento a salvo de los dos grandes problemas documentales: el ruido, -exceso de información encontrada no pertinente - y el silencio - ausencia de información pertinente que existe en la base de datos y no es recuperada (p. 3).

También Rubio Liniers (2020) menciona que los principales problemas del lenguaje natural son:

Normalización de singulares y plurales. Para dicho tema la indización debe ajustarse en lo posible a la Norma UNE 50-106-90. Normas para el desarrollo de tesauros monolingües. Además, el significado de muchos términos varía según su número (Derecho; derechos/Obligación; Obligaciones).

Ambigüedad y falta de especificidad. La utilización de unitérminos, propio del lenguaje natural crea además ambigüedad y falsas combinaciones en la recuperación. Según el ejemplo anterior un buen sistema controlado debería recoger formas como: Obligación de declarar; Obligaciones del Estado (p. 2).

- **Controlados.**

Al hablar de los lenguajes controlados se entiende que son los lenguajes documentales dentro de estos podemos encontrar los tesauros, las listas de encabezamientos de materia y las clasificaciones, estos presentan un vocabulario elaborado con anterioridad que nos ayuda al momento de la indización y por ello a la recuperación de la información, he podido analizar que coinciden en esta tipología de lenguaje controlado tanto Van Slype (1991) y Martín Gavilán (2011) además Martín Gavilán enuncia lo siguiente “El vocabulario controlado proporciona al usuario un punto de búsqueda, en vez de dos o más, y reduce la posibilidad de que la búsqueda sea incompleta. Sin embargo, puede perderse alguna información debido a la falta de especificidad y errores en el análisis” (p. 3).

2.2.2. Según la coordinación

- **Precoordinados**

Gil Urdiciain (1996) dice que en los sistemas precoordinados los términos están unidos de acuerdo con unas reglas sintácticas establecidas por el propio lenguaje. Los lenguajes precoordinados contribuyen mejor que los postcoordinados a conseguir precisión en la búsqueda, es decir, tienen mayor capacidad para rechazar los documentos irrelevantes, mismos que producen ruido documental en el momento de la recuperación porque las relaciones que se crean mediante el álgebra de Boole, o los conocidos operadores booleanos, son genéricas y no impiden las falsas combinaciones por lo que se pierde la información.

- **Postcoordinados.**

Gil Urdiciain (1996) considera que el lenguaje postcoordinado carece de sintaxis; salvo excepciones, su vocabulario consiste en términos simples o unitérminos que se pueden conectar en el momento de la recuperación por medio de una sintaxis booleana o conocidos como operadores booleanos. Atendiendo a los procedimientos seguidos para relacionar los conceptos al recuperar la información, la autora llega a la conclusión de que en los lenguajes precoordinados la relación entre los términos es gramatical, mientras que en un lenguaje postcoordinado la relación es lógica. Como contrapartida, los postcoordinados tienen la ventaja de dar respuesta a necesidades de indización reales.

2.2.3. Según la estructura

En cuanto a la estructura, se reconoce la existencia de tres tipos: jerárquica, combinatoria y sintáctica

- **Jerárquica**

Gil Urdiciain (1996) sostiene que los lenguajes de estructura jerárquica organizan sistemáticamente el conocimiento, es decir, agrupan las materias en áreas categoriales.

A la estructura jerárquica responden determinadas clasificaciones, como se puede enumerar a la clasificación jerárquica de Dewey o la CDU, donde los conceptos se distribuyen en clases o categorías y dependen los unos de los otros, estos tipos de lenguajes van de lo general a lo particular, resultan poco efectivos en la recuperación de la información, estos se representan por códigos alfabéticos, numéricos o alfanuméricos que facilitan su localización en la estructura arborescente.

- **De estructura combinatoria y asociativa.**

Para Gil Urdiciain (1996) esta tipología permite la libre combinación de los términos que los componen entre sí, eliminando la rigidez de los lenguajes de estructura jerárquica según las necesidades de indización. Las listas de encabezamientos de materia y los tesauros corresponden a esta categoría, si bien es cierto que en la organización multidimensional de los tesauros participa también la estructura jerárquica, estos lenguajes tienen una presentación alfabética y los conceptos se asocian o combinan libremente entre sí para definir el tema o temas a utilizar.

- **De estructura sintáctica**

En concordancia y análisis sobre este tema algunos autores concuerdan con Gil Urdiciain (1996) quien menciona que los lenguajes documentales de estructura sintáctica surgen por la cantidad de ruido que producen los lenguajes combinatorios, que pueden poseer una red de relaciones semánticas, pero no tienen gramática. La estructura sintáctica permite concretar y ordenar los conceptos. Este tipo de lenguaje requiere un cierto grado de automatización en los procesos tanto de indización como de recuperación, por lo que su aparición está asociada al desarrollo tecnológico. Según su complejidad, se pueden diferenciar dos tipos: lenguajes sintácticos de gramática simple y lenguajes sintácticos de gramática elaborada.

2.2.4. Según su especialización

El lenguaje documental según la especialización puede clasificarse en general y en especializado, esto es más utilizado en los archivos como también para las bibliotecas.

- **Índices generales**, aquí podemos encontrar las clasificaciones universales, que intentan abarcar todo el conocimiento humano.
- **Índices especializados**, estos se refieren o están restringidos a una rama determinada del saber, como se observa en el caso de los tesauros especializados.

2.3. Clasificación del lenguaje documental

2.3.1. Lenguaje natural

Se conoce como lenguaje natural aquel lenguaje que se utiliza para describir un documento, el personal de la información en algunas ocasiones realiza la indización utilizando términos que no tienen ningún tipo de control ni normalización, esto dificulta al momento de recuperar la información produciendo el llamado silencio informacional, o también puede darse que se recupera documentos que no tengan ninguna relación con la búsqueda, provocando el ruido documental.

Pinto Molina (1995) manifiesta que el lenguaje natural es aquel en el que están escritos todos los documentos primarios, sometidos al análisis documental. Es también el lenguaje hablado y el que se emplea en prácticamente todas las informaciones originales, cualquiera que sea su soporte. Todos nos expresamos en una amplia variedad de lenguajes naturales para comunicarnos con los demás y con ellos manifestamos ideas, sentimientos y emociones, con una enorme multiplicidad de matices, sin atender a cuestiones de espacio o de tiempo o al tipo y cantidad de términos que han de ser empleados para la comunicación que con ellos pretendemos conseguir.

Martínez y Murillo (2009) sostiene que el lenguaje natural conlleva a que aparezca la ambigüedad, que crea dificultades en la recuperación y organización de la información. Surgen entonces accidentes lingüísticos propios del lenguaje natural como son la homonimia y la sinonimia. La primera provoca ruido en la búsqueda, y la segunda provoca el llamado silencio documental. Sólo el logro de la deseable entropía

puede evitar estos inconvenientes y facilitar una recuperación eficaz de la información.
(p.73)

Para Estrada Villacís (2015) el lenguaje natural se usa para la comunicación inmediata, la que se realiza ahora mismo, coexisten diferentes significados para un sola palabra o diversas palabras sinónimas, en este contexto esta autora se plantea la siguiente interrogante: ¿Y por qué hay que controlar los términos del lenguaje natural? y su respuesta es la siguiente, porque el lenguaje natural es ambiguo, los conceptos se pueden representar de formas diversas, dando lugar a problemas de recuperación. Este tipo de lenguaje es usado hoy en día para buscar información en línea y la web semantica trata de ayudar a recuperar la información, en este caso se debería conocer bien del tema para poder obtener los resultados de la búsqueda.

El lenguaje natural es rico en terminología, en formas (plurales y singulares), tiempos verbales, acrónimos, sinónimos, polisemias, etc. La principal diferencia entre el lenguaje natural y el documental controlado es precisamente el control terminológico, que permite representar los conceptos de forma unívoca, sin ambigüedades. Para ser más concretos, las diferencias se dan en el número de términos del vocabulario, el control de las formas, el control del significado y las relaciones de significado entre términos, a continuación agrego un cuadro de las ventajas y las desventajas de usar el lenguaje natural y el lenguaje controlado.

Tabla 6
Ventajas y desventajas de los lenguajes controlado y natural

Lenguaje controlado	Lenguaje natural o libre
Ventajas	
Representación lógica y sistemática	Rapidez en la representación
Reduce la ambigüedad	Lenguaje cercano al ausuario
Aporta precisión	Fácil actualización
Ayuda en el análisis	Funcionales en OPACS
Desventajas	
Insumen más tiempo en la representación	Alto grado de ambigüedad

Problemas en su actualización	Distorsión en la comunicación
Lenguaje distante al usuario	Dificultades en su control
Inconvenientes para su uso en OPACS	Búsquedas fallidas ruidos y silencios

Fuente: Mónica Elizabeth Estrada Villacís

Slype (1991) también menciona que existen dos tipos principales de lenguajes combinatorios libres: las listas de palabras clave y las listas de descriptores libres.

Lista de palabras clave una lista de palabras clave está constituida por una colección no ordenada (sino puesta por orden alfabético) de las palabras significativas, denominadas también no vacías (es decir, todas las palabras que no son artículos, conjunciones, pronombres, preposiciones, numerales y ciertos verbos y adverbios), extraídas, de forma automática, por el ordenador, a partir del título, del resumen y, cada vez más a menudo, del texto completo de los documentos registrados dentro de un sistema documental dado.

Ejemplos: biblioteca; documentación; documental.

La lista de palabras clave es, la mayoría de las veces, monolingüe; igualmente puede ser plurilingüe, es decir, puede contener palabras de dos o más lenguas, pero sin equivalencias entre las palabras de las diferentes lenguas.

Lista de descriptores libres una lista de descriptores está constituida por una colección no ordenada (sino puesta por orden alfabético) de conceptos destacados, por un proceso intelectual, a partir de los documentos registrados dentro de un sistema documental dado; esos conceptos son expresados por palabras o por expresiones extraídas de los documentos, o propuestos por los documentalistas, sin verificar si existen previamente en una lista establecida a priori.

Ejemplos: biblioteca; servicio de documentación; servicio documental. La lista de descriptores libres es generalmente monolingüe (p. 1).

El uso de lenguaje libre o natural también puede conllevar a que existe el llamado silencio documental como también el ruido documental, ya que los términos no están

normados, y por ende también se puede indicar que este lenguaje da paso a la conocida sinonimia, polisemia, homonimia, que se detallará a continuación.

- **Sinonimia**

El uso de los sinónimos afecta en gran tamaño la recuperación de la información en los sistemas donde se almacena la información ya sean estos bases de datos, repositorios o catálogos, los términos de un tema se deben conceptualizar bien para poder indizar un documento, se debe tener claro que cada termino debe tener un significado y cada significado nos debe llevar a un único término, de esta manera coincido con Rubio Liniers (2020) quien menciona los siguiente "la utilización de sinónimos, que aporta valor y riqueza desde el punto de vista literario es el mayor problema en la recuperación en texto libre", recomienda el uso de los tesauros para evitar la sinonimia y para una mejor acepción coloca el siguiente ejemplo.

Enfermedad laboral

Usado por Enfermedad profesional

Enfermedad profesional

Use Enfermedad laboral

- **Homonimia**

Para Juncà Campdepadrós y Martínez Ferreras (2009) la polisemia y variantes provocan ruido documental, por lo que a continuación se presenta su explicación.

La homonimia se da cuando dos conceptos han llegado a tener el mismo nombre, la misma forma, pero vienen de orígenes diferentes y, por lo tanto, tienen etimologías diferentes. Por ejemplo, metro puede ser el transporte urbano, una unidad de medida o el utensilio para medir. Pero el origen etimológico entre el transporte y los otros dos significados es evidente: el primero es una abreviación de la palabra inglesa metropolitan, y en el segundo caso viene del griego μέτρον y significa medida. Otro ejemplo: la palabra castellana botín puede venir del latín bota y significará ‘calzado hasta el tobillo’, o puede venir del alemán bytin y significará ‘premio de una conquista’.

En castellano y catalán este fenómeno es menos frecuente que en otras lenguas, como el inglés o el francés, en las que abundan las palabras homónimas que dan mucho juego en los chistes. Dentro de la homonimia podemos diferenciar las palabras que escribiéndose igual tienen significados diferentes, llamadas homógrafas, como las anteriores metro o botín, de las palabras que sonando igual también tienen significados diferentes, conocidas como palabras homófonas: vell/bell en catalán, o tubo/tuvo en castellano (p. 24).

- **Polisemia**

Decimos que dos palabras son polisémicas cuando el mismo signo lingüístico, palabra o sonido tiene más de un significado. Habitualmente el contexto de la conversación o lectura donde está insertada la palabra deshace los problemas de ambigüedad, pero una palabra polisémica introducida en un sistema documental, sin el contexto, puede dar lugar a ruido documental.

Ejemplo. Un usuario puede estar buscando sobre columnas en arquitectura y recuperar datos sobre columnas tipográficas de diarios. Los lenguajes documentales controlan la polisemia diferenciando cada significado con paréntesis, usando el plural o el singular, adjetivando, etc. Un tipo de polisemia es la homonimia. La diferencia entre ellas radica en la etimología de la palabra. Si la etimología de las dos palabras es la misma, hablamos de polisemia; si la etimología es diferente, hablamos de homonimia, esto es lo que aportan (Juncà Campdepadrós & Martínez Ferreras, 2009, p. 23).

2.3.2. Lenguaje controlado

Se entiende por lenguaje controlado a aquellos términos normalizados que proporcionan al usuario facilidad al momento de buscar información, aquí podemos

encontrar las listas de encabezamientos de materia, tesauros y las clasificaciones, también se puede decir que es un vocabulario previamente elaborado que representa el contenido específico de la materia que se está tratando y de esta manera localizar la información que se requiere.

De acuerdo a lo antes mencionado también se dice que se trata de un lenguaje controlado porque los términos que lo componen han sido depurados de las ambigüedades del lenguaje natural y están destinados a representar de manera unívoca el contenido de un documento y de las consultas dentro de un sistema documental. Es un lenguaje documental con voluntad de univocidad, que busca eliminar los fenómenos de sinonimia, polisemia y homonimia; por eso exige establecer el control sobre todos los términos que reúne (Juncà Campdepadrós y Martínez Ferreras, 2009).

Al lenguaje controlado Gil Urdiciain (1996) le conoce como lenguaje documental, ella menciona con referente a este tema lo siguiente, “lenguaje documental como todo sistema artificial de signos normalizados, que facilitan la representación formalizada del contenido de los documentos para permitir la recuperación, manual o automática, de información solicitada por los usuarios” mismo que coincide con Martínez y Murillo (2009) “el lenguaje documental resuelve igualmente problemas planteados por el multilingüismo, como en el caso de los tesauros multilingües, que sirven de puente a diferentes lenguas, o en el caso de los lenguajes de clasificación, fundados en la simbolización numérica, que los hace independientes de toda lengua y constituyen una escritura comprensible en todos los países; de ahí su fácil empleo internacional” (p. 72).

Para Slype (1991) existen dos tipos principales de lenguajes combinatorios controlados: las listas de autoridades y los tesauros de descriptores.

Lista de autoridades una lista de autoridades está constituida por una colección no ordenada (sino puesta por orden alfabético) de conceptos destinados a representar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental dado; estos conceptos son

expresados por palabras o por expresiones extraídas de una lista finita, establecida a priori; sólo los términos que figuran en esta lista pueden ser utilizados para indizar los documentos y las consultas. Ejemplos: biblioteca; servicio de documentación. La lista de autoridades es, la mayoría de las veces, monolingüe.

Thesaurus de descriptores un thesaurus es una lista estructurada de conceptos, destinados a representar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental determinado, y a ayudar al usuario en la indización de los documentos y de las consultas los conceptos son extraídos de una lista finita, establecida a priori; sólo los términos que figuran en esta lista pueden ser utilizados para indizar los documentos y las consultas; la ayuda al usuario la proporciona la estructura semántica del thesaurus: fundamentalmente las relaciones de equivalencia, de jerarquía y de asociación.

Ejemplos:

-biblioteca

genérico: sistema de información

específico: biblioteca pública:

: biblioteca escolar

-servicio de documentación

equivalente: sistema documental

genérico: sistema de información

específico: agencia de resúmenes

asociado: base de datos bibliográficos.

2.3.3. Tesoros

Los tesauros son lenguajes documentales, que con frecuencia se utiliza en los Centros de Información para realizar la indización, existen tesauros generales, como también especializados, etimológicamente, el término proviene del latín thesaurus, cuyo significado es tesoro y que poco tiene que ver con el concepto que tiene hoy para el profesional de la información.

La norma ISO 2788-1986 define un tesoro como: “Un vocabulario controlado y dinámico, compuesto por términos que tienen entre ellos relaciones semánticas y genéricas y que se aplica a un dominio particular del conocimiento”.

La ISO 2788/TC46 define la palabra tesoro desde el punto de vista funcional y estructural. Funcionalmente es un instrumento para controlar la terminología al trasladar a un lenguaje más estricto la lengua natural de los documentos. Estructuralmente es vocabulario controlado y dinámico de términos compuesto por descriptores relacionados entre sí semánticamente (jerarquía, asociación, equivalencia) y que se aplica a un dominio particular del conocimiento.

Por su parte, Van Slype (1991) define un tesoro como: “Una lista estructurada de conceptos destinados a representar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental determinado y a ayudar al usuario en la indización de los documentos y de las consultas” (p. 7).

Para Gil Urdiciain (1996) como cita Martínez y Murillo (2009), los tesauros son: “un lenguaje documental postcoordinado, de estructura combinatoria. Un vocabulario controlado dotado de relaciones que son las que le distinguen de otros instrumentos terminológicos y le proporcionan la categoría de lenguaje.

El Tesoro es un lenguaje documental que recoge palabras simples o compuestas que integran las expresiones utilizadas por los autores de los documentos de un área especializada que acreditan la garantía literaria de los términos; es un lenguaje de indización que utilizan los especialistas de la información para representar los temas

de los documentos y tiene la capacidad para indizar los documentos de una biblioteca o de una base de datos especializada o para marcar términos de recuperación en una biblioteca digital, también es un lenguaje controlado que los términos que incluye están validados por un proceso en el que se establecen relaciones entre ellos para mantener la univocidad de sus descriptores y se determina la preferencia de un término sobre otros posibles y de esta manera facilitar la recuperación de la información este es uno aporte de (Sánchez Luna, 2012).

Estaríamos hablando del tesoro como un vocabulario controlado de descriptores con un significado previamente establecido y cuyo fin último sería el de definir de manera unívoca el contenido de un documento, el fin último de este lenguaje es el de ser puente de conexión entre los usuarios y las unidades de información, esto es, de servir tanto para la indización como para la recuperación documental así lo manifiesta (Martínez Ferreras, 2017).

A continuación, Martínez Ferreras (2017) aporta lo siguiente:

Los tesauros pueden ser de los siguientes tipos:

- **Generales o especializados**, según su campo de aplicación.
- **Multidisciplinares o monodisciplinares**, según el número de disciplinas o temas que incluyen.
- **Macrotesauros o microtesauros**, según la naturaleza y compatibilidad con otros tesauros. Los macrotesauros, aunque poseen las características de un tesoro en cuanto a estructura y función, y al mismo tiempo, sus diferentes campos semánticos o microdisciplinas, pueden funcionar como si fueran tesauros independientes, con total autonomía. En este sentido, se podría hablar de un conjunto de tesauros más pequeños o microtesauros.
- **Monolingües o multilingües**, según las lenguas utilizadas.
- **Alfabéticos o sistemáticos**, según la presentación utilizada para mostrar los términos y sus relaciones semánticas.
- **Sistemas automatizados de gestión de tesauros**

En la actualidad con el avance de la tecnología todo va innovando por lo que en el área de la bibliotecología cada día se hace necesario la utilización de los sistemas automatizados de gestión de tesauros, estos facilitan la gestión y edición de un tesoro de manera automatizada, permitiendo de esta forma reducir el tiempo y los costes a la hora de su elaboración, su utilización permite:

- a) Incluir términos y registrar automáticamente las relaciones invertidas con otros términos del tesoro.
- b) Editar diferentes presentaciones del tesoro y seleccionar el tipo de relación que queremos visualizar en el listado.
- c) Utilizarse directamente en la consulta.
- d) Contar con prestaciones para los administradores del tesoro como las siguientes:
 - Posibilidad de mantener un fichero histórico donde se reflejen las modificaciones y los cambios que se hacen en los diferentes descriptores.
 - Facilidades a la hora de eliminar un descriptor del
 - Elaboración automática de estadísticas.

A continuación, cito dos ejemplos de sistemas de gestión de tesauros y se puede utilizar de acuerdo al requerimiento del centro de información de acuerdo al idioma en el que se presenta:

- **Multites** (en inglés): permite la conversión y generación de ficheros en lenguaje HTML y por ende la inclusión del tesoro en la red Internet.
- **TemaTres** (en español): es una aplicación web para la gestión de lenguajes documentales. Se encuentra orientada especialmente al desarrollo de tesauros jerárquicos, pero también puede utilizarse para desarrollar estructuras de navegación web, o como complemento de un gestor de contenidos en bibliotecas digitales o en una biblioteca tradicional.

2.4. Recuperación de la información

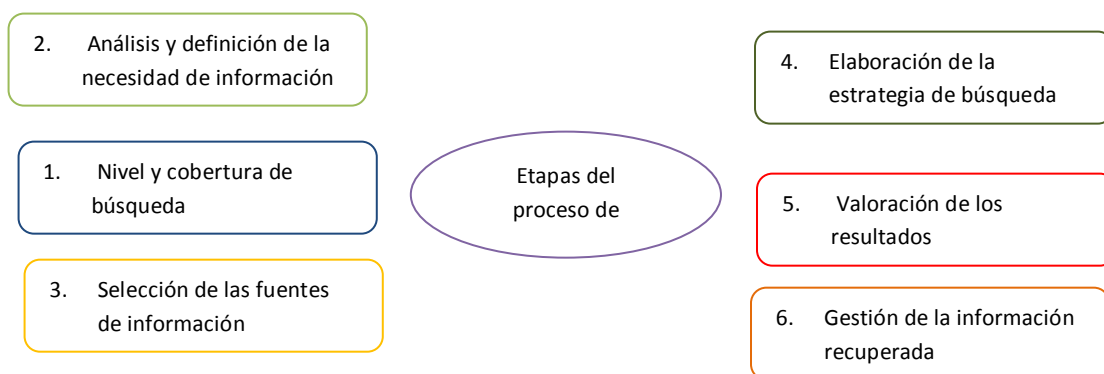
La recuperación de la información es un proceso global donde intervienen factores interrelacionados entre sí, la finalidad de la recuperación de la información o el tema, determina el tipo de fuente de información que se requiere consultar y los recursos que nos permite acceder, en la actualidad el desarrollo de las nuevas tecnologías y la facilidad de publicar en línea y la información disponible en la web, hacen necesario saber conocer las fuentes fiables y de calidad, como los catálogos y bases de datos, para ellos las búsquedas deben ser precisas y eficaces para obtener los resultados requeridos.

2.4.1. Etapas del proceso de búsqueda de información

La Biblioteca Universitaria de Alicante (2020) presenta en el siguiente gráfico a las etapas del proceso de búsqueda de información:

Figura 1

Proceso de búsqueda de información



“La búsqueda bibliográfica es el procedimiento que nos permite localizar, en una serie de fuentes de información distintas, un conjunto de documentos y de información necesarios para resolver cualquier problema de investigación o información que se nos plantee”, (Biblioteca Universitaria de Alicante, 2020, p. 3).

- **Analizar la situación o el problema**

En primer lugar, se debe tener clara la finalidad y los objetivos de la búsqueda, los antecedentes investigativos, la fundamentación teórica y la actualización son muy

importantes en una investigación, luego encuadrar el tema de búsqueda en un ámbito científico concreto, para evitar confusiones y recuperar información no necesaria. Seguidamente se tiene que concretar el punto de vista desde el que se enfocará el tema de búsqueda, se debe tener bien claro desde que punto de vista se va a realizar la investigación por ejemplo esta puede ser que se investigue un tema dentro del área social.

- **Establecer el nivel y la cobertura de la búsqueda**

Lo primero que se debe tener en cuenta es que se debe establecer el periodo temporal que debe cubrir la búsqueda, se debe delimitar la cobertura geográfica y saber seleccionar la tipología documental que desee manejar estos pueden ser revistas, monografías, libros entre otros, también podemos incluso delimitar a en qué tipo de formato se quiere recuperar la información como PDF, WORD, ECXEL, etc.

- **Seleccionar las fuentes de información que vamos a utilizar**

Es una de las fases más importantes del proceso de búsqueda. Nos garantiza la recuperación de la información más adecuada a nuestra necesidad de información, estas fuentes varían en función de la disciplina o ámbito de estudio, podemos recuperar la información de fuentes multidisciplinarias o también de fuentes especializadas, de acuerdo al requerimiento del investigador, algunos ejemplos de fuentes multidiciplinares se puede citar por ejemplo, las bases de datos Web of Science y Scopus, base de datos multidisciplinaria, referencial y de citas bibliográficas revisada por pares con mayor cobertura a nivel mundial y una base de datos especializada como Vlex, plataforma especializada en el suministro de información en todas las áreas jurídicas, ofrece una amplia colección de libros y revistas a texto completo, con la posibilidad de descarga.

- **Elaborar la estrategia de búsqueda**

Es la manera en que interrogamos a un recurso o un conjunto de recursos de información para localizar y acceder a la información que necesitamos, según la Biblioteca Universitaria de Alicante (2020), para elaborar una estrategia de búsqueda se debe seguir los siguientes pasos:

1. Escribe una frase que describa lo que quieres buscar.
2. Identifica los conceptos o ideas principales y elige los términos que mejor los representen.
3. Elimina conceptos de contenido vago o impreciso.
4. Busca términos alternativos: sinónimos u otros términos relacionados de contenido más amplio o específico.
5. Puedes utilizar diccionarios y tesauros.
6. No te olvides de los términos en otros idiomas, las variantes gramaticales o las palabras derivadas.
7. Si existe la posibilidad, traduce esos términos al lenguaje documental propio del recurso que se está consultando (tesauro, lista de materias)
8. Establece las relaciones entre los términos seleccionados con el lenguaje de interrogación propio del recurso (campos de búsqueda y operadores booleanos, de proximidad o de comparación).
9. Ejecuta la búsqueda.
10. Evalúa los resultados de tu búsqueda y revisa el proceso en función de los resultados.

- **Ejecutar la búsqueda y valorar los resultados**

Debes revisar los resultados obtenidos y comprobar si son pertinentes, es decir, si se adecuan a tus necesidades de información, ese proceso de revisión te permitirá detectar si has recuperado muchos documentos que en realidad no te interesan (ruido documental) o, si, por el contrario, te falta información relevante que deberías haber obtenido con la búsqueda (silencio documental).

- **Recuperar y almacenar la información localizada**

Se debe revisar los resultados obtenidos y comprobar si son pertinentes, es decir, si se adecuan a tus necesidades de información, los profesionales de la información de la Biblioteca de Alicante consideran que se debe tener en cuenta las siguientes opciones:

- Descargarte un fichero con las referencias bibliográficas localizadas o los textos completos de los documentos que te interesen si están disponibles.
- Enviarte por correo electrónico el listado con las referencias bibliográficas localizadas y el enlace al texto completo.
- Proquest y otras bases de datos permiten enviar, además, el texto completo de los documentos seleccionados si está disponible.
- Crearte un perfil personalizado de usuario en el recurso utilizado donde guardar tus búsquedas o tus documentos.
- La Biblioteca siempre dispone del gestor de referencias bibliográficas Refworks, que es una herramienta que permite crear una base de datos personalizada, alimentada automáticamente con todas las referencias localizadas en los distintos recursos. Además de poder almacenar y organizar las referencias y los archivos adjuntos, si están disponibles, te permitirán generar bibliografías e insertar citas de manera automática en tus trabajos y publicaciones. Además, existen otras herramientas similares de libre acceso, como Zotero o Mendeley.

2.4.2. Técnicas de búsqueda

En el siguiente cuadro se plasma las técnicas que se usa para proceder con la búsqueda de la información:

Figura 2
Técnicas de búsquedas



- **Definición del tema y selección de las palabras clave**

El éxito de la búsqueda estará determinado en gran parte por la adecuada selección de los términos que se utilice en la estrategia de búsqueda, utiliza términos significativos, que representen conceptos relevantes y que diferencien los documentos que le puedan interesar del resto de documentos que pueda contener el recurso consultado.

Olvidarse de preposiciones, conjunciones, artículos, pronombres, adjetivos o adverbios.

Cuando buscamos documentos por las palabras clave que pueden contener en el título, el resumen o su texto completo, estamos utilizando un lenguaje natural, y debemos tener en cuenta las distintas formas en que puede representarse un concepto. Además, se tiene que tener en cuenta las variantes de género y número

(masculino / femenino; singular / plural), aunque hay recursos, como bases de datos y portales de revistas electrónicas, que automáticamente recuperan tanto el singular como el plural del término seleccionado para la búsqueda.

Este modo de realizar la búsqueda puede llevar a perder información relevante o a recuperar información que no es de interés, en ocasiones, es más apropiado buscar la información por términos que reflejen el contenido de los documentos, por lo general, esos términos forman parte del lenguaje documental y están disponibles en el recurso a través de la lista de materias o de los tesauros, para ello se ha de utilizar el campo materia, descriptor o palabras clave si están disponibles en el recurso (Biblioteca Universitaria de Alicante, 2020, p. 12).

- **Campos de búsqueda y modos de búsqueda**

Los campos de búsqueda más destacados son: título, autor, palabras clave, materia o descriptor, resumen y todos los campos, este último campo es el que permite buscar información en cualquier parte del registro y engloba a todos los demás, sólo se debe usar cuando no se encuentre información por otros campos, ya que es demasiado genérico. En recursos que tienen información a texto completo, se dispone también de la opción campo texto completo.




El uso del campo materia, palabra clave o descriptor permite realizar búsquedas más exhaustivas, el número de campos de búsqueda disponibles varía en función del tipo de recurso, en general, las bases de datos o los catálogos disponen de más campos que los portales de revistas electrónicas, hay algunos portales de revistas electrónicas que tienen opciones de búsqueda similares a las de las bases de datos, pero es frecuente encontrar casos en los que los únicos campos son, al margen del autor: título, título y resumen o título, resumen y texto completo. (Biblioteca Universitaria de Alicante, 2020, p. 20).

- **Operadores**

Son los elementos que expresan las relaciones entre los términos incluidos en la estrategia de búsqueda, hay 4 tipos principales de operadores.

- **Lógicos o booleanos:** permiten combinar tanto palabras dentro de un mismo campo (por ejemplo, el título), como varios campos entre sí (por ejemplo, título y autor). Reflejan las relaciones entre conceptos que identifican documentos. Suelen escribirse en inglés y los más importantes son: OR (O), AND (Y) y NOT (NO).

Figura 3
Operadores booleanos

OPERADOR	DESCRIPCIÓN	EJEMPLO
AND (Intersección)	Reduce y especifica la búsqueda.	Contabilidad AND Auditoría 
OR (Union)	Amplia la búsqueda.	Contabilidad OR Auditoría 
NOT (Exclusión)	El término o expresión que le sigue.	Contabilidad NOT Auditoría. 

Fuente: blogs.unc.edu.ar

- **Posicionales o de proximidad:** permiten buscar un término según en el contexto en que se encuentra, es decir, en función de su posición respecto a otros términos en el campo que seleccionemos. Los más frecuentes son NEAR (cerca) y ADJ (adyacencia). Suelen utilizarse para campos de búsqueda que contienen muchas palabras, como el resumen o el texto completo. Si van seguidos de un número, indican la cantidad máxima de palabras que puede haber entre los términos de búsqueda.

- **Comparación o rango:** sirven para limitar la búsqueda mediante una expresión numérica, datos o, generalmente, fechas: menor que (2000), igual que (=2000), o en un periodo concreto (2000-2013).
- **Truncamientos:** Permiten buscar por la raíz de las palabras, pueden colocarse a la izquierda, a la derecha o en medio del término, los truncamientos por la derecha son los más habituales, se utilizan para buscar plurales y variantes de términos, se representa con un signo: *, \$,?. (Biblioteca Universitaria de Alicante, 2020)

2.4.3. Los metadatos

Para definir lo que es metadatos se revisa información en varios sitios ya sean estos artículos, libros, blogs, entre otros, a continuación, adjunto algunos de ellos, los metadatos en la navegación y recuperación de información se utilizan para detectar información relevante de una forma rápida y eficaz. Las etiquetas describen el contenido del recurso web, que posteriormente utilizan las herramientas de búsqueda para localizar y acceder al recurso. Principalmente son las etiquetas de palabra clave y título las que dan paso a localizar el documento.

“Es común describir el término metadatos como datos que describen otros datos o "datos sobre datos". De forma general, en efecto, el concepto de metadatos se refiere a aquellos datos que hablan de los datos, es decir, describen el contenido de los archivos o la información de los mismos” (PowerData, 2016)

“Precisamente, los metadatos son “datos que hablan acerca de los datos”, en el sentido de que describen el contenido de los archivos o la información que estos traen en su interior”. (Docunecta, 2020).

A continuación, la biblioteca de la CEPAL define a los metadatos de la siguiente manera:

Cuando hablamos de documentar los datos, es frecuente que se mencione a los metadatos. Esto ocurre ya que tanto la documentación como los metadatos abarcan al conjunto de información que describe los datos generados en la investigación. La principal diferencia entre la documentación en general y los metadatos, radica en que estos últimos poseen una estructura que les permite ser procesados mediante computadoras, por lo cual pueden transmitirse con mayor facilidad y ser interoperables.

Los metadatos juegan un papel importante en el contexto de los datos FAIR, ya que son considerados incluso más importantes que los propios datos. Esto se debe a su disposición constante y abierta, la cual permite enlazar datos de investigación y publicaciones en la web sobre datos y servicios FAIR (CEPAL, 2020).

- **Tipos de metadatos**

Según la CEPAL (2020) existen tres tipos de metadatos que se cita a continuación.

- **Metadatos descriptivos**

Información sobre el contenido del recurso, que es de utilidad para su recuperación y comprensión.

Por ejemplo:

- Título
- Autor
- Abstract
- Palabras clave

- **Metadatos administrativos**

Información necesaria para manejar un recurso o que se relaciona con su creación. Dentro de esta categoría se encuentran:

- metadatos técnicos
- los metadatos de
- y los metadatos de derechos
- **Metadatos estructurales**

Describen las relaciones que existen entre las partes de un conjunto de datos. Por ejemplo, un esquema que representa las relaciones entre tablas de una base de datos o una tabla de contenidos.

- **Funciones de los metadatos**

Algunas de las funciones más importantes de los metadatos son las siguientes:

- **Búsqueda:** los metadatos deben proporcionar suficiente información
- **Recuperación:** los metadatos deben proporcionar información a los usuarios para que puedan adquirir la información que sea de su interés. En este sentido, también se considera la “función recuperación” a aquella información que describe cómo localizar fuera de línea los datos, la persona de contacto, los formatos de distribución de los datos o cualquier restricción de acceso a los datos.
- **Transferencia:** los metadatos deben facilitar la información necesaria para que los usuarios hagan uso de los archivos recuperados en sus máquinas. Este componente incluiría información sobre el tamaño del conjunto de datos, y sus metadatos, la estructura tanto lógica como física de los datos y metadatos.
- **Evaluación:** los metadatos deben considerar información que asista a los usuarios a determinar si los datos van a ser útiles para una aplicación.
- **Archivo y conservación:** los metadatos son una pieza clave para garantizar que los recursos de información se documenten, se definan sus responsables y continúen siendo accesibles en el futuro (NISO, 2004).
- **Interoperabilidad:** los metadatos facilitan la interoperabilidad, puesto que se han definido estándares de metadatos y existen protocolos compartidos para el intercambio de esta información. Protocolos como el Z39.50 o el CSW han ayudado en búsquedas simultáneas de datos en sistemas distribuidos.

CAPITULO III

MARCO METODOLÓGICO

3.1. Ubicación

La investigación se centró en el análisis y definición de los tipos de lenguajes documentales, los listados de descriptores libres, lenguaje natural y lenguaje controlado para la recuperación de la información, para lo cual se ha seleccionado los sistemas de gestión de información de la bibliotecas universitarias caso puntual del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, ya que está conformado por cuatro bibliotecas universitarias tres de ellas especializadas y una general, esta se encuentra ubicada en la cabecera cantonal del Cantón Cuenca, de la provincia del Azuay.

Figura 4
Ubicación Universidad de Cuenca



Fuente: google maps

El Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, está conformado por cuatro bibliotecas en cuatro campus diferentes, en el Campus Central se encuentra la coordinación general a cargo de la Licenciada Rocío Campoverde, también está conformado por un especialista de biblioteca, el Área de Normalización y Procesos Técnicos está conformada por tres bibliotecarios y un técnico de biblioteca quienes son los encargados de la cadena documental, encargados también de ingresar la información en el repositorio institucional y el Área de Servicios

conformada por ocho bibliotecarios y tres técnicos de biblioteca, la biblioteca presta servicio de estantería abierta para la comunidad universitaria en especial usuarios de las siguientes facultades que se encuentran en el campus central, Facultad de Psicología, Facultad de Filosofía, Facultad de Jurisprudencia y Facultad de Ingeniería, también se presta el servicio a las Facultades de Ciencias Químicas y Facultad de Ingeniería que se encuentra en el Campus Balzay pero el fondo bibliográfico aún se encuentra en el campus central ya que el nuevo campus no cuenta con Biblioteca.

La biblioteca del Campus Paraíso cuenta con un especialista de biblioteca y tres bibliotecarios mismos que se encargan de todo el procesos de gestión bibliotecaria, basados en las directrices que les proporcionan desde el área de procesos técnicos del Campus Central, prestan sus servicios a usuarios de las Facultades de Ciencias Médicas y Odontología, de la misma manera en el Campus Yanuncay la biblioteca cuenta con un especialista de Biblioteca y dos bibliotecarios, presta los servicios a usuarios de las Facultades de Ciencias Agropecuarias, Medicina Veterinaria y la Facultad de Artes, y por último, la Biblioteca del Campus Centro Histórico cuenta solo con un bibliotecario encargado de todos los procesos y brinda los servicios a usuarios de la Facultad de las Ciencias de la Hospitalidad.

3.2. Equipos y materiales

Para el desarrollo de la investigación se consideró los siguientes equipos y materiales los mismos que se detalla a continuación:

Recursos humanos

- Investigadora
- Directora del Trabajo de Titulación
- Personal bibliotecario del Centro de Documentación Regional Juan Bautista Vázquez

Recursos Tecnológicos

- Internet

- Computadora
- Celular
- Sistemas de Gestión Bibliotecaria.
- Google Drive

Recursos materiales

- Hojas de papel bond
- Esferos

3.3. Tipo de investigación

Para la realización del trabajo, se utilizó el método exploratorio, se analizó la tipología del lenguaje documental y también se analizó el estado actual en el que se organiza el contenido de la información en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca.

Concordando con lo que menciona Hernández Sampieri *et al.* (2014) “Los estudios exploratorios sirven para familiarizarnos con fenómenos relativamente desconocidos, obtener información sobre la posibilidad de llevar a cabo una investigación más completa respecto de un contexto particular, indagar nuevos problemas, identificar conceptos o variables promisorias, establecer prioridades para investigaciones futuras, o sugerir afirmaciones y postulados”, p. 70.

Otro de los métodos que se utilizó, es la revisión bibliográfica, para contextualizar y conceptualizar la investigación, se realizó una revisión bibliográfica exhaustiva en artículos científicos, libros digitales y tesis, que tratan de temas similares sobre lenguaje controlado, lenguaje natural y la recuperación de la información, mismo que da cumplimiento al primer objetivo específico.

También, para la recopilación de la información se usó un cuestionario para realizar una encuesta, misma que se compone de preguntas cerradas y se aplicó a los profesionales de la información del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca.

3.4. Idea a defender

La normalización del lenguaje natural mediante el lenguaje controlado permitirá optimizar los procesos de la recuperación de la información.

3.5. Población o muestra

La población que se estudió se refiere como lo menciona Triola (2009), es el conjunto completo de todos los elementos (puntuaciones, personas, medidas, etc.) que se va estudiar. El conjunto es completo porque incluye a todos los sujetos que se estudiarán. Por lo antes mencionado la población a la cual se aplicó la encuesta es a todo el personal del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, no fue necesario aplicar ningún tipo de muestra ya que el universo es muy reducido por lo tanto a continuación se detalla el número de profesionales:

Tabla 7
Población a ser encuestada

Detalle	Número	Porcentaje
Especialista de biblioteca	3	12.50%
Bibliotecario	17	70.83%
Técnico de biblioteca	4	16.67%
Total	24	100%

3.6. Recolección de información

Para la recolección de la información se aplicó una encuesta a los profesionales de la información que laboran en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, para lo cual se utilizó un cuestionario diseñado en Google Drive con preguntas cerradas para facilitar la tabulación, se aplicó de manera virtual enviando el enlace del cuestionario a los correos electrónicos de cada profesional.

3.6.1. Validación de los instrumentos.

El instrumento de recolección de información para la encuesta, fue remitido a dos profesionales siendo ellos el Magíster Luis Enrique Lescano Borrego, especialista de Biblioteca Universitaria del Centro de Documentación Regional Juan Bautista Vázquez y a la Magíster María del Carmen Gaibor Orbea, directora de la Biblioteca Universidad Central del Ecuador, para la respectiva validación previa a su aplicación.

3.7. Procesamiento de la información y análisis estadístico

Los resultados de la encuesta se recolectaron y tabularon mediante la aplicación Google forms, estos se presentan mediante un análisis estadístico descriptivo a través de gráficos y tablas de doble entrada con lo que se analizó los datos correspondientes para su análisis e interpretación y con los resultados definir las conclusiones y recomendaciones.

3.8. Variables respuesta o resultados alcanzados

Para la presente investigación previo a aplicar la encuesta se realizó la validación del instrumento, por dos expertos en el área de bibliotecología, para posterior aplicación de la encuesta al personal que labora en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca.

CAPITULO IV
RESULTADOS Y DISCUSIÓN

4.1. Análisis de los resultados

Los resultados obtenidos de las 24 encuestas aplicadas a los profesionales de la información que laboran en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, involucrados directamente con el uso de los lenguajes documentales, dan los siguientes resultados que serán presentados y analizados a continuación.

4.2. Interpretación de datos

Los profesionales que laboran directamente con el análisis de la información son 24, con respecto a la edad, sexo y nivel de instrucción el resultado es el siguiente:

Tabla 8
Sexo, edad e instrucción académica

	Alternativas	Cantidad	Porcentaje
Sexo	Hombres	7	29.2%
	Mujeres	17	70.8%
	Total	24	100%
Edad	20-30	6	25%
	31-40	9	37.5%
	41-50	6	25%
	+51	3	12.5%
	Total	24	100%
Instrucción	Bachiller	2	8.3%
	Egresado	1	4.2%
	Tercer nivel	9	37.5%
	Cuarto Nivel	12	50%
	Total	24	100%

Análisis e interpretación

En el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca entre los profesionales que trabajan con el análisis de la información son 24 personas que corresponden al 100%, de ellos 17 son mujeres que corresponde al 70.8% y 7 hombres que corresponde al 29.2%, en los rangos de edades se encuentra que 6 personas que corresponden al 25% se encuentran dentro del rango de 20 a 30 años de edad, 9 personas que corresponde al 37.5% se encuentran dentro del rango de 31 a 40 años de edad, 6 personas que corresponde al 25% se encuentran en el rango de edad de 41 a 50 años y 3 personas que corresponde el 12.5% están dentro del rango mayores de 51 años de edad. También se puede observar que existen 2 profesionales son bachilleres que corresponden al 8.3%, 1 egresado de tercer nivel que corresponde al 4.2%, 9 que corresponde al 37.5% tienen tercer nivel y 12 que corresponde al 50% cuenta con cuarto nivel, lo que se evidencia que existen profesionales de diversas edades y grados académicos.

También se verifica que 12 profesionales que corresponde al 50%, son graduados de la carrera de bibliotecología y el otro 50% son de otras áreas como Ciencias de la Educación, Turismo, Administración de Empresas, Ingeniería Comercial, Ingeniería de Empresas entre otras áreas. Esto por un lado puede ser positivo ya que pueden intercambiar conocimientos tanto desde su experiencia laboral como también los nuevos conocimientos adquiridos en las aulas de estudio y que están a la par de la tecnología.

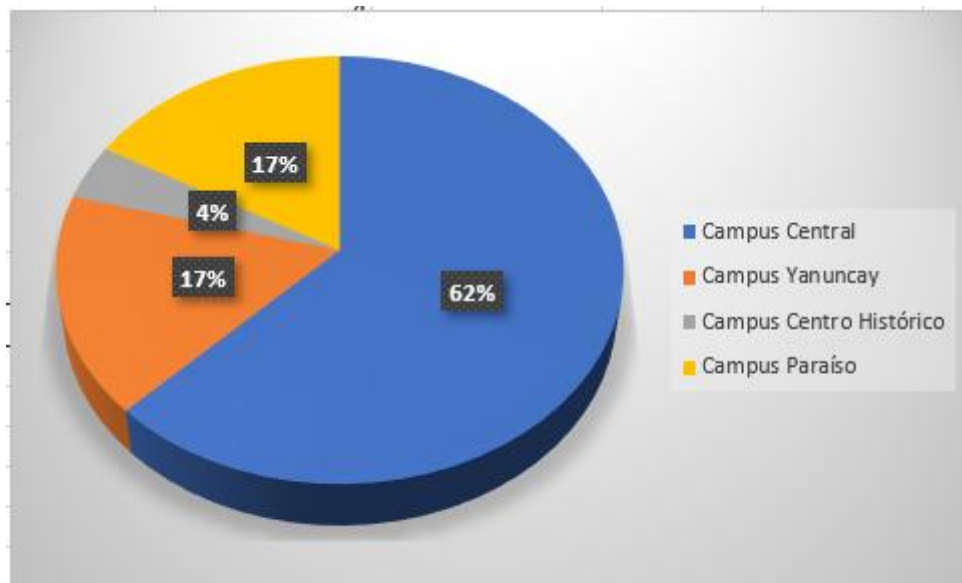
1. Biblioteca en la que labora

Tabla 9

Biblioteca en la que labora

Biblioteca	Cantidad	Porcentaje
Campus Central	15	62.50%
Campus Yanuncay	4	16.70%
Campus Centro Histórico	1	4.20%
Campus Paraíso	4	16.70%
Total	24	100%

Figura 5
Biblioteca en la que labora



Análisis e interpretación

Se puede apreciar en el gráfico que 15 profesionales que corresponde al 62%, se encuentra en la Biblioteca del Campus Central, esto es debido a que la biblioteca es más grande y que también el Área de Normalización y Procesos Técnicos, encargados del procesamiento de la información se encuentra en dicha biblioteca, mientras que en las bibliotecas de los otros Campus pocos profesionales trabajan con el tratamiento de la información.

2. Tiempo de experiencia en Bibliotecología o Archivología

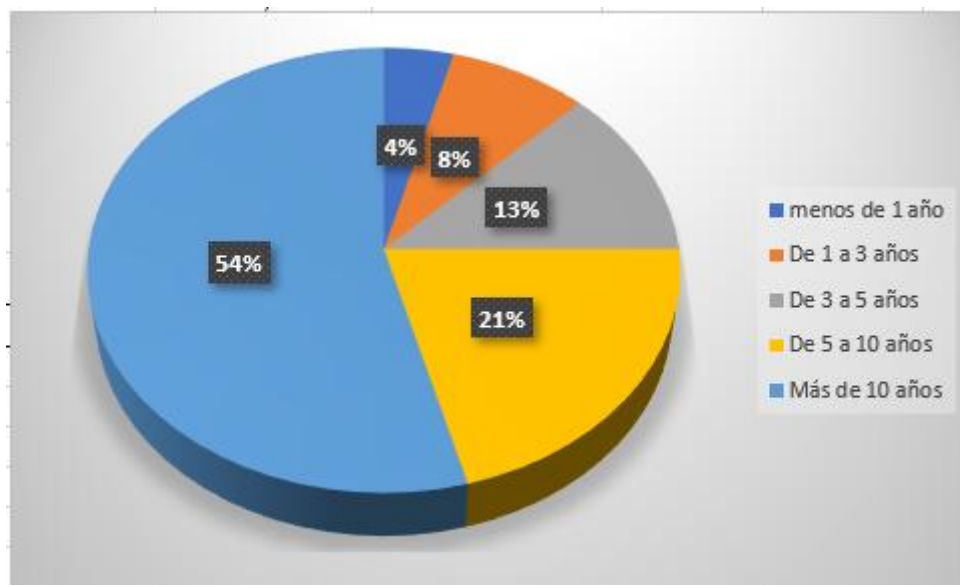
Tabla 10

Años de experiencia

Años	Cantidad	Porcentaje
menos de 1 año	1	4.17%
De 1 a 3 años	2	8.33%
De 3 a 5 años	3	12.50%

De 5 a 10 años	5	20.83%
Más de 10 años	13	54.17%
Total	24	100%

Figura 6
Años de experiencia



Análisis e interpretación

Los profesionales de la información que laboran en el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca de los 24 encuestados 13, que corresponde al 54% tienen una amplia experiencia en las ciencias de la información con más de 10 años, seguido por 5 personas que corresponde al 21% que tienen experiencia entre 5 a 10 años, por lo que se puede manifestar que el personal tiene una experiencia basta para manejar una biblioteca universitaria.

3. ¿Cuál es su denominación del puesto que desempeña actualmente?

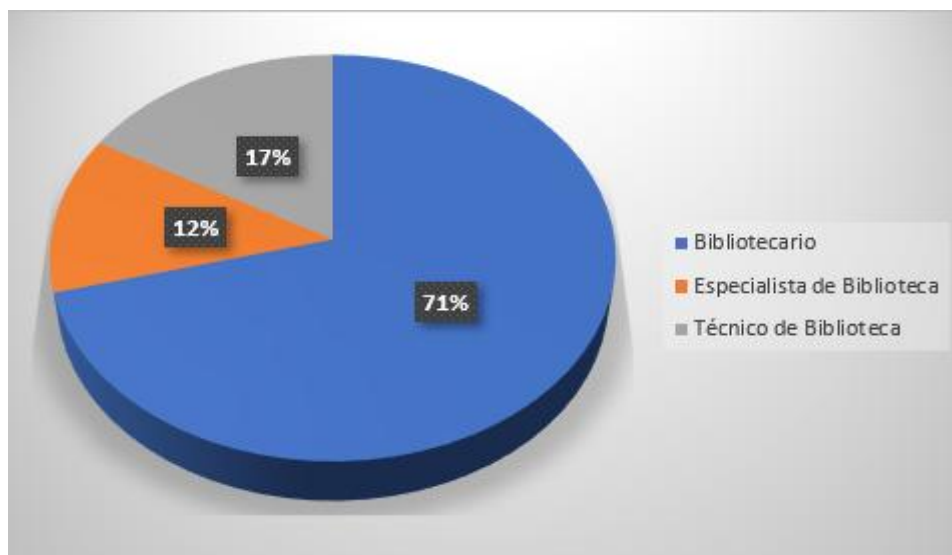
Tabla 11

Denominación del puesto

Denominación de puesto	Cantidad	Porcentaje
Especialista de Biblioteca	3	12.50%
Bibliotecario	17	70.83%
Técnico Bibliotecario	4	16.67%
Total	24	100%

Figura 7

Denominación del puesto



Análisis e interpretación

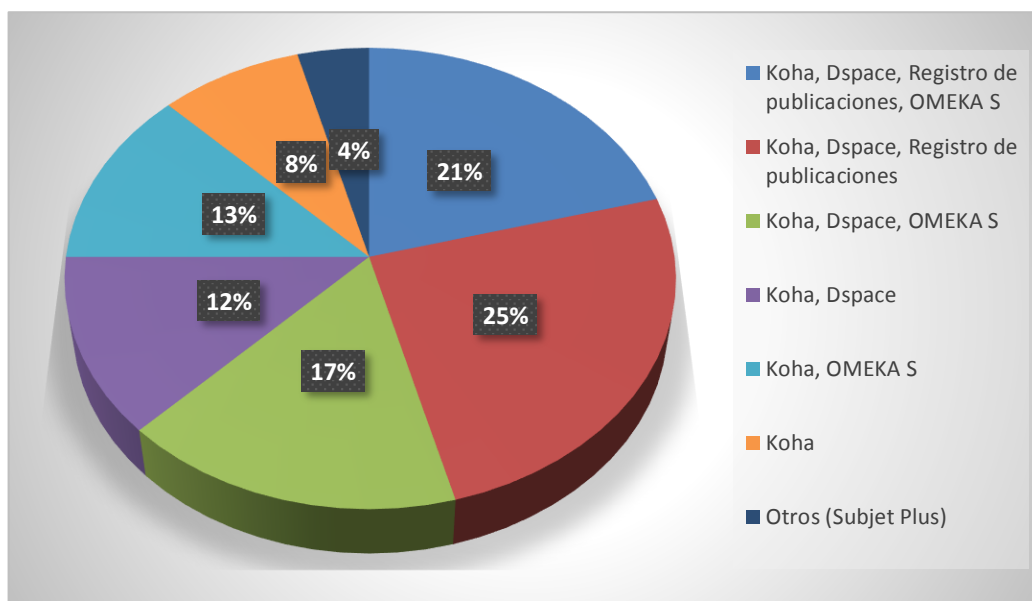
Los profesionales del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, de los 24 encuestados 17, que corresponde al 71% son Bibliotecarios, encargados en realizar las actividades bibliotecarias de manejo y tratamiento de la información.

3 ¿Qué plataforma de ingreso de información maneja usted?

Tabla 12
Plataforma de gestión bibliográfica

Plataforma	Cantidad	Porcentaje
Koha, Dspace, Registro publicaciones, OMEKA S	5	20.83%
Koha, Dspace, Registro de publicaciones	6	25.00%
Koha, Dspace, OMEKA S	4	16.67%
Koha, Dspace	3	12.50%
Koha, OMEKA S	3	12.50%
Koha	2	8.33%
Otros (Subjet Plus)	1	4.17%
Total	24	100%

Figura 8
Plataforma de gestión bibliotecaria



Análisis e interpretación

Seis bibliotecarios, que corresponde al 25% del personal encuestado trabajan con las plataformas, Koha, Dspace y Registro de Publicaciones; 5 bibliotecarios que corresponde al 21% manejan todas las plataformas, otros manejan 1 ó 2 plataformas diferentes. Pero lo que se puede visualizar es que el 100% de encuestados trabajan con el Koha, debido a que es el Sistema de Gestión Bibliotecaria del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca. Los bibliotecarios, especialistas de biblioteca y técnicos de biblioteca tienen permisos para catalogar o también para depurar los registros bibliográficos en el S.G.B Koha, a diferencia que en las otras plataformas acceden solo determinados bibliotecarios y especialistas de biblioteca que se les ha otorgado los correspondientes permisos.

4. Tipos de registros bibliográficos que usted procesa

Tabla 13

Tipos de registros bibliográficos

Tipo de documento	Cantidad	Porcentaje
Libros, Revistas, Trabajos de titulación, Artículos científicos	5	20.83%
Libros, Revistas, Trabajos de titulación	4	16.67%
Libros, Trabajos de titulación	1	4.17%
Libros, Trabajos de titulación, Artículos científicos	3	12.50%
Libros, Trabajos de titulación, Material audiovisual	1	4.17%
Libros	4	16.67%
Todos los anteriores	6	25.00%
Total	24	100%

Figura 9

Tipos de registros bibliográficos



Análisis e interpretación

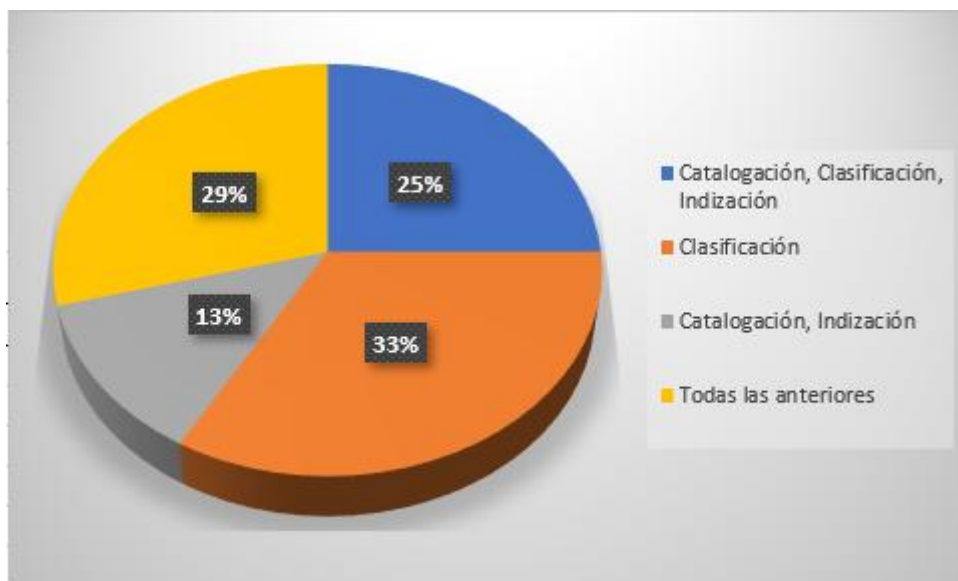
En concordancia con la pregunta anterior se puede comprobar que los resultados son similares por lo que el 100% de los encuestados manifiestan que trabajan con libros ya que la catalogación y la depuración de los registros bibliográficos se lo hace en el S.G.B. Koha, por lo tanto 6 personas que corresponde al 25% responden que trabajan con Libros, Revistas, Trabajos de Titulación y Artículos científicos debido a que a ellos se les otorga los permisos para acceder a plataformas como Koha, Dspace, Registro de Publicaciones y Omeka S.

3. ¿Qué actividades usted realiza en el procesamiento de la información?

Tabla 14
Actividades de Indización

Actividades	Cantidad	Porcentaje
Catalogación, Clasificación, Indización	6	25.00%
Clasificación	8	33.33%
Catalogación, Indización	3	12.50%
Todas las anteriores	7	29.17%
Total	24	100%

Figura 10
Actividades de Indización



Análisis e interpretación

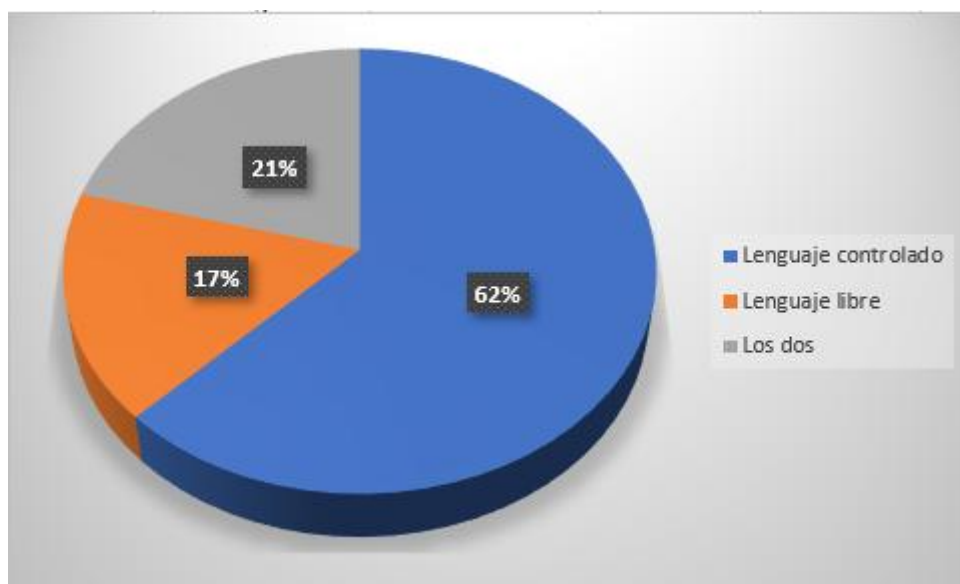
7 profesionales que corresponde al 33%, realizan todas las actividades de tratamiento documental como Catalogación, Clasificación, Indización incluyendo resumen, 6 encuestados que corresponde al 25% hacen actividades similares excluyendo el resumen, el resto de encuestados solo clasifican o solo indizan.

7. ¿Qué tipo de lenguaje documental utiliza?

Tabla 15
Uso del lenguaje documental

Lenguajes documentales	Cantidad	Porcentaje
Lenguaje controlado	15	62.50%
Lenguaje libre	4	16.67%
Los dos	5	20.83%
Total	24	100%

Figura 11
Uso del lenguaje documental



Análisis e interpretación

Para dar cumplimiento al objetivo general de este trabajo de investigación mismo que dice analizar los diferentes tipos de lenguajes documentales, para la recuperación de la información en las bibliotecas universitarias, para lo cual se ha visto la necesidad de preguntar a los profesionales que tipo de lenguaje documental utilizan, obteniendo los siguientes resultados los encuestados responden al 62% que se debe utilizar el lenguaje controlado, mientras que en menor porcentaje responden que utilizan el lenguaje natural o los dos lenguajes tanto controlado como el natural.

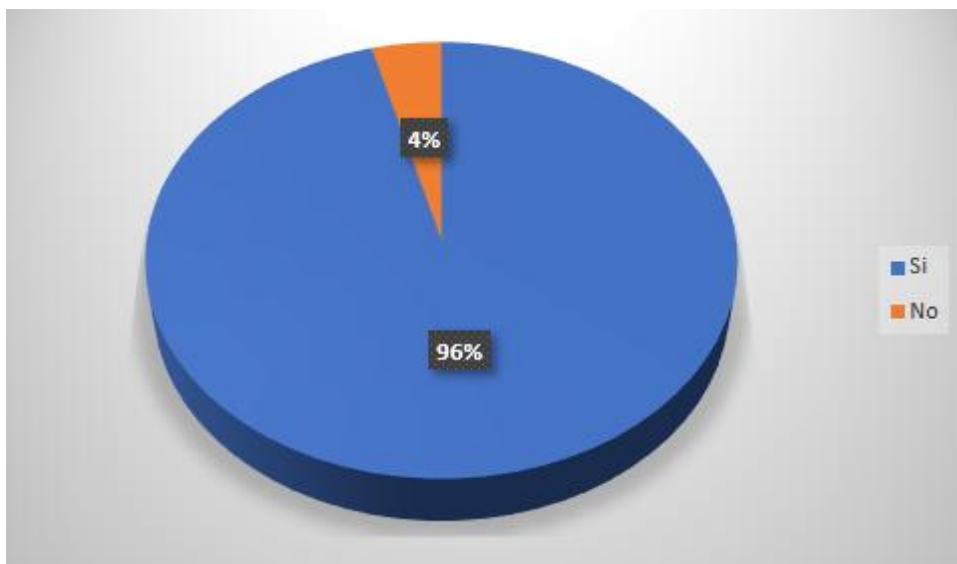
8. ¿Es importante el proceso de la indización en el tratamiento documental?

Tabla 16
Importancia de la indización

Alternativas	Cantidad	Porcentaje
Si	23	95.83%
No	1	4.17%
Total	24	100%

8. ¿Es importante el proceso de la indización en el tratamiento documental?

Figura 12
Importancia de la indización



Análisis e interpretación

La Indización es fundamental en el tratamiento documental puesto que su finalidad es identificar, clasificar y localizar los documentos, para de esta manera poder recuperar la información, utilizando los conceptos, las materias o los descriptores en los tradicionales catálogos manuales u también hoy en día en el Opac, en los Repositorios institucionales o Bases de datos. En concordancia con lo mencionado 23 encuestados que corresponde al 96% responde que sí es muy importante la Indización, frente al 4.17% responde negativo.

9. ¿Qué lenguaje documental usted utiliza para la indización?

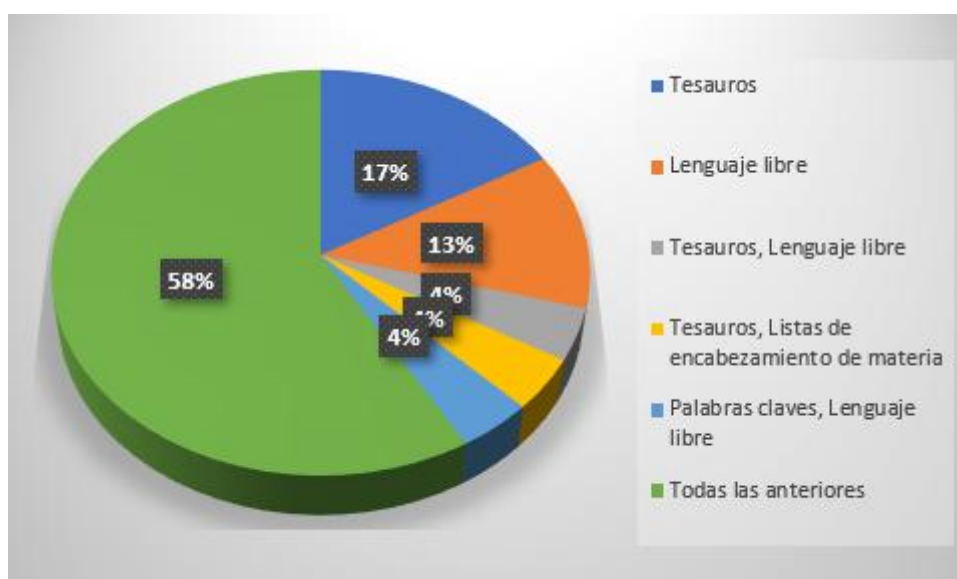
Tabla 17
Instrumentos de indización

Lenguaje documental	Cantidad	Porcentaje
---------------------	----------	------------

Tesauros	19	79.20%
Listas de encabezamiento de materia	9	37.50%
Listas de descriptores	7	29.20%
Clasificaciones (CDU, CDD)	10	41.70%
Palabras claves	11	45.80%
Lenguaje libre	12	12.00%
Total	24	100%

9. ¿Qué lenguaje documental usted utiliza para la indización?

Figura 13
Instrumentos de indización



Análisis e interpretación

Entre los lenguajes documentales tanto de vocabulario controlado o libre tenemos los Tesauros, Listas de encabezamientos de materia, Listados de Palabras Claves, Sistemas de Clasificaciones, el lenguaje libre, que nos ayudan en la indización y por ende a la recuperación de la información, el 58% de los encuestados utilizan todas las herramientas mencionadas por lo que esta actividad está siendo bien encaminada en el procesamiento de la información, mientras que el 42% utilizan una o dos herramientas.

10. ¿Cuáles son las dificultades que usted atraviesa al momento de la indización?

Tabla 18

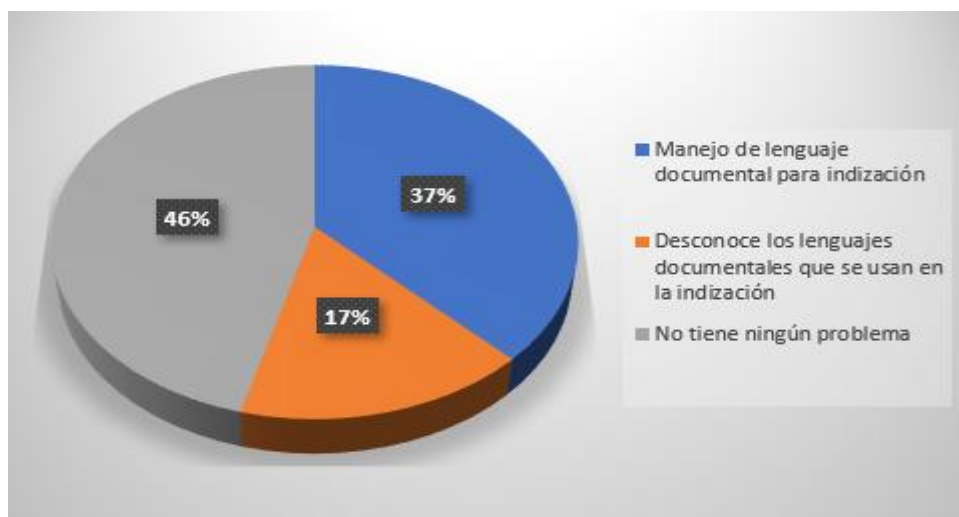
Dificultades de indización

Alternativas	Cantidad	Porcentaje
Manejo de lenguaje documental para indización	9	37.50%
Desconoce los lenguajes documentales que se usan en la indización	4	16.67%
No tiene ningún problema	11	45.83%
Total	24	100%

10. ¿Cuáles son las dificultades que usted atraviesa al momento de la indización?

Figura 14

Dificultades de indización



Análisis e interpretación

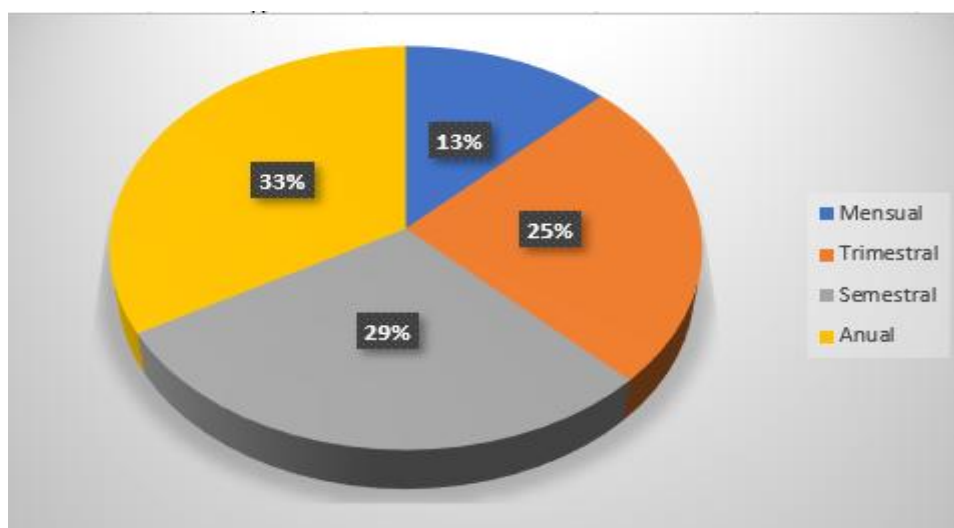
De los 24 encuestados 11 personas que corresponde al 45%, responde que no tiene ningún problema al momento de realizar la Indización, frente al 37% tiene algún problema con el manejo del lenguaje documental para la indización debido a que cuentan con profesiones diferentes a la bibliotecología y solamente maestrías en el área de bibliotecología o Archivología, por los que sus bases no están bien fundamentadas y existe algún tipo de desconocimiento.

11. ¿Con qué frecuencia cree usted que se debe realizar capacitaciones sobre el manejo del lenguaje documental?

Tabla 19
Frecuencia de capacitaciones

Periodo	Cantidad	Porcentaje
Mensual	3	12.50%
Trimestral	6	25.00%
Semestral	7	29.17%
Anual	8	33.33%
Total	24	100%

Figura 15
Frecuencia de capacitación



Análisis e interpretación

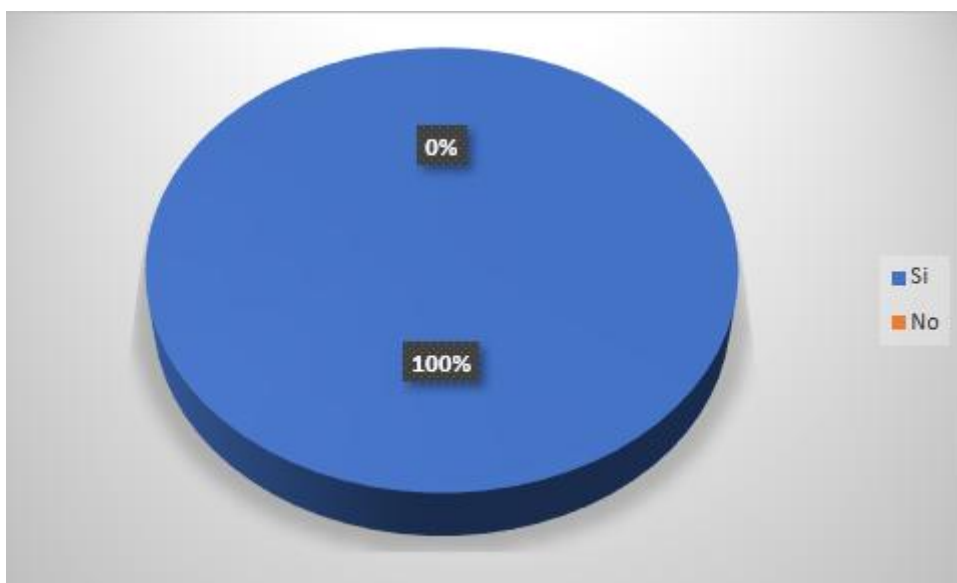
De los encuestados 8 profesionales que corresponde al 29% y 6 que corresponde al 6 responden que se debe realizar capacitaciones anuales o semestrales sobre el manejo del lenguaje documental puesto que no tienen problema con el manejo del lenguaje documental, y sería por conocer sobre algunas actualizaciones o por ayudar a los demás profesionales que, si tienen problemas en el manejo o algún tipo de desconocimiento, ellos son los que si piden la capacitación sea más frecuente.

12. ¿Usted considera que el uso de lenguaje controlado ayuda en la interoperabilidad?

Tabla 20
Interoperabilidad

Alternativas	Cantidad	Porcentaje
Si	24	100.00%
No	0	0.00%
Total	24	100%

Figura 16
Interoperabilidad



Análisis e interpretación

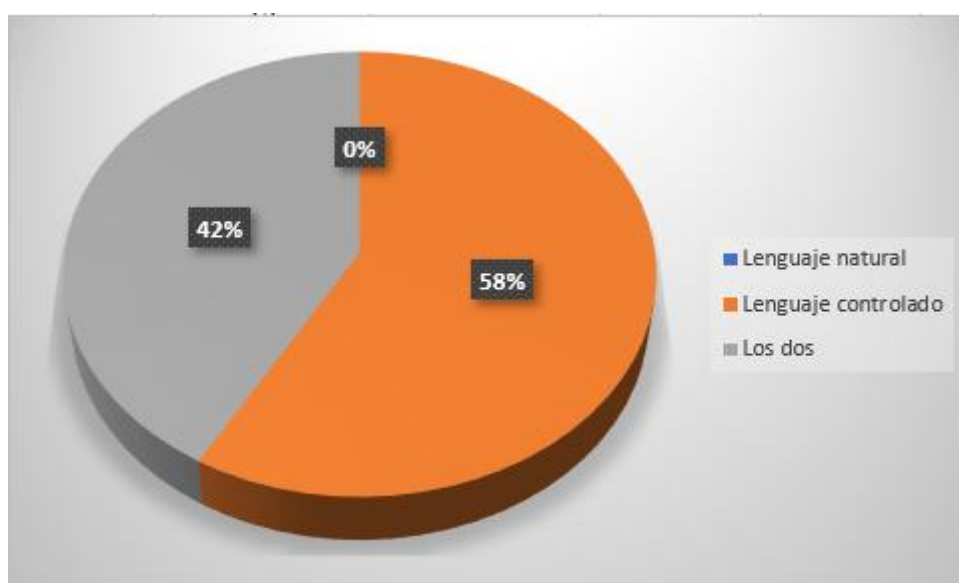
En el centro de Documentación Regional Juan Bautista Vázquez de la Universidad el 100% de encuestados están conscientes que de la correcta indización depende la interoperabilidad entre otros sistemas de recuperación de la información, puesto que así los documentos estén en acceso libre no podrán ser recuperados si está mal indizado.

13. ¿Cuál lenguaje documental sería el adecuado utilizar en el registro bibliográfico?

Tabla 21
Lenguaje documental recomendado

Tipo de lenguaje	Cantidad	Porcentaje
Lenguaje natural	0	0.00%
Lenguaje controlado	14	58.33%
Los dos	10	41.67%
Total	24	100%

Figura 17
Lenguaje documental recomendado



Análisis e interpretación

El 58% de encuestados responde que se debe utilizar siempre el lenguaje controlado en los registros bibliográficos, para indizar el material bibliográfico, mientras que el 42% manifiesta que se debe utilizar el lenguaje controlado como también el lenguaje libre puesto que algunos investigadores buscan información por palabras claves, partes de título o del contenido de los documentos.

14. ¿Cree que sería útil unificar el lenguaje natural con el controlado?

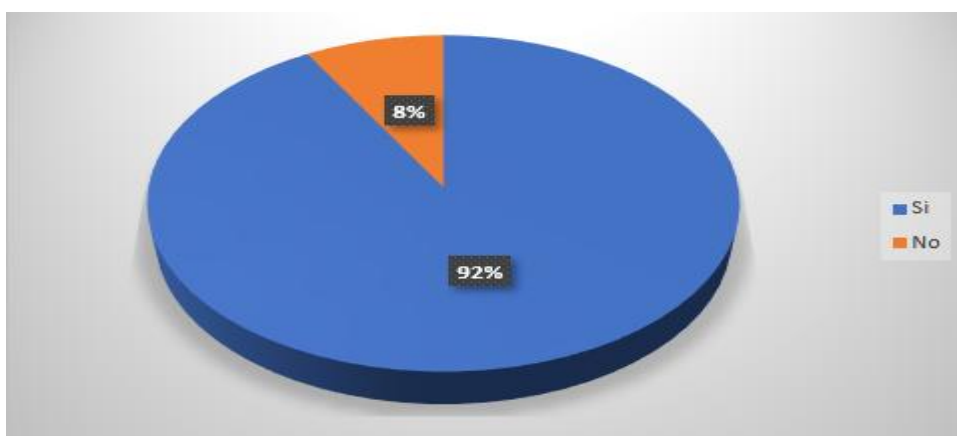
Tabla 22

Unificación de lenguaje documental

Alternativas	Cantidad	Porcentaje
Si	22	91.67%
No	2	8.33%
Total	24	100%

Figura 18

Unificación de lenguaje documental



Análisis e interpretación

El 92% de los encuestados responden que se debería unificar el lenguaje controlado con el lenguaje natural, esto en los registros que se están depurando se puede realizar debido a que existen registros que como términos de indización se ha colocado el título

del documento dividido en varias partes o palabras más frecuentes que se encuentra dentro del documento, pero esto conlleva a que los documentos no sean recuperados y se produce el silencio documental.

15. ¿A su criterio las bibliotecas de la Universidad de Cuenca deberían contar con un instrumento unificado de indización?

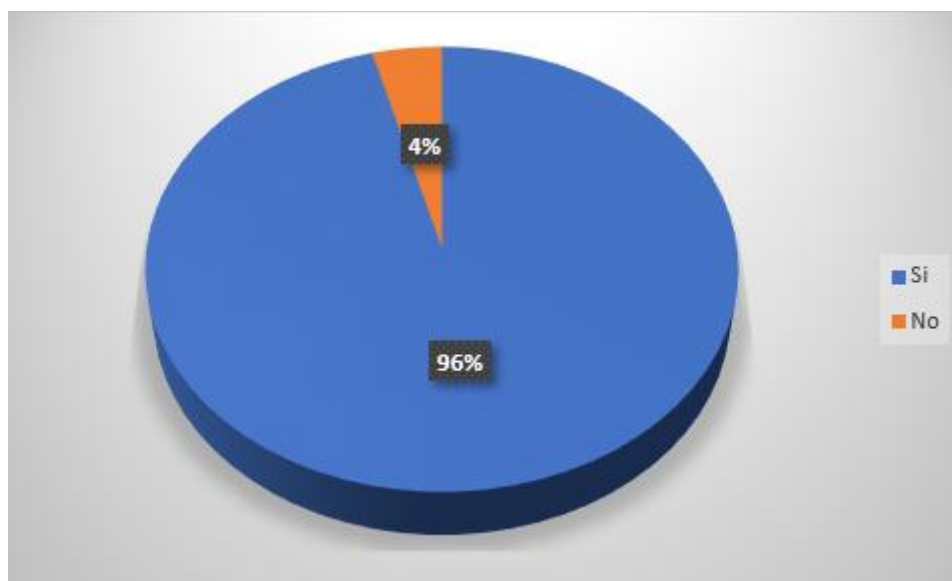
Tabla 23

Instrumento de indización

Alternativas	Cantidad	Porcentaje
Si	23	95.83%
No	1	4.17%
Total	24	100%

Figura 19

Instrumento de indización



Análisis e interpretación

El 96% de los encuestados responde que sí están de acuerdo en contar con un documento propio para el Centro, analizando sus respuestas que se les pidió que argumenten el porqué de su respuesta manifiestan que al contar con un instrumento

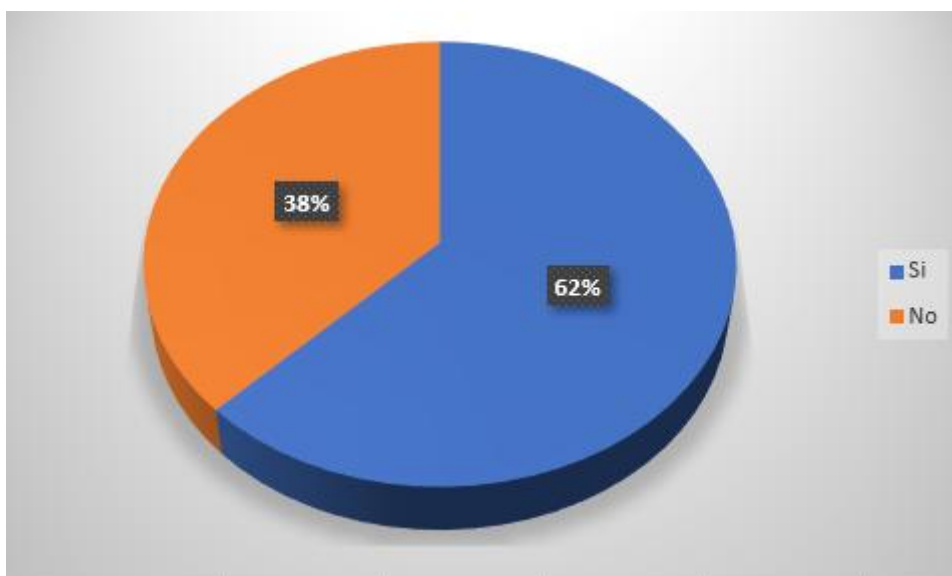
propio del Centro de Documentación, estaríamos hablando un mismo lenguaje, estaríamos dentro de un estándar internacional, ayudaría en la recuperación de la información, para poder controlar el vocabulario utilizado y ser cosechados a nivel internacional, también manifiestan que se está trabajando en la normalización de todos los servicios porque esto ayudaría para optimizar el tiempo y las actividades que se realizan.

16 ¿Usted considera que los usuarios recuperan información relevante en el OPAC?

Tabla 24
Recuperación de la información

Alternativas	Cantidad	Porcentaje
Si	15	62.50%
No	9	37.50%
Total	24	100%

Figura 20
Recuperación de la información



Análisis e interpretación

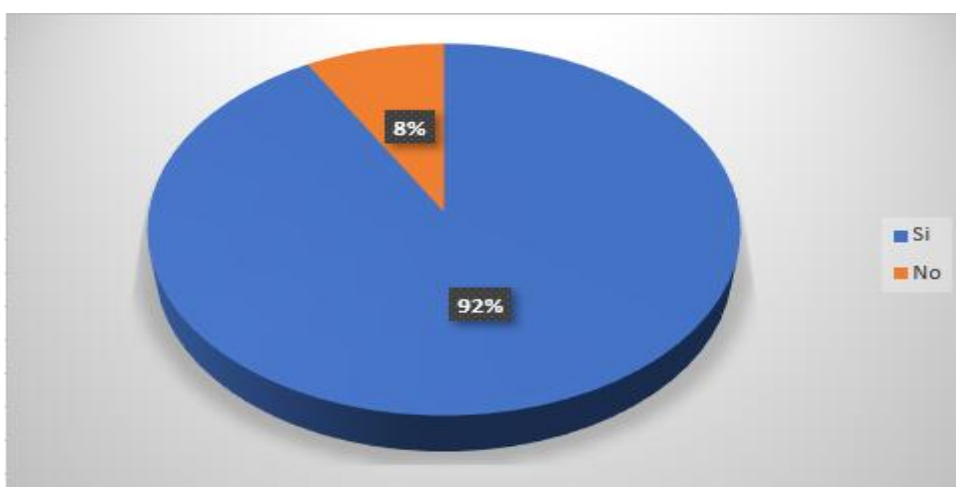
El 68% de los encuestados responden que si recuperan información relevante del Opac, su argumento es que los usuarios del Centro de Documentación Regional, reciben capacitaciones en manejo de los sistemas bibliotecarios como también de bases digitales y gestores bibliográficos, también manifiestan que si recuperan información relevante puesto que la indización está acorde a las necesidades de los usuarios, también que en el Opac pueden recuperar información tal como títulos, autores, resumen, índice de contenido, descriptores y la ubicación del material bibliográfico.

17 ¿Usted considera que la información del Repositorio Institucional es cosechada por usuarios externos?

Tabla 25
Cosecha de información

Alternativas	Cantidad	Porcentaje
Si	22	91.67%
No	2	8.33%
Total	24	100%

Figura 21
Cosecha de información



Análisis e interpretación

Con respecto a la interrogante que si la información alojada en el repositorio institucional es cosechada por usuarios externos la respuesta es el 92% que sí, puesto que todos conocemos que varios investigadores de otras universidades o de otras ciudades piden información a los bibliotecarios encargados de referencia presencial y virtual de turno cuando encuentran documentos restringidos, solicitando se les facilite la información de la misma manera se puede evidenciar en los reportes bibliométricos que lo realizan periódicamente para conocer el estado de como la información está siendo cosechada.

CAPÍTULO V

CONCLUSIONES, RECOMENDACIONES, BIBLIOGRAFÍA Y ANEXOS

5.1. Conclusiones

Luego de haber realizado un análisis cualitativo y cuantitativo sobre los diferentes tipos de lenguajes documentales utilizados en las bibliotecas universitarias que ayudan a la recuperación de la información, por medio de una recopilación bibliográfica de terminología variada, así como también de los resultados obtenidos de la encuesta aplicada a 24 profesionales de la información del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, se llegó a las siguientes conclusiones:

- El termino lenguajes documentales no es igual que lenguajes de indización, puesto que lenguaje documental puede ser de vocabulario controlado o de vocabulario natural, mientras que los lenguajes de indización son explícitamente de vocabulario controlado, por lo que, el 100% de encuestados consideran que este último sí es importante para el tratamiento documental.
- En base a los resultados obtenidos de la encuesta, se estima que para el proceso de la recuperación de la información el 83,3% del total de profesionales de la información del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca, emplean el lenguaje controlado, mientras que a comparación del 37,5% restante hacen uso del lenguaje libre
- Por otro lado, la tipología del lenguaje documental es muy variada, se ha podido destacar que existen dos grandes grupos como son: el lenguaje controlado y el lenguaje libre o natural, siendo el primero el más recomendado para su uso debido a que este tipo de lenguaje evita el ruido o el silencio documental.
- Con relación al enunciado anterior se puede confirmar que para la Indización de documentos bibliográficos se ayudan de instrumentos de lenguaje controlado como: Tesoros, Listas de encabezamiento de materia, Clasificaciones (CDU, CDD) y Listados de palabras clave.

- Para la recuperación de la información, los lenguajes de interrogación se componen de técnicas y símbolos para extraer la información de acuerdo a ciertas restricciones o descripciones emitidas por el usuario, dada esta situación los encuestados en un 62% consideran que si se recupera información relevante, debido a que se realiza educación informacional a los usuarios, ya que se imparte capacitaciones en el manejo del Opac, repositorio institucional y bases digitales, en la cual se les enseña la manera de recuperar información validada.
- Los encuestados afirman que la información producida por el Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca es cosechada o recuperada por usuarios externos, esto se puede evidenciar mediante los análisis bibliométricos realizados por el personal bibliotecario y los rankings de las universidades ecuatorianas, en donde la Universidad de Cuenca se encuentra en octavo lugar,
- Finalmente se puede evidenciar que es importante la normalización del lenguaje documental debido a que permite optimizar los procesos de la recuperación de la información

3.2. Recomendaciones

De acuerdo a los resultados de la encuesta realizada a los profesionales de la información del Centro de Documentación Regional Juan Bautista Vázquez de la Universidad de Cuenca puedo recomendar lo siguiente:

- Que los profesionales de la información tanto bibliotecarios como especialistas de Biblioteca deben contar con los debidos permisos para poder acceder a cada uno de los módulos de las plataformas y sistemas de Gestión bibliotecaria, ya que al existir restricciones el personal se ve limitado al momento de realizar correctamente el tratamiento documental.
- Es importante que la capacitación al personal de biblioteca en el manejo de lenguajes documentales se debe realizar constantemente debido a que la información evoluciona a la par con la tecnología.
- También es necesario contar con un instrumento de lenguaje controlado, como un tesoro, propio para el Centro de Documentación, el cual será de beneficio para mejorar la indización del material bibliográfico, estandarización de terminología, agilizar la selección de términos y optimizar esfuerzos y el tiempo.

3.3.BIBLIOGRAFÍA

Biblioteca Universitaria de Alicante. (2020). *La búsqueda de información científica*.

Aliante: Universidad de Alicante.

CEPAL, B. d. (18 de diciembre de 2020). *Gestión de datos de investigación*.

Obtenido de Gestión de datos de investigación:

<https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/metadatos>

Chaumier, J. (1986). *Análisis y lenguajes documentales : el tratamiento lingüístico de la información documental*. Barcelona: Mitre.

Docunecta, E. d. (10 de diciembre de 2020). *Gestion documental*. Obtenido de Qué son los metadatos: definición, tipos y ejemplos:

<https://www.docunecta.com/blog/que-son-los-metadatos>

Estrada Villacís, M. E. (2015). *Lenguajes documentales y lenguaje natural en la recuperación de información*. Obtenido de <https://www.infotecarios.com/>: <https://www.infotecarios.com/lenguajes-documentales-y-lenguaje-natural-en-la-recuperacion-de-informacion/#.Y2sC93bMJkh>

García-Jiménez, A., Rodríguez-Mateos, D., & Catalina-García, B. (2019). Estudio sobre la indización /etiquetado y los lenguajes documentales en cinco diarios españoles. *Scire: representación y organización del conocimiento*, 25(1), 55-64.

Gil Urdiciain, B. (1996). *Manual de lenguajes documentales*. Madrid: Ediciones Trea.

Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. d. (2014). *Metodología de la Investigación* (Sexta ed.). México: McGraw Gill Education.

Iglesias Maturana, M. T. (1997). Beneficios de la utilización de lenguajes controlados en el análisis y recuperación de información. *Investigación Bibliotecológica*, 32-45.

- Juncà Campdepadrós, M., & Martínez Ferreras, D. (2009). *Lenguajes documentales*. Cataluña: Universitat Oberta de Catalunya.
- Martín Gavilán, C. (2011). Lenguajes documentales Principales tipos de clasificación Encabezamientos de materia, descriptores y tesauros. En *Temas de Biblioteconomía* (págs. 1-25).
- Martínez Ferreras, D. (2017). *Los tesauros*. Guadalajara: Universidad de Guadalajara.
- Martínez, L. A., & Murillo, J. F. (2009). Tesauros y lenguajes documentales. En G. d. Aragón (Ed.), *Actas de las I Jornadas de Museos Aragoneses* (págs. 20-22). Zaragoza: Departamento de Educación, Cultura y Deporte.
- Mochón-Bezares, G., Méndez-Rodríguez, E., & Sorli-Rojo, Á. (2017). Etiquetado social y blog-scraping como alternativa para la actualización de vocabularios controlados: aplicación práctica a un tesoro de Biblioteconomía y Documentación. *Información, Cultura y Sociedad*, páginas 20-29.
- Naumis Peña, C. (2006). *Los tesauros documentales y su aplicación en la información impresa, digital y multimedia*. México: Alfagrama.
- Pinto Molina, M. (1995). Lenguajenatural y lenguaje documental. En M. Pinto Molina, *Lenguajes documentales* (págs. 173-185).
- PowerData. (19 de marzo de 2016). *PowerData*. Obtenido de PowerData: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/que-son-los-metadatos-y-cual-es-su-utilidad>
- Rubio Liniers, M. C. (2020). El análisis documental: indización y resumen en bases de datos especializadas. En *El análisis documental* (págs. 1-13).
- Sánchez Luna, B. E. (2012). Lenguajes documentales. *Anuario de Bibliotecología*, 61-77.
- Slype, G. V. (1991). *Los lenguajes de indización: concepción, construcción y utilización en los sistemas documentales*. Fundación Germán Sánchez.

Soler Monreal, M. C. (2009). *Evaluación de vocabularios controlados en la indización de documentos mediante índices de consistencia entre indizadores*. Valencia: Universidad Politécnica de Valencia. Obtenido de <http://hdl.handle.net/10251/6341>
<https://dx.doi.org/10.4995/Thesis/10251/6341>

Triola, M. F. (2009). *Estadística*. México: Pearson Educación.

3.4. ANEXOS

CUESTIONARIO PARA PROFESIONALES DE LA INFORMACIÓN DEL CDR

“JBV”

Este instrumento tiene por objetivo “Analizar los diferentes tipos de lenguajes natural y controlado, para la recuperación de la información en las bibliotecas universitarias”. Se solicita contestar con sinceridad las preguntas planteadas.

1. Biblioteca a la que pertenece

Biblioteca Campus Central
Biblioteca Campus Centro Histórico
Biblioteca Campus Paraíso
Biblioteca Campus Yanuncay

2. Tipo de biblioteca

General
Especializada

3. ¿Cuál es su grado académico?

Bachiller
Tercer nivel
Cuarto Nivel
Otro ---- -----

4. ¿Qué título académico posee?

5. ¿Cuál es su denominación del puesto que desempeña actualmente?

Especialista de Biblioteca
Bibliotecario
Técnico bibliotecario

6. ¿Qué plataforma de ingreso de información maneja usted? (puede seleccionar más de 1

opción)

Koha
Dspace
Registro de publicaciones
Biblioteca digital

7. Tipos de registros bibliográficos que usted procesa (puede seleccionar más de 1 opción)

Libros
Revistas

Tesis
Artículos
Material audiovisual
Todos los anteriores

8. ¿Qué actividades usted realiza en el procesamiento de la información? (puede seleccionar más de 1 opción)

Catalogación
Clasificación
Indización
Todas las anteriores

9. ¿Qué tipo de lenguaje documental utiliza?

Lenguaje controlado

Lenguaje natural

10. ¿Tiene usted conocimiento si se maneja algún tipo de lenguaje documental estandarizado para la indización de material bibliográfico? ¿Cuál?

**11. ¿Usted cree que es importante el proceso de la indización en el tratamiento documental?
¿Por qué?**

12. ¿Qué lenguaje documental usted utiliza para la indización? (puede seleccionar más de 1 opción)

Tesauros

Listas de encabezamiento de materia

Listas de descriptores

Lenguaje natural

Palabras claves

13. ¿Cuáles son las dificultades que usted atraviesa al momento de la indización?

Manejo de lenguaje documental para indización

Desconoce los lenguajes documentales que se usan en la indización

14. ¿Con qué frecuencia cree usted que se debe realizar capacitaciones sobre el manejo del lenguaje documental?

Mensual

Trimestral

Semestral

Anual

15. ¿Desde su percepción, los descriptores que se coloca a los registros bibliográficos ayudan a recuperar la información?

SI

NO

A veces

16. ¿Usted considera que el uso de lenguaje controlado ayuda en la interoperabilidad?

Sí

No

17. ¿Cuál lenguaje documental sería el adecuado utilizar en el registro bibliográfico?

Lenguaje natural

Lenguaje controlado

Ambos

18. ¿Cree que sería útil unificar el lenguaje natural con el controlado? ¿Por qué?

19. ¿A su criterio las bibliotecas de la Universidad de Cuenca deberían contar con un instrumento unificado de indización?
