

## DISEÑO DE REPOSITORIO DIGITAL PARA USO EN INSTITUCIONES DE NIVEL SUPERIOR

Ing. Itzel López Miranda  
Estudiante de la Maestría en  
Sistemas Computacionales en un  
Programa PNPC.  
*Tecnológico Nacional de México*  
*/ IT de Acapulco* Acapulco,  
Guerrero,  
México.ilm.miranda18@gmail.co  
m

M.I.D.S. Alma Delia de Jesús  
Islao  
Docente de la Maestría en  
Sistemas  
Computacionales.*Tecnológico*  
*Nacional de México / IT de*  
*Acapulco* Acapulco, Guerrero,  
México.  
alma.islao.ita@gmail.com

M.T.I. Rafael Hernández Reyna  
Docente de la Maestría en  
Sistemas  
Computacionales.*Tecnológico*  
*Nacional de México / IT de*  
*Acapulco* Acapulco, Guerrero,  
México.rafaelhernandez.reyna@g  
mail.com.mx

Dra. Miriam Martínez Arroyo  
Docente de la Maestría en  
Sistemas Computacionales.  
*Tecnológico Nacional de México /*  
*IT de Acapulco* Acapulco  
Guerrero, México.  
Miriamms\_ds@hotmail.com

**Resumen**— En este artículo, se plasma el trabajo interdisciplinario de la Maestría en Sistemas Computacionales con apoyo del CONACYT, impartida en el Tecnológico Nacional de México / Instituto Tecnológico de Acapulco. El artículo tiene por meta presentar el desarrollo analítico de los repositorios de tesis, como parte del movimiento de acceso abierto a la información, el uso y la aplicación de trabajos realizados. Esto se ha convertido en el medio por excelencia para lograr el adecuado registro, acceso, visibilidad y preservación de la producción científica que crean las comunidades universitarias a largo plazo.

**Palabras clave**— Python, MVC, Metadatos, Repositorio Digital. Tesis Electrónicas.

### INTRODUCCIÓN

La cantidad de repositorios que se encuentran en las instituciones hoy en día ha ido en aumento significativo debido a que sea expandido rápidamente en la última década. La información digital ha sido el factor principal en este crecimiento. Un ejemplo de este crecimiento en el número de *RIS* que se observa en el *Ranking Web de Repositorios* del Mundo, en su edición de julio de 2017, que contabiliza 2.283 repositorios institucionales en el mundo, de los cuales 72 corresponden a España en esta categoría, por poner un ejemplo. [1].

Cabe destacar que los repositorios digitales institucionales comenzaron a crearse en todo el mundo a principios del 2000, como una forma de reunir y ofrecer acceso digital, libre y gratuito de las evidencias de trabajos realizados por los estudiantes de las mismas unidades académicas. [2].

En años recientes se ha implementado el desarrollo de herramientas que cubran la necesidad de gestionar diversos documentos como son tesis, tesinas, artículos, reporte de actividades, ponencias etc. Que permitan obtener el grado académico correspondiente al nivel de estudios que se esté cursando, los cuales brindan conocimiento e

información útil para las consultas requeridas por los estudiante, académicos e investigadores que con fines educativos, empresariales, gubernamentales se pueda beneficiar de la información recopilada en estos centros.

Estos son llamados repositorios digitales que permiten tener un control para preservar y almacenar una colección de datos, y de esta manera disminuir costos dentro de los espacios delimitados para el resguardo de información dentro de las organizaciones e instituciones.

Es por eso que en este artículo se propone el diseño de un repositorio, por el cual los usuarios puedan consultar y obtener información previamente almacenada.

A continuación se describe la estructura por la que está conformada este artículo, en la sección 2 se explica el marco conceptual de la estructura a seguir, en la sección 3 se muestra la explicación del diseño de la herramienta a desarrollar, en la sección 4 se explican los diferentes tipos de repositorios existentes y sus diferencias, en la sección 5 se explica la organización y el uso de las tecnologías en los repositorios.

### OBJETIVO GENERAL

Se hace mencionar las tecnologías y herramientas necesarias para la implementación del repositorio digital para la Universidad Americana de Acapulco, comprender el proceso con la finalidad de almacenar, difundir, visualizar y proporcionar acceso a los trabajos digitales de investigación.

### MARCO CONCEPTUAL

A continuación, se describen las herramientas principales utilizadas para el análisis y el desarrollo del sistema:

*Enterprise Architect*

Es una herramienta de diseño y modelado visual basada en UML, la plataforma soporta: el diseño y construcción de sistemas de software; modelado de procesos de negocio; y modelado de dominios basados en la industria. Es utilizado por empresas y organizaciones no solo para modelar la arquitectura de sus sistemas, sino también para procesar la implementación de estos modelos en todo el ciclo de vida del desarrollo de la aplicación. [3].

Con el apoyo de esta herramienta se han creado los diagramas para dejar más clara la comprensión del funcionamiento del sistema, además de identificar los roles que tienen las personas involucradas en el manejo de esta.

#### *Visual Studio Code*

Es un editor de código fuente desarrollado por *Microsoft* para *Windows*, *Linux* y *macOS*. Incluye soporte para depuración, control de *Git* integrado, resaltado de sintaxis, finalización de código inteligente, fragmentos de código y refactorización de código. También es personalizable, de modo que los usuarios pueden cambiar el tema del editor, los métodos abreviados de teclado y las preferencias. Es gratuito y de código abierto. [4].

El código combina la interfaz de usuario optimizada de un editor moderno con asistencia y navegación de código enriquecido y una experiencia de depuración integrada, sin la necesidad de un *IDE* completo. *Visual Studio Code*, cuenta con herramientas de *Debug* hasta opciones para actualización en tiempo real de nuestro código en la vista del navegador y compilación en vivo de los lenguajes que lo requieran.

#### *Python*

Python es un lenguaje de programación el cual basa su estructura en una sintaxis que favorece la creación de un código legible y reutilizable. [5]. Se trata de un lenguaje de programación multiparadigma, este soporta programación orientada a objetos, programación imperativa y programación funcional, esta última a menor medida. Es un lenguaje interpretado, dinámico y multiplataforma. Una de las ventajas principales de este es la posibilidad de crear un código con gran legibilidad, que ahorra tiempo y recursos, lo que facilita su comprensión e implementación.

Entre los usos que se le da a este lenguaje es en el análisis de datos e implementación de inteligencia artificial, cuenta con un gran número de librerías para procesamiento de datos, y

herramientas como *Django* para el desarrollo web. [6].

La elección de estas herramientas se debe al hecho de que se busca la creación de una herramienta de bajo costo y que logre ser desarrollada de manera rápida, por lo que se utilizan el lenguaje de programación de *Python* en conjunto con el *Framework de Django* y la biblioteca de estilos de *Bootstrap*.

#### *Django*

Django es un *framework web* de alto nivel que permite el desarrollo rápido de sitios *web* seguros y mantenibles. Desarrollado por programadores experimentados, Django se encarga de gran parte de las complicaciones del desarrollo *web*, por lo que puedes concentrarte en escribir tu aplicación sin necesidad de reinventar la rueda. Es gratuito y de código abierto, tiene una comunidad próspera y activa, una gran documentación y muchas opciones de soporte gratuito y de paga. [7].

#### *Bootstrap*

Bootstrap tiene un soporte relativamente incompleto para *HTML5* y *CSS 3*, pero es compatible con la mayoría de los navegadores web. La información básica de compatibilidad de sitios *web* o aplicaciones esta disponible para todos los dispositivos y navegadores. Existe un concepto de compatibilidad parcial que hace disponible la información básica de un sitio web para todos los dispositivos y navegadores. [8].

#### DISEÑO DE LA HERRAMIENTA PROPUESTA

##### *Diagrama de arquitectura de conexión.*

A continuación se presenta el modelado de la solución a la problemática, la cual permitirá que la recopilación de información de las tesis y tesisnas de las instituciones de nivel superior se puedan realizar por medio del desarrollo de un repositorio digital, el cual centralizará las diferentes fuentes de información que hoy en día se recopilan de manera impresa y otorgará más control de la información, permitiendo el registro de cada uno de los formatos que resulten ser fuente de información, como resultado de alguna investigación o estancia por parte de los alumnos, guardando los datos en el servidor del sistema, otorgando confiabilidad de la información generada por cada registro. Siendo de suma importancia para la recaudación de información en menor tiempo posible, con respecto a la validación de los contenidos en cuanto título y descripción de cada registro, y de esta manera tener el control en el sistema de cada

consulta de las diferentes secciones del repositorio siendo estas de vital importancia para el uso y manejo del mismo.

El sistema propuesto para el repositorio de los trabajos efectuados por los alumnos del plantel es un complemento, por lo que cada alumno podrá darse de alta por medio de un usuario en el sistema digital que será desarrollada, obteniendo acceso al sistema web y con la posibilidad de realizar consultas correspondientes al interés que cada alumno respectivamente este efectuando.

Con el nuevo sistema, los alumnos ya no tendrán la necesidad de realizar impresos para el resguardo de su trabajo dentro de la institución, si no podrán efectuarlo desde cualquier dispositivo con acceso a internet, debido a que este sistema estará alojado en el servidor del instituto, gracias a la infraestructura con la que cuenta el plantel.

El sistema del repositorio estará siempre en comunicación con el servidor del plantel y conectado a la base de datos, podemos observar la propuesta de conexión que se propone utilizar en el plantel. [9]. (Véase figura 1)

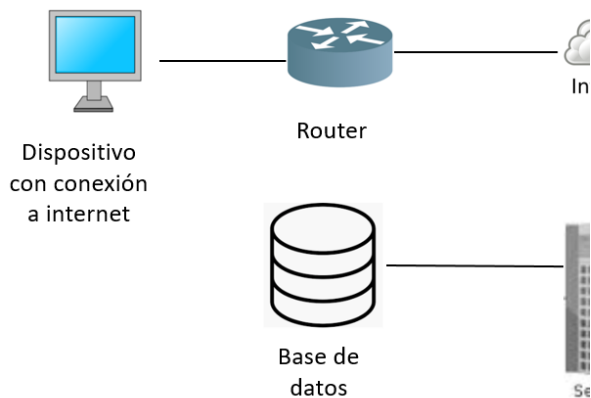


Figura 1. Diagrama de conexión

#### Diagrama de contexto

Un diagrama de contexto consiste de flujos de datos y flujos de control, almacenes de datos y un solo proceso, que consiste en una sola burbuja. El nombre dentro del proceso suele ser el nombre del sistema completo o un acrónimo convenido.

Agregando a lo anterior, se observa un diagrama de contexto, el cual nos ayuda en nuestro modelado, con este tipo de proyecciones podemos entender los objetivos que se pueden conseguir, (Véase figura 2). El diagrama de contexto ayuda a tener claro nuestro entorno y utiliza por lo menos cuatro

puntos como base, el primero nos permite visualizar como es que se pretende implementar el sistema, el segundo nos arroja la estructura y arquitectura global y el comportamiento del desarrollo, y el tercero nos permite entender las planillas que nos darán las guías por las cuales podremos empezar la construcción de la estructura y por último el cuarto punto este tipo de diagramas de modelado otorgan la documentación de las decisiones que hemos adoptado.[10]

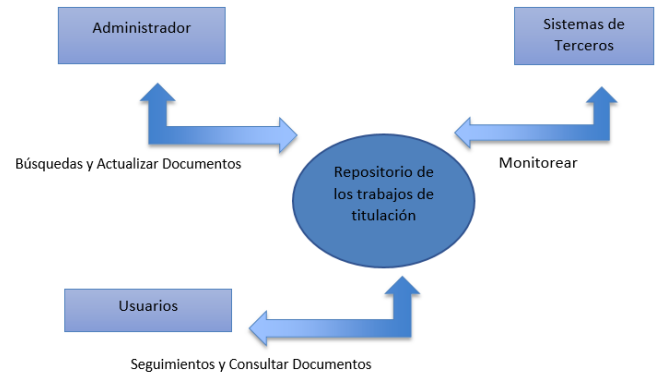


Figura 2. Diagrama de contexto.

- **Administrador:** se encarga de mantener el sistema completo en cuanto a información, documentar correctamente las fuentes de información entrantes y actualizar las existentes, claro está que la administración solo será encargada a personal seleccionado por los directivos y así mantener este tipo de información aun que publica es verdad, pero a buen resguardo para su consulta.
- **Usuarios:** son encargados de actualizar informaciones importantes, como los datos correspondientes y los diferentes reportes que continuamente suban para el seguimiento de sus avances, los cuales les permitirán poder realizar búsquedas con mayor eficiencia y una completa relación en las palabras seleccionadas con este fin.
- **Sistema de tercero:** pueden visualizar de forma cotidiana que tipo de información se mantiene al instante y así poder obtener más información, esto debido a que no podemos mantener la información de manera centralizada y cerrada al público en general, las personas que quieran hacer uso de las fuentes de información podrán visualizar y el repositorio digital sin hacer uso de las posibilidades de modificación.

### Diagramas de casos de usos

Los diagramas de casos de uso son importantes para visualizar, especificar y documentar el comportamiento del elemento, estos diagramas facilitan que los sistemas, subsistemas y clases sean abordables y comprensibles, al presentar una vista externa de cómo pueden utilizarse estos elementos en un contexto dado. Los diagramas de casos de uso también son importantes para probar sistemas ejecutables a través de ingeniería directa y para comprender sistemas ejecutables a través de ingeniería inversa. En este apartado se describen los actores que intervienen en el uso de repositorio digital. (Véase Fig. 2,3). [10]

**Administrar recursos** es el encargado de mantener el sistema así como actualizado, así de manera viable y de fácil acceso, manteniendo todo de la mejor manera posible.

**Administrar Usuarios** es el encargado de mantener el orden a los usuarios del sistema y así estos mismos podrán consumir del repositorio la información, pero sin poder modificar caso alguno.

**Administrar categoría** este caso de uso hace referencia al control de roles en los usuarios, pues claro está que esto delimita quien logra acceder y modificar cada uno de los apartados del repositorio.

**Crud de recursos, Crud de categoría, Crud de usuarios** este apartado de los casos de uso hace referencia a la parte interna y como se interacciona con la base de datos en la actualización, modificación, creación y eliminación de datos, respectivamente en cada uno de los módulos que comprenden el sistema digital.

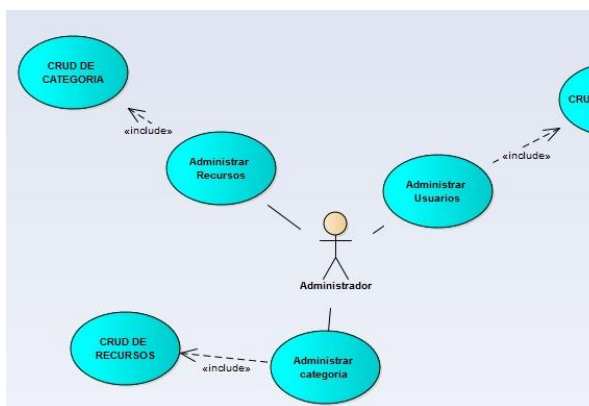


Figura 3. Diagrama de caso de uso (Administrador).

**Leer detalle recurso** cada uno de los apartados del repositorio debe de ser visible para el usuario, pero esta ventana deberá de estar llena con la descripción de los detalles del apartado.

**Utilizar repositorio** uno de los casos de uso más importantes es el de utilizar el repositorio. Y esto con el fin de poder hacer uso de los apartados y así beneficiarse de la información y los diferentes beneficios que esto conlleva.

**Buscar recursos** dentro del repositorio tendremos tesis, tesinas, reportes de residencias y más recursos, producto de la recopilación de información, así el usuario podrá buscar cada uno de ellos y enriquecer su ser con información en abundancia de manera ordenada y con fácil acceso.

**Registrarse** para ayudar a recopilar la información el usuarios deberá de realizar un registro en nuestro sistema, y así podrá el mismos rellenar los campos y depositar en el repositorio algunas de las fuentes de información que se le solicite en su camino como estúdiate.

**Loginearse** una vez que el usuarios se registre y este dado de alta, podrá entrar en el sistema, mediante un **login** que requería el nombre de usuario y contraseña para entrar al sistema.

**Navegar por categoría** los módulos serán los que delimitaran este tipo de categorías, y son las cuales nos permitirán tener un sistema ordenado y ágil en cuanto a búsqueda se refiere.

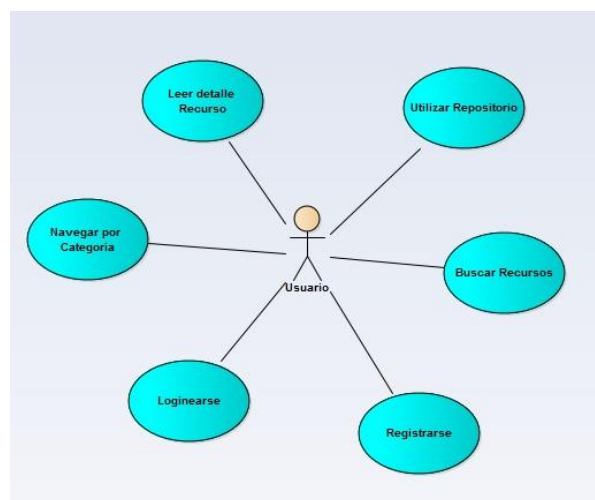


Figura 4. Diagrama de caso de uso (Usuario)

Los repositorios están constituidos por un conjunto de archivos digitales en representación de productos científicos y académicos que pueden ser accedidos por los usuarios. El contenido de un repositorio puede ser alojado en su servidor privado o referenciar hacia la ubicación del alojamiento original del archivo.

#### *Características de un repositorio*

- Almacenan diferentes tipos de archivos
- Cada objeto contiene metadatos
- Ofrece un sistema de gestión y validación de las publicaciones
- Cuenta con un sistema de brusquedad que agiliza la localización de los objetos
- Incrementa la difusión y visualización de contenidos

Hay varias formas de clasificar los repositorios digitales, el criterio más común distingue entre los institucionales y los temáticos.

#### *A. Repositorios Institucionales.*

Desarrollado por organismos políticos, sociales y educativos como universidades e institutos o asociaciones, para depositar, usar y preservar la producción científica y académica que generan en formato digital y haciéndolo accesible al público. Y así de esta manera ofrece un servicio de acceso abierto.

#### *B. Repositorios Temáticos.*

Son creados por grupos de investigadores, una institución etc. que reúnen documentos relacionados con un área temática particular donde la temática suele ser social, de educación ciudadana o académica.

#### *C. Repositorios Datos.*

Repositorios que almacenan, conservan y comparten los datos de las investigaciones, además se entienden como el material registrado, aceptado por la comunidad científica y necesario para validar los resultados de la investigación que sirven para determinar más claro lo que se puede compartir y bajo que contexto:

- Según el formato
  - Textos
  - Números
  - Imágenes
  - Otros
- Procesos de obtención
  - Experimentales
    - Secuencias genéticas
    - Cromatografías
  - Simulaciones
    - Modelos climáticos
    - Modelos económicos
  - Observaciones
    - Encuestas
    - Experimentos irrepetibles
- Según objetivo recogida
  - Específicos
    - Solo de interés para un proyecto de investigación
  - Alcance medio
    - De interés para una disciplina concreta
  - De interés general
    - De interés para la ciencia en su conjunto e incluso de interés social
- Según fase de investigación
  - Datos preliminares
    - Datos recién extraídos sin ningún tipo de procesamiento. Denominados en inglés raw data
  - Datos finales

Datos que ya han sido procesados y combinados con otros (en inglés final research data.)

#### *D. Acumulativo y perpetuo.*

La aceptación de materiales es constante y procura establecer una normativa para el mantenimiento de los recursos digitales en el tiempo a partir de su preservación en este tipo de ambiente.

#### *E. Abierto e interoperable*

Se requiere libre acceso a usuarios fuera de la institución, acceso vía múltiples herramientas de búsqueda y descubrimiento

#### *Repositorio Propuesto.*

Este repositorio académico e institucional, el repositorio es una plataforma digital que recoge, preserva y difunde la producción académica de una institución y/o de una disciplina científica, permitiendo el acceso a los objetos digitales que contiene y a sus metadatos. Contiene además una serie de mecanismos que permiten importar,

identificar, almacenar, preservar, recuperar y exportar un conjunto de objetos digitales, normalmente desde un portal web. Esos objetos son descritos mediante una serie de metadatos que facilitan su recuperación.

Los tecnológicos son instituciones públicas de educación superior, su papel es servir la sociedad de acuerdo a las necesidades que ella tiene y que se detectan directamente a través de los organismos del estado. El origen de este sistema educativo se remonta al año 1948, cuando eran inaplazables, a las entidades federativas, las oportunidades de acceder a la educación tecnológica superior, ahí encontramos la razón fundamental de su creación.

Como se observa en la Figura 5 la mayoría de los repositorios son institucionales (85.6 %), con una proporción mucho más pequeña de los demás tipos de repositorios, los temáticos (8.8). Sin embargo los repositorios, temáticos son los responsables de la aportación mayor de contenidos a los repositorios de acceso abierto.

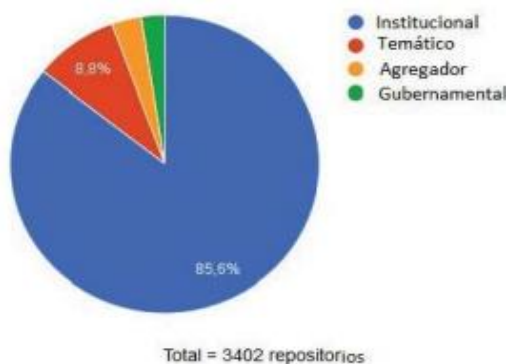


Figura 5. Tipos de repositorios de acceso abierto en el mundo

#### ORGANIZACION Y TECNOLOGÍAS UTILIZADAS EN REPOSITARIOS DIGITALES

A partir del año 1990 se dieron varios movimientos de intelectuales, en torno al acceso libre al conocimiento, donde se dan a conocer permitieron a las actividades o estándares para la creación de los repositorios digitales.

Para implementar un repositorio digital se consideran herramientas libres más utilizadas que permiten la implementación de estos repositorios son:

A continuación, se describen las herramientas principales utilizadas para los repositorios digitales:

#### *Dspace*

Es un repositorio digital donde captura, guarda, indexa y permite la consulta de la producción intelectual de los grupos y centros de investigación de universidades. Es una herramienta libre disponible para instituciones de investigación a nivel mundial, su uso se ha extendido también a instituciones privadas.

#### *E-prints*

Es un software libre que facilita la creación de repositorios virtuales, creado por la universidad de Southampton. Creado con la finalidad de crear un repositorio institucional de edición electrónica para la investigación académica, pero puede ser usado para otros propósitos.

#### *OAI-PMH (Open Archive Initiative-Protocol for Metadata Harvesting)*

OAI-PMH (Open Archive Initiative-Protocol for Metadata Harvesting) es un protocolo para la transmisión de contenidos en internet, creado por investigadores a nivel mundial a partir de su primera reunión en Octubre de 1999 en Nuevo México, USA. Cuya finalidad es desarrollar y promover estándares de interoperabilidad para facilitar la difusión eficiente de contenidos en Internet

#### *Lucene*

Es un software que permite crear buscadores de contenidos, básicamente permite indexación y búsqueda de documentos. Utilizado por numerosos proyectos, y es software libre respaldado por la fundación Apache.

#### *OpenDoar*

Es un directorio de repositorios académicos de acceso abierto. Además, proporciona información estadística de estos repositorios y permite la búsqueda de repositorios o sus contenidos.

#### *Fedora*

Posee arquitectura de gestión de activos digitales (Digital Asset Management, DAM), sobre la cual se pueden construir muchos tipos de biblioteca digital, repositorios (archivos) institucionales, archivos digitales, y sistemas de bibliotecas digitales. El repositorio tiene actualmente más de 11.000 artículo. El Contenido está disponible gratuitamente al público.

En la figura 6. Se muestra el porcentaje de los tipos de producción intelectuales que almacenan las organizaciones en sus repositorios.

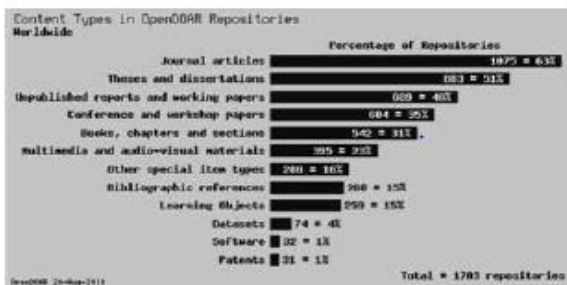


Figura 6. Repositorios por Herramientas Utilizadas.

## CONCLUSIÓN

Este artículo consta de una serie de trabajos, donde se describen las tecnologías y herramientas a utilizar, un diseño de implementación de un repositorio digital para el instituto. El repositorio permitirá organizar, mantener y consultar documentación relacionado a la producción de investigación en el instituto. Se ha hecho un estudio y selección de la herramienta que permitirá la construcción del repositorio. Así como una explicación del uso y mejora que este repositorio dará a la institución de educación superior donde sea implementada.

## REFERENCIA

- [1][2] Gonzalez, L. A. (s.f.). Diseño de un Repositorio de Objetos de Apoyo al Aprendizaje Colaborativo.
- [3] Systems, S. (09 de Febrero de 2019). *Systems, S. Enterprise Architect*. Obtenido de Sparx Systems. Obtenido de <https://sparxsystems.com/products/ea/>.
- [4] *Visual Studio Code*. (s.f.). Obtenido de [http://www.ecured.cu/Visual\\_Studio\\_Code](http://www.ecured.cu/Visual_Studio_Code).
- [5] Challenger-Pérez, I. D.-G. (2014). *El lenguaje de programación Python*. *Ciencia Holguín*, XX(2), 1-13.
- [6] Loaiza Sojos, E. L., Molina Ríos, J. R., & Loja Mora, M. N. (2016). *Evolución de los Frameworks en el Desarrollo de Aplicaciones Web con Python*.
- [7] <https://developer.mozilla.org/es/docs/Learn/Server-side/Django/Introducci%C3%B3n>. (s.f.).
- [8] <https://www.bootstrap-y-como-funciona-en-el-diseno-web>. (s.f.).

- [9] Miranda, I. L. (15 de Agosto de 2019). Implementación de un Repositorio Digital que Permita el Almacenamiento y Difusión de la Producción Académica y de Investigación. *Memorias del Congreso Internacional de Investigación Academia Journals Puebla 2019*. Puebla, México.
- [10] Alejandro, J. I. (2019). PROPUESTA DE UN SISTEMA WEB PARA REALIZAR PAGOS ELECTRÓNICOS INTERBANCARIOS DE TRÁMITES EN ESTANCIAS PUBLICAS. Congreso Internacional de Investigación Academia Journals Puebla.
- [11] Gonzalez, M. A. (2017). Implementación de un repositorio Digital para el entorno local de la Facultad de Comunicación de la Universidad de la Habana.