

ESCUELA SUPERIOR POLITECNICA DEL LITORAL



Facultad de Ingeniería en Electricidad y Computación

TRABAJO DE TITULACIÓN

“IMPLEMENTACIÓN DE UN SOFTWARE PARA GENERAR REPORTES
DINÁMICOS DE LAS NOTAS DEL MOOC HCD DE LA ESCUELA
SUPERIOR POLITÉCNICA DEL LITORAL MEDIANTE USO DE TÉCNICAS
ETL Y CUBOS DE INFORMACIÓN”

Previa a la obtención del Título de:

MAGISTER EN SISTEMAS DE INFORMACION GERENCIAL

Presentado por:

LSI. STIVEN ALEXIS PRADO PIN

ING. DOLORES MARGARITA PESANTEZ PESANTEZ

GUAYAQUIL – ECUADOR

2021

AGRADECIMIENTO

Nuestro agradecimiento a Dios y a nuestra familia por sus oraciones y apoyo constante.

A nuestra tutora y administradores HCD por su valiosa ayuda en la orientación y desarrollo de este proyecto.

A nuestros amigos que frecuentemente estaban pendientes para celebrar junto a nosotros el cumplimiento de esta meta.



LSI. Stiven Alexis Prado Pin



ING. Dolores Margarita Pesantez Pesantez

DEDICATORIA

El presente trabajo de titulación lo dedicamos a Dios, a nuestras familias, a nuestros amigos por su constante acompañamiento emocional y espiritual.

LSI. Stiven Alexis Prado Pin

ING. Dolores Margarita Pesantez Pesantez

TRIBUNAL DE SUSTENTACIÓN



MSIG. LENIN FREIRE COBO

COORDINADOR MSIG

Ph.D. OTILIA ALEJANDRO MOLINA

DIRECTORA DEL PROYECTO DE GRADUACIÓN



MSIG. OMAR MALDONADO DAÑÍN

MIEMBRO DEL TRIBUNAL

RESUMEN

El presente trabajo de titulación fue realizado en la Escuela Superior Politécnica del Litoral (ESPOL) en la Facultad de Ingeniería en Electricidad y Computación (FIEC) dirigido al curso MOOC Herramientas de Colaboración Digital (HCD) del término 2020-2T.

En cada semestre se abren dos cursos MOOC HCD, uno denominado MOOC HCD Completo, dirigido a estudiantes que deben aprobar componentes teórico y práctico y otro curso MOOC HCD Repetidores dirigido a estudiantes que tienen pendiente aprobar componentes prácticos.

Estos cursos MOOCs usan tres sistemas para el manejo de la información de los estudiantes, cuyas funciones están distribuidas de la manera: El Sistema Académico contiene las listas de los estudiantes que serán inscritos en los dos cursos MOOCs. El sistema EdX mantiene el registro de tres encuestas, de las notas del curso MOOC y del examen teórico y el Sistema Sidweb almacena los archivos y las notas de los exámenes prácticos.

El objetivo principal para el desarrollo de este proyecto es integrar los datos provenientes de los tres sistemas en una única fuente destino.

La integración de los datos beneficia a los administradores en varios aspectos, como: se reduce la carga de trabajo dedicada a la generación de reportes

consolidados y gráficos estadísticos, los datos están a la mano y son de fácil acceso.

Se plantea desarrollar reportes consolidados dirigidos a los dos cursos MOOCs, integrando los datos de los diferentes archivos sobre notas MOOC, examen teórico, examen práctico, datos demográficos, encuestas e ilustrando esto en gráficos representativos que permitan visualizar el comportamiento de los estudiantes en el curso.

Las herramientas utilizadas para la implementación del proceso ETL y la integración de los datos son MySQL con el ID Workbench y para la ilustración de reportes y gráficas se utiliza la librería Highchart.

ÍNDICE GENERAL

AGRADECIMIENTO	ii
DEDICATORIA.....	iii
TRIBUNAL DE SUSTENTACIÓN.....	iv
RESUMEN	v
ÍNDICE GENERAL.....	vii
ABREVIATURAS Y SIMBOLOS	xi
ÍNDICE DE FIGURAS.....	xiii
ÍNDICE DE TABLAS	xvii
INTRODUCCIÓN.....	xviii
CAPITULO 1	1
GENERALIDADES.....	1
1.1. Antecedentes.....	1
1.2. Descripción del Problema	3
1.3. Solución Propuesta.....	7
1.4. Objetivo general	8
1.5. Objetivos específicos.....	9
1.6. Alcance del proyecto.....	9

CAPÍTULO 2.....	12
MARCO TEÓRICO.....	12
2.1. Integración de datos	12
2.1.1. Importancia	13
2.1.2. Retos	14
2.1.3. Técnicas.....	16
2.2. Proceso ETL.....	18
2.2.1. Extracción	18
2.2.2. Transformación.....	19
2.2.3. Carga.....	20
2.3. Metodologías ágiles.....	20
2.3.1. Metodología Scrum	22
CAPÍTULO 3.....	28
ÁNÁLISIS DE LA SITUACIÓN ACTUAL.....	28
3.1. Definición de la situación actual	28
3.2. Levamiento de requerimientos y expectativa funcional.....	38
3.2.1. Fuente de datos.....	40
3.2.2. Estructura de reportes	42
3.3. Definición de infraestructura TI.....	44

3.4.	Definición de actores y casos de uso.....	47
3.5.	Planificación del uso de Scrum.....	49
CAPÍTULO 4.....		51
ANÁLISIS, DISEÑO Y DESARROLLO		51
4.1.	Análisis de requerimientos del proceso de desarrollo	51
4.1.1.	Requerimientos funcionales.....	51
4.1.2.	Requerimientos no funcionales.....	62
4.2.	Diseño del modelo de datos.....	64
4.2.1.	Definición de la arquitectura.....	64
4.2.2.	Modelo de datos fuente	66
4.3.	Proceso ETL.....	68
4.4.	Proceso de Integración de datos.....	78
4.5.	Pruebas de usabilidad	79
4.6.	Planteamiento del uso de Scrum.....	82
4.6.1.	Sprint 1 Planificación Inicial.....	82
4.6.2.	Sprint 3 – Datos integrados en la base	84
4.6.3.	Sprint 5 – Modelos de reportes.....	85
4.6.4.	Sprint 7 – Gráficas ilustrativas	86
CAPÍTULO 5.....		88

ANÁLISIS DE RESULTADOS.....88

5.1. Implementación.....88

5.2. Análisis de resultados.....100

CONCLUSIONES Y RECOMENDACIONES.....103

ABREVIATURAS Y SIMBOLOS

E:	Excel
EP:	Estudiante Práctico
ETL:	Extract, Transform y Load (Extracción, Transformación y Carga)
ETP:	Estudiante Teórico Práctico
GTSI:	Gerencia de Tecnologías y Sistemas de Información
HCD:	Herramientas de Colaboración Digital
IDD:	Integración de datos
LMS:	Learning Management System (Sistema de Gestión de Aprendizaje)
MOOC:	Massive Open Online Courses
P:	Práctico
PAE:	Periodo Académico Extraordinario
PAO:	Primer Periodo Académico Ordinario
PPT:	PowerPoint

TP: Teórico Práctico

W: Word

ÍNDICE DE FIGURAS

Figura 1.1: Proceso curso MOOC HCD	3
Figura 1.2: Repositorio de archivos MOOC HCD.....	11
Figura 2.1: Uso de Tecnologías Ágiles en la industria.....	22
Figura 2.2: Desarrollo secuencial del sprints	23
Figura 3.1: Proceso de aprobación MOOC HCD completo	29
Figura 3.2: Proceso de aprobación MOOC HCD repetidores	30
Figura 3.3: Proceso curso MOOC completo y MOOC Repetidores	38
Figura 3.4: Infraestructura TI.....	47
Figura 3.5: Caso de uso general	48
Figura 3.6: Planificación de entregables por cada Sprint	50
Figura 4.1: Estructura del Sistema	65
Figura 4.2: Proceso de carga de datos	68
Figura 4.3: Proceso de extracción Sistema Académico	69
Figura 4.4: Proceso de extracción notas MOOC.....	69
Figura 4.5: Proceso de extracción encuesta demográfica	70
Figura 4.6: Proceso de extracción encuesta de satisfacción	71
Figura 4.7: Proceso de extracción encuesta diagnóstico estudiantes MOOC repetidores	71
Figura 4.8: Proceso de extracción encuesta diagnóstico estudiantes MOOC completo.....	72
Figura 4.9: Proceso extracción archivos Sidweb.....	73

Figura 4.10: Proceso para convertir archivo csv a xlsx	74
Figura 4.11: Proceso de Transformación y Carga	75
Figura 4.12: Base de datos integrada MOOC HCD	76
Figura 4.13: Reportes MOOC HCD	79
Figura 4.14: Proceso de creación de reportes integrados.....	79
Figura 4.15: Menú MOOC HCD - Reporte Niveles de Conocimiento.....	81
Figura 4.16: Reporte Exámenes Prácticos - MOOC Repetidores	81
Figura 4.17: Reporte General elaborado por los administradores.....	83
Figura 4.18: Archivos descargados plataforma edX	83
Figura 4.19: API Proceso ETL y Base de datos	84
Figura 4.20: Creación de querys Reportes MOOC HCD	85
Figura 4.21: Querys MOOC HCD	86
Figura 4.22: Reporte examen teórico	86
Figura 4.23: Gráfica de niveles de conocimiento estudiantes teórico-práctico.....	87
Figura 4.24: Graficas del rendimiento del examen teórico por módulo.....	87
Figura 5.1: Reporte encuesta demográfica.....	89
Figura 5.2: Reporte prueba de diagnóstico.....	90
Figura 5.3: Reporte encuesta de satisfacción	91
Figura 5.4: Reporte notas curso MOOC HCD	92
Figura 5.5: Reporte examen teórico	92
Figura 5.6: Reporte de aciertos y fallos examen teórico – módulo el computador	93
Figura 5.7: Reporte de aciertos y fallos examen teórico – módulo el internet.....	93

Figura 5.8: Reporte de aciertos y fallos examen teórico – módulo PowerPoint	93
Figura 5.9: Reporte de aciertos y fallos examen teórico – módulo Excel	93
Figura 5.10: Reporte de aciertos y fallos examen teórico – módulo Word	93
Figura 5.11: Reporte de aciertos y fallos examen teórico – módulo herramientas colaborativas.....	93
Figura 5.12: Reporte rúbrica examen Excel	94
Figura 5.13: Reporte rúbrica examen Word.....	95
Figura 5.14: Reporte rúbrica examen PPT	95
Figura 5.15: Reporte puntos extras encuestas vs exámenes prácticos.....	96
Figura 5.16: Reporte nivel de conocimiento vs componentes MOOC HCD	97
Figura 5.17: Reporte nivel de conocimiento de los módulos vs cada uno de los componentes MOOC HCD	97
Figura 5.18: Reporte nivel de conocimiento El Computador	97
Figura 5.19: Reporte nivel de conocimiento El Internet	97
Figura 5.20: Reporte nivel de conocimiento PowerPoint.....	98
Figura 5.21: Reporte nivel de conocimiento Word	98
Figura 5.22: Reporte nivel de conocimiento Excel	98
Figura 5.23: Reporte nivel de conocimiento Herramientas Colaborativas	98
Figura 5.24: Nivel de conocimiento por componente MOOC HCD	98
Figura 5.25: Nivel de conocimiento prueba de Diagnóstico MOOC HCD	99
Figura 5.26: Nivel de conocimiento curso edX MOOC HCD	99
Figura 5.27: Nivel de conocimiento examen teórico MOOC HCD	99

Figura 5.28: Nivel de conocimiento examen práctico MOOC HCD	100
Figura 5.29: Reporte GTSI	100
Figura 5.30: Aplicación Web MOOC HCD	101

ÍNDICE DE TABLAS

Tabla 1: Ponderación de los módulos MOOC HCD	30
Tabla 2: Ponderación del examen teórico.....	31
Tabla 3: Rúbrica Word	32
Tabla 4: Rúbrica PowerPoint.....	33
Tabla 5: Rúbrica Excel.....	34
Tabla 6: Fuentes de datos MOOC HCD.....	41
Tabla 7: Reportes MOOC HCD.....	42
Tabla 8: Actores y Casos de uso	47
Tabla 9: Caso de uso 1 - Generar Reportes.....	48
Tabla 10: Caso de uso 2 - Gestión de archivos fuente para el proceso ETL	48
Tabla 11: Miembros Scrum	49
Tabla 12: Nivel de Conocimiento	52
Tabla 13: Rúbrica examen PowerPoint.....	52
Tabla 14: Rúbrica examen Word	55
Tabla 15: Rúbrica examen Excel	57
Tabla 16: Detalle de Reportes MOOC HCD	59
Tabla 17: Gama de colores por categoría.....	63
Tabla 18: Gama de colores por componente	63
Tabla 19: Fuente de datos.....	67
Tabla 20: Proceso ETL MOOC HCD.....	77

INTRODUCCIÓN

Los MOOC (Massive Open Online Courses) son cursos que se imparten en línea vía Web y están abiertos a una audiencia masiva con niveles de conocimientos heterogéneos.

Los cursos MOOCs desarrollados sobre una plataforma y para su administración requieren de otras plataformas académicas para evaluar exámenes presenciales sincrónicos, sumado al volumen de estudiantes registrados, implica tiempo, organización y administración de archivos complicada, lo cual conlleva a un alto nivel de errores, por lo que, la administración de estos cursos actualmente es muy demandante en carga de trabajo y tiempo de dedicación.

La ESPOL, entre los cursos MOOCs que ofrece se encuentra el de Herramientas de Colaboración Digital - HCD, dirigido a estudiantes de pregrado y admisiones. Este curso MOOC está publicado sobre la plataforma EdX y es un requisito que el estudiante debe cumplir antes de graduarse, debido a que, con él obtiene la suficiencia en ofimática.

Cada semestre se lanzan dos cursos con contenido específico, el MOOC HCD completo dirigido para estudiantes que deben aprobar los componentes teórico y práctico y el MOOC HCD repetidores para estudiantes que les falta aprobar componentes prácticos.

El curso MOOC HCD completo está compuesto por 6 módulos: El Computador, Internet, PowerPoint (PPT), Word, Excel y Herramientas Colaborativos.

El curso MOOC HCD repetidores está compuesto por 3 módulos: PowerPoint, Word y Excel.

El curso MOOC HCD maneja ocho cohortes: Teórico-Práctico, PPT, WORD, EXCEL, PPT-WORD, PPT-EXCEL, WORD-EXCEL y Práctico. Las cohortes son grupos creados para clasificar a los estudiantes registrados en el curso.

En el curso MOOC HCD Completo se utiliza únicamente la cohorte Teórico-Práctico, por ende, los estudiantes registrados en este curso son denominados estudiantes teórico práctico.

En el curso MOOC HCD Repetidores se utilizan las cohortes PPT, WORD, EXCEL, PPT-WORD, PPT-EXCEL, WORD-EXCEL y Práctico. Los estudiantes registrados en cada cohorte dependerán de los componentes prácticos que le faltan aprobar, de tal forma que, solo tendrán visible el contenido que deben aprobar. Los estudiantes registrados en este MOOC son denominados estudiantes prácticos.

Cada componente del curso MOOC debe ser aprobado con mínimo 60 puntos.

Un estudiante teórico-práctico debe aprobar el MOOC completo, luego aprobar un examen teórico cronometrado y finalmente un examen práctico con

supervisión. El componente práctico consta de 3 exámenes PowerPoint (p), Word (w) y Excel (e).

Un estudiante práctico debe aprobar el o los módulos del curso MOOC que tenga pendiente y finalmente el o los exámenes prácticos correspondientes.

En los cursos MOOC completo y MOOC repetidores se realizan tres encuestas denominadas demográfica, diagnóstico y satisfacción para extraer información relevante de los estudiantes.

El curso MOOC HCD utiliza tres plataformas para su administración: la plataforma EdX para el contenido del curso, encuestas y examen teórico, un Sistema académico desarrollado sobre .net para generar las listas de registros de los estudiantes MOOC completo y MOOC repetidores, y Sidweb desarrollado sobre el LMS Canvas para la evaluación de los componentes prácticos (Word, Excel y PowerPoint).

Para conciliar reportes de notas teóricas, prácticas y encuestas con los respectivos datos demográficos, actualmente se descargan los archivos de las plataformas mencionadas en formato .csv y para procesarlos se hace uso de la herramienta Excel aplicando filtros, tabla dinámica, funciones básicas y condicionales que permitan generar los diversos reportes consolidados para posterior visualizar e interpretar el comportamiento y rendimiento de los estudiantes.

El tiempo de dedicación para el desarrollo de cada reporte consolidado es de aproximadamente dos días y para la generación de los 21 reportes que requiere la administración del curso, es de dos semanas, dado que esta tarea está distribuida entre 3 administradores.

El año académico contiene tres semestres, dado que en cada semestre se liberan dos cursos MOOC HCD, por año se liberan 6 cursos.

El objetivo principal de este proyecto es implementar un sistema que integre todos los archivos generados por los tres sistemas, en una misma fuente de datos, de tal forma que, permita realizar consultas, generar reportes y visualizar gráficas descriptivas que ayuden a los administradores a tomar decisiones y aplicar mejoras continuas en el contenido, distribución del contenido y evaluaciones teórica y práctica.

Para la integración de datos se hará uso de un repositorio en OneDrive donde se alojarán los archivos fuentes de los tres sistemas, posterior se hará uso del proceso ETL para la respectiva extracción, transformación y carga a la nueva fuente de datos, una vez completada esta fase, se generarán los reportes consolidados y gráficos estadísticos solicitados por los administradores.

CAPITULO 1

GENERALIDADES

1.1. Antecedentes

La ESPOL con una larga y reconocida trayectoria por su modelo de aprendizaje, incursionó incorporando en el año 2017, los MOOCs como complemento al modelo curricular.

El primer curso MOOC que se lanzó fue el de Herramientas de Colaboración Digital (HCD), el cual comprende seis unidades, tres teóricas y tres prácticas.

Este curso es prerrequisito para tomar la materia Fundamentos de Programación, la cual, para ciertas carreras está ubicada en el primer nivel y para otras en el segundo nivel de las mallas curriculares de las carreras de ESPOL y además prepara al estudiante en conocimientos de ofimática que es un requisito de graduación.

El MOOC HCD se difunde a los estudiantes objetivo mediante correo electrónico y charlas informativas por parte del personal a cargo de la

administración del curso, dando a conocer las políticas, fechas importantes y preguntas frecuentes.

El curso MOOC HCD registra un aproximado entre 1000 a 2000 estudiantes por periodo, el primer periodo se da de febrero a abril (PAE – Periodo Académico Extraordinario), el segundo de mayo a septiembre (I PAO – Primer Periodo Académico Ordinario) y el tercero de octubre a enero (II PAO – Segundo Periodo Académico Ordinario). Hasta el PAE 2021 se han atendido a 25682 estudiantes.

Durante el curso MOOC los estudiantes deben completar tres encuestas, desarrollar las actividades del curso MOOC, rendir un examen teórico y 3 exámenes prácticos, correspondiente a los tres módulos prácticos comprendidos en el curso MOOC HCD. Si alguno de los estudiantes no aprueba el curso MOOC en línea o el examen teórico debe repetir todo el curso MOOC. En el caso de que apruebe el curso MOOC en línea, el examen teórico y no apruebe algunos de los componentes prácticos debe repetir solo el o los componentes prácticos reprobados.

El curso MOOC completo en donde se registra a los estudiantes teórico-práctico (tp) tiene una duración de 6 semanas y el curso MOOC repetidores en donde se registran a los estudiantes prácticos (p), tiene una duración de 3 semanas.

Cada semestre se lanzan 2 cursos, un curso MOOC HCD completo y un curso MOOC HCD repetidores, es decir, que durante el año académico se lanzan 6 cursos.

1.2. Descripción del Problema

La administración del curso MOOC HCD conlleva varias actividades como se ilustra en la Figura 1.

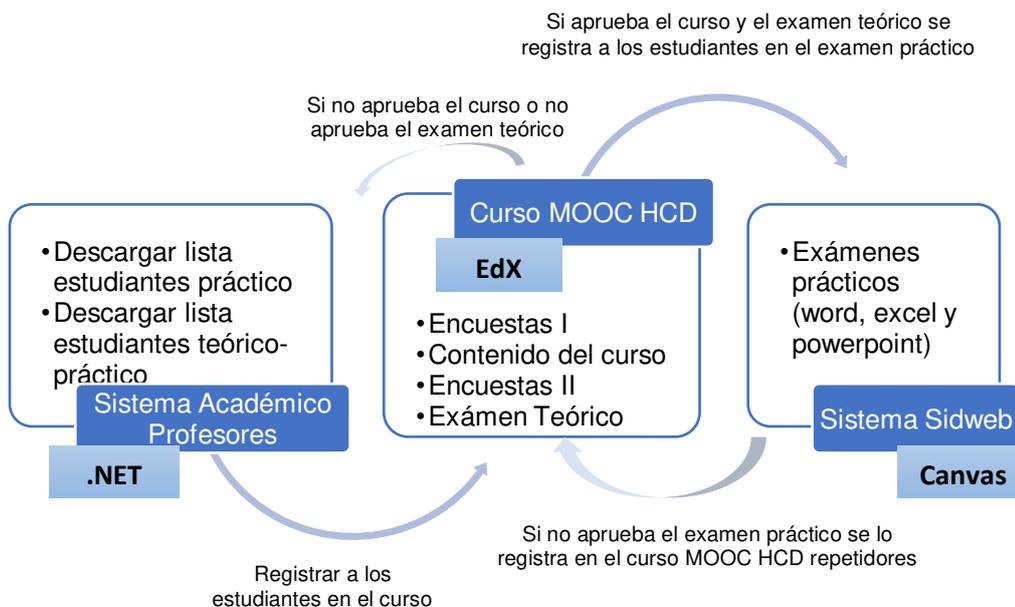


Figura 1.1: Proceso curso MOOC HCD

El curso MOOC HCD, utiliza tres plataformas para su administración el sistema académico desarrollado sobre .net para generar las listas de registros, la plataforma edX para el contenido del curso y Sidweb desarrollado sobre el LMS Canvas para la evaluación de los componentes prácticos (Word, Excel y PowerPoint).

El MOOC HCD requiere del sistema académico, debido a que, es el que administra las notas y determina el estado de aprobado y reprobado. En caso de ser aprobado el estudiante podrá registrarse en la materia Fundamentos de Programación. En caso de ser reprobado, el sistema valida que componentes son los que tienen una nota menor que 60 entre el examen teórico y práctico (Word, PPT y Excel). Si el estudiante tiene una nota menor que 60 en el examen teórico, será asignado a la lista de registro del MOOC Completo, es decir lo debe repetir. Si la nota del examen teórico es mayor o igual a 60, valida que cada una de las notas de los exámenes prácticos sea mayor o igual a 60, en el caso de que todas cumplan con la nota de aprobación el estudiante aprueba el requisito de Ofimática, caso contrario es asignado a la lista de registro MOOC repetidores.

Adicionalmente, el sistema académico cumple con el proceso de registro de los estudiantes ESPOL y Admisiones, por lo que, aquellos estudiantes que nunca han tomado el curso MOOC, denominados estudiantes novatos, son asignados a la lista MOOC completo.

El curso MOOC HCD hace uso de la plataforma edX para impartir el contenido del curso, realizar las encuestas y evaluar el examen teórico cronometrado.

Finalmente hace uso de la plataforma Sidweb para evaluar exámenes prácticos, debido a que, los docentes que ayudan con la toma de

exámenes prácticos están más familiarizados con esta plataforma, en lo que respecta a descarga de archivos entregables y asignación de calificaciones. La plataforma SIDWeb permite una mejor administración y distribución de estudiantes y docentes en las diferentes franjas horarias en que se desarrollan los exámenes prácticos.

Otro motivo por el cual se hace uso de Sidweb, es porque edX en ocasiones presenta problemas en el servidor con ejercicios donde se solicita a los estudiantes cargar archivos. Los administradores han realizado pruebas con estudiantes que suben archivos en la plataforma edX, pero el servidor los borra, esto perjudicaría a los estudiantes, debido a que, en 2 de los 3 exámenes prácticos, se requiere que los estudiantes carguen sus archivos resultados.

Actualmente, ESPOL está solicitando a una empresa externa el mantenimiento de la plataforma para controlar los fallos que se producen en esta.

Una vez finalizado el periodo de un curso MOOC, los administradores deben generar reportes para analizar el rendimiento de los estudiantes en cada etapa, curso MOOC, examen teórico, exámenes prácticos junto con los datos demográficos de los estudiantes. Adicionalmente, generan reportes de las encuestas para conocer otros datos demográficos del estudiante, el nivel de conocimiento con el que ingresa al curso y el nivel de satisfacción con el que finaliza el curso.

Para generar los reportes consolidados deben unificar datos de los tres sistemas de forma manual con ayuda de la herramienta Excel, consumiendo un tiempo aproximado de 2 semanas y con ello, generando una sobrecarga de trabajo.

Con el objetivo de sistematizar la administración del curso y optimizar los tiempos de trabajo, en este proyecto se ha planteado desarrollar la integración de la información, la generación de reportes consolidados y la visualización de gráficas estadísticas correspondientes a cada reporte.

Se han contabilizado 21 reportes, los cuales, consumen un alto tiempo de elaboración y que son solicitados para hacer análisis. Estos reportes son:

- Reporte de calificaciones MOOC HCD estudiantes teórico-práctico y práctico. (2 reportes)
- Reporte detallado de examen teórico estudiantes teórico-práctico. (1 reporte)
- Reporte detallado de las tres encuestas de los estudiantes teórico-práctico y práctico. (6 reportes)
- Reporte detallado de las calificaciones de los exámenes prácticos de los estudiantes teórico-práctico y práctico. (6 reportes)
- Reporte de las calificaciones del examen práctico con el incremento de los puntos de la encuesta de los estudiantes teórico-práctico y práctico. (2 reportes)

- Reporte nivel de conocimiento de las tres encuestas, curso MOOC HCD, examen teórico y exámenes prácticos de los estudiantes teórico-práctico. (1 reporte)
- Reporte nivel de conocimiento de las tres encuestas, curso MOOC HCD y exámenes prácticos de los estudiantes práctico. (1 reporte)
- Reporte de calificaciones consolidado estudiantes teórico-práctico y práctico. (2 reportes)

Los reportes consolidados de calificaciones se envían a la Unidad de Gerencia de Tecnologías (GTSI) para que estas notas sean asentadas en el Sistema Académico de estudiantes, de tal forma que, se libere el prerrequisito a los estudiantes que aprobaron todos los componentes del MOOC HCD y puedan registrarse en la materia Fundamentos de Programación.

1.3. Solución Propuesta

En este proyecto se propone desarrollar una aplicación web que integre los archivos generados de tres sistemas y mediante un proceso ETL se proceda a extraer, transformar y cargar los datos provenientes de diferentes fuentes para integrarlos en una base de datos robusta. A partir de esta base se generan reportes e ilustración de gráficas estadísticas utilizando la información de calificaciones, datos demográficos y encuestas de los estudiantes del curso MOOC HCD.

Este proyecto trabajará con la data del PAO 2020-2T, que servirá como proyecto piloto para los próximos cursos.

Se mostrará la reducción del tiempo de dedicación por parte de los administradores en las tareas de consolidación de los datos, generación de reportes y consultas representadas mediante tablas y gráficos estadísticos.

Para la obtención de información se realizarán sesiones de trabajo con el personal a cargo de la administración del curso MOOC HCD, con el fin de recopilar los archivos de las diversas plataformas y definir los requerimientos del equipo, validación en el diseño del sistema y finalmente la ejecución de pruebas.

Finalmente, para el desarrollo del proyecto se va a aplicar la metodología ágil Scrum, que facilitará el avance progresivo y efectivo del proyecto, lo cual involucra entregables semanales de los avances definidos.

1.4. Objetivo general

Sistematizar el proceso y optimización del tiempo de trabajo en la administración, consolidación y generación de reportes y gráficos estadísticos de las calificaciones, encuestas y datos demográficos del curso MOOC HCD a través de una aplicación web.

1.5. Objetivos específicos

- Analizar los requerimientos funcionales necesarios para identificar las mejoras a implementar, de tal forma que, se reemplace tareas manuales por tareas de sistema.
- Definir los archivos y datos a analizar provenientes del Sistema Académico, Sidweb y edX.
- Determinar los requerimientos de análisis por parte de los administradores del curso MOOC HCD.
- Integrar los datos de las diferentes fuentes haciendo uso del proceso ETL, para generar reportes.
- Creación de gráficos estadísticos por cada reporte mediante la librería Highchart.

1.6. Alcance del proyecto

El alcance de este proyecto se enfoca en los estudiantes de Admisiones y en estudiantes de ESPOL que tomaron el curso MOOC HCD completo y el curso MOOC HCD repetidores en el II PAO 2020.

Los archivos descargados de los diferentes sistemas serán depositados por el administrador en un repositorio de OneDrive, que cumpla con la estructura como se ilustra en la Figura 1.2.

Los archivos de rúbricas de exámenes prácticos de Word y PPT deben contener una identificación ya sea cédula o matrícula, de tal forma que, las notas se puedan integrar con los datos de los otros sistemas.

Adicionalmente, los archivos de exámenes prácticos Word, Excel y PPT, deben contener el siguiente formato de nombre:

tipoExamen_#Franja_#Paralelo_NombresApellidosProfesor, dado que dicha información es requerida para el reporte GTSI.

Los archivos del Sistema Académico deben contener la fecha de inicio del curso MOOC HCD, la cual es importante registrar en el reporte GTSI.

El administrador debe crear un archivo que contenga los códigos con las respuestas respectivas correspondientes a las encuestas. Al generar el reporte se realiza la comparación de códigos entre el archivo fuente y el archivo creado por el administrador para presentar la información con sus preguntas y respuestas asociadas.

El proceso ETL que ejecuta el administrador desde la aplicación, tomará los archivos del repositorio, los depurará y cargará en una base de datos MySQL. Si el administrador no deposita los archivos no se tendrá la información disponible para generar los reportes ni los gráficos.

Los reportes específicos requeridos e identificados por los administradores son 21 y estarán cargados en la aplicación en un menú vertical, de tal forma que, el administrador al dar clic sobre alguno de los

enlaces a los reportes éste se mostrará en forma de tabla con sus respectivas gráficas.

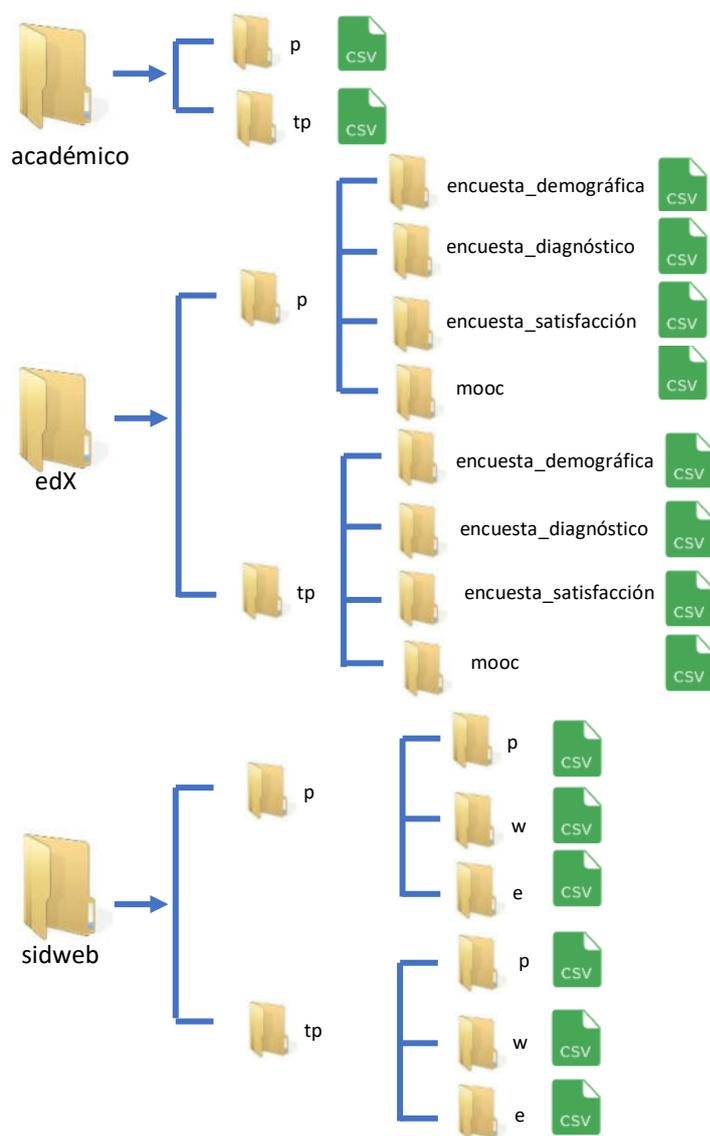


Figura 1.2: Repositorio de archivos MOOC HCD

CAPÍTULO 2

MARCO TEÓRICO

2.1. Integración de datos

Administrar diferentes plataformas para un mismo modelo de negocio, puede generar conflictos ya que los datos se encuentran almacenados en distintos servidores y al revisar los datos de forma separada no aportan para un análisis, es aquí donde es vital que se aplique la integración de datos.

La integración de datos (IDD) consiste en un conjunto de procesos técnicos y de negocio que unifica los datos de diferentes fuentes de información donde es común que se mantengan formatos distintos, con el fin de ubicarlos en una única fuente de valor y fiable, esto le brinda al usuario una vista de 360° de toda la información relevante de la institución [1] [2].

Para realizar este proceso se requiere extraer los datos de las múltiples fuentes, limpiar la data, depurarla y finalmente convertir los datos en información consistente y relevante para la toma de decisiones estratégica y de esta forma aumentar su ventaja competitiva [3].

El objetivo de la IDD es extraer el máximo conocimiento de los datos y evitar la fragmentación de la información [1].

No hay un enfoque universal en materia de integración, muchas de las técnicas que los expertos en TI utilizan están en evolución, sin embargo, la integración de datos conserva elementos comunes como: una red fuente de datos, una base de datos central y usuarios realizando peticiones de consultas [3].

2.1.1. Importancia

La IDD hoy en día es una tarea fundamental en las instituciones por las siguientes razones [3] [4]:

- Mejora la colaboración y la fusión de los datos: Todas las áreas dentro de una compañía generan y manipulan data diariamente, esta actividad debe ser colaborativa para que posibilite una visión global de la información.

- Ahorra tiempo: Al mantener la buena práctica de integración de datos, el equipo de trabajo puede generar reportes inmediatos y consolidados.
- Minimiza falencias: Recabar datos manualmente exige una alta demanda de tiempo y la mano de obra de varios empleados, ya que deben combinar los datos para generar información, esto puede conllevar errores ya que los datos son muy manipulados, por eso es importante la automatización con el uso de un software.
- Proporciona datos valiosos: Al centralizar los datos en una misma fuente se pueden identificar problemas a nivel de negocio y aplicar mejoras preventivas, permitiendo a la organización ser más productiva y competitiva.

2.1.2. Retos

La IDD involucra grandes desafíos referente a unificar datos de fuentes dispares, algunos de estos son [3] [2] [4]:

- Entender los datos: El paso inicial para una IDD es que la persona a cargo entienda los datos que va a recabar,

este paso parece muy obvio y puede estar sobreestimado, sin embargo, recopilar los datos, entenderlos y analizarlos es el cimiento crucial para alcanzar el objetivo deseado.

- **Datos heterogéneos:** Involucra unificar datos de diferentes fuentes, en diversos formatos para estandarizarlos y transformarlos a un formato compatible con la fuente destino.
- **Manejo de volúmenes de datos:** Recopilar, administrar y generar información a partir del volumen de datos es un desafío, por lo que, se debe plantear al inicio del proyecto el nivel de crecimiento del negocio y la capacidad de almacenamiento.
- **Generación de datos constantes:** Una vez integrados los datos no se debe perder de vista el objetivo de mantener las acciones de integración, para conservar una vista holística de los datos.

2.1.3. Técnicas

Para realizar una IDD se deben plantear preguntas cruciales para analizar la viabilidad y grado de dificultad del proceso, tales como: ¿De qué sistemas se obtendrán los datos? ¿Están todos los datos disponibles para cumplir con los requisitos? ¿Cuáles son las reglas del negocio? ¿quién será el administrador del sistema? Adicional definir requisitos no funcionales como el tiempo de procesamiento de datos, el tiempo de respuesta del sistema y el número de usuarios concurrentes [5].

Al seleccionar las herramientas de desarrollo se debe considerar que sean escalables, costo de implementación y soporte asequible y con preferencia que sean herramientas libres de licencias [5].

Teniendo presente las consideraciones antes expuestas, la modalidad de la IDD se basa en la magnitud del negocio, la necesidad a satisfacer y los recursos disponibles [3] [5].

- Integración de datos manual: es cuando el proceso es realizado por un usuario, el cual reúne los datos de las distintas fuentes de forma manual, los depura y los unifica en una nueva fuente. Esta técnica es aplicada en

organizaciones pequeñas donde el volumen de datos no es significativo.

- Integración de datos mediante middleware: es cuando una aplicación de middleware es la encargada de normalizar los datos y transferirlos a la fuente unificada. Esta técnica se aplica cuando el IDD no cuenta con los permisos para acceder a los datos fuente.
- Integración virtual: es cuando a partir del sistema origen se define un conjunto de vista para obtener acceso a todos los datos de la compañía. Su principal beneficio radica en que la latencia de los datos actualizados en el sistema origen a la vista consolidada es mínima.
- Integración de datos físicos: consiste en un nuevo sistema que mantiene una copia de los archivos fuente, los cuales se procesan para obtener una visión unificada de los datos. Esta técnica es la más tradicional.
- Integración a partir de aplicaciones: consiste en que una aplicación de software extrae e integra los datos en la fuente, pasando por un proceso de normalización para que los datos sean compatibles.

2.2. Proceso ETL

El proceso de extracción, transformación y carga, conocido por sus siglas en inglés como ETL (Extract, Transform and Load), es parte fundamental para la implementación de un modelo de negocios inteligente. Este proceso permite transportar datos desde fuentes de archivos estructuradas, semiestructuradas y no estructuradas.

- **Estructurados:** contenidos en bases de datos.
- **Semiestructurados:** en formatos legibles para máquinas: HTML tabulado, Excel, CSV, entre otros.
- **No estructurados:** formato no legible para máquinas: Word, HTML no tabulado, PDF, TXT, entre otros [6].

El proceso ETL permite analizar los datos una vez alojados en el destino, este puede ser una base de datos, un data mart o un data warehouse [4]. A continuación, se verán cada uno de los pasos que lo componen.

2.2.1. Extracción

La fase I del proceso ETL consiste en extraer los datos desde los sistemas origen, analizarlos e interpretarlos, de tal forma

que, cumplan con la estructura deseada, para su posterior proceso de transformación.

En esta fase se debe precautelar el análisis y la interpretación, ya que los formatos y la organización de los datos pueden variar en cada sistema [4].

2.2.2. Transformación

En esta fase se aplican a los datos las normas y reglamentos del modelo de negocio, para que presenten los formatos y estándares que garanticen la calidad de los datos. Estos son:

Normalización: Definir los atributos que intervendrán, con su respectivo tipo de datos y cómo serán almacenados en la base de datos.

Eliminación de duplicados: excluir los datos redundantes.

Verificación: realizar pruebas automatizadas para corroborar la información y los tiempos de ejecución. Esta fase permite continuar depurando los datos para evitar falencias en el sistema.

Clasificación: Maximiza la eficiencia de la base de datos, agrupando y clasificando los tipos de datos. Este proceso

permite crear tablas adicionales para los reportes e informes resumidos.

2.2.3. Carga

La fase final de este proceso es cargar los datos a un nuevo destino, que permite hacer auditoría de los datos y mantener un histórico de los mismos.

Existen dos métodos para ejecutar el proceso de carga:

Acumulación: Consiste en hacer un resumen de todas las transacciones y cargarlas como una transacción única.

Rolling: Se aplica cuando la información requiere que sea registrada en niveles jerárquicos o de mayor granularidad [4].

2.3. Metodologías ágiles

Las metodologías ágiles tuvieron origen en la industria del software, cuando las empresas de este medio se percataron que su forma tradicional de trabajar no estaba dando resultados óptimos y las entregas de los productos salían con un alto retardo.

En el año 2001 nace las metodologías ágiles, donde los CEOs de las compañías de software más grandes plantearon las mejores prácticas y desarrollaron el llamado 'Manifiesto Agil' [7].

La metodología ágil es una filosofía que propone una forma de trabajar y organizarse basados en pequeños avances progresivos hasta completar el objetivo deseado. Estas metodologías ayudan a la compañía a dirigir decisiones, proyectos y presupuestos de forma rápida, flexible y con un resultado de alta calidad que cumpla con las expectativas del cliente [8] [9].

Según la firma de servicios profesionales más grande del mundo PriceWaterhouseCoopers (PwC) los proyectos que se aplican con metodologías ágiles son un 28% más exitosos [10].

Las metodologías más usadas en el medio de desarrollo se la aprecian en la Figura 2.1 [11]:

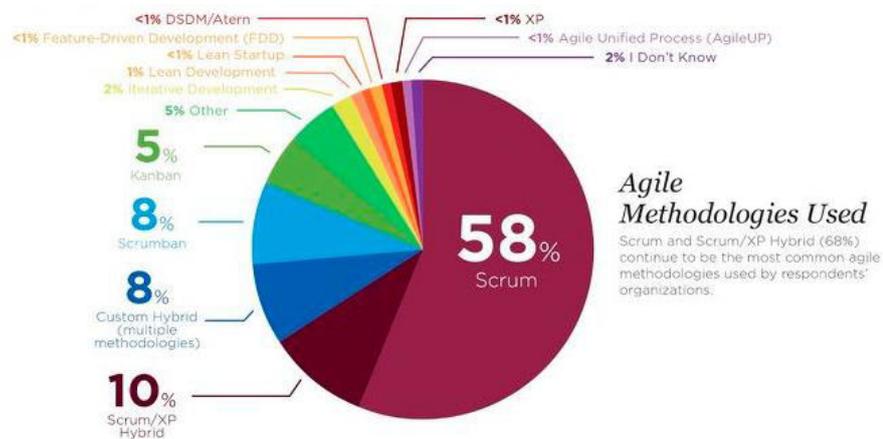


Figura 2.1: Uso de Tecnologías Ágiles en la industria
Fuente: <https://bit.ly/3hyOXgS>

Todas las metodologías ágiles aportan a las compañías con la disciplina de entregas de calidad orientadas a la flexibilidad, ya que se presentan cambios continuos en la planificación y en los requerimientos del cliente. A pesar de que las metodologías ágiles nacieron para las áreas de desarrollo de software son aplicables para varios tipos de proyectos.

2.3.1. Metodología Scrum

Consiste en trabajar por iteraciones o los denominados 'sprints'. El periodo de duración del requerimiento es de una semana. Para la entrega de cada sprints se realizan reuniones donde se definen los avances y ajustes en los requerimientos y en la planificación [10].

La secuencia de esta metodología es aplicar mejoras a las fases anteriores e ir desarrollando el nuevo sprint secuencial, de forma que se vaya puliendo todos los requerimientos hasta llegar a la fase final, como se aprecia en la Figura 2.2 [10].



Figura 2.2: Desarrollo secuencial del sprints
Fuente: <https://bit.ly/3hyon7p>

Los pilares fundamentales de esta metodología son [11]:

- Transparencia entre el cliente y el equipo Scrum.
- Inspección constante del desarrollo fase a fase.
- Adaptación a los cambios durante todo el proceso.

Los usuarios que intervienen en esta metodología son [12]:

Dueño del producto (Product Owner): Es el responsable de dar seguimiento al proyecto y por ende al equipo de desarrollo, es el que impulsa la lista del producto (Product Backlog), en la cual se detallan claramente los avances o sprint del proyecto, de tal forma que sean claras, viables y transparentes para el equipo de desarrollo.

Equipo de desarrollo (Development Teams): Son la mano de obra para la implementación del proyecto, son los encargados de entregar los sprint semanalmente. La sinergia entre el equipo de desarrollo optimiza la eficiencia y efectividad del trabajo. Las características del equipo de desarrollo son: autoorganizados, multifuncionales, no hay categorías todos están al mismo nivel, la responsabilidad de sprint recae sobre todo el equipo de desarrollo equitativamente.

El equipo de desarrollo debe ser lo suficientemente pequeño, mínimo tres, para conservar la característica de ágil, menos de tres miembros puede presentar limitaciones en cuanto a habilidades en el grupo y lo suficientemente grande, menos de nueve, para completar los requerimientos del proyecto, un grupo mayor a nueve representa mayor demanda de coordinación.

Scrum Master: Es el líder que brinda apoyo a todo el equipo Scrum.

Apoya al dueño del producto asegurando el cumplimiento de los objetivos, el alcance, orientándolo a que la lista del producto sea lo más clara y precisa para maximizar su valor.

Apoya al equipo de desarrollo impulsando la eficiencia y eficacia en la entrega de cada sprint a lo largo del proyecto, eliminando

obstáculos que limiten la productividad. Es asesor en las reuniones Scrum y en los entornos organizacionales donde no han adoptado esta metodología.

Apoya a la organización guiándolos en el correcto uso de la metodología, en la planificación de los proyectos y aumentando la productividad del equipo Scrum.

Características

Las características más relevantes de esta metodología son [12] [11]:

- Esta metodología se basa en dividir un proyecto completo en pequeños entregables, definiendo tiempos de desarrollo cortos.
- Los entregables se priorizan de mayor a menor, es decir los de mayor impacto o rentabilidad para el cliente son los primeros en ser liberados.
- Es posible liberar el producto con los requerimientos mínimamente viables, sin necesidad de esperar a la fase de desarrollo final.

- El desarrollo incremental, la obligación de entregar sprint funcionales y las constantes revisiones y demostraciones brinda la garantía de que el proyecto es de calidad.
- Validar la practicidad del proyecto en las etapas iniciales, determina la factibilidad de la implementación, sin la necesidad de realizar la inversión, evaluar la calidad software y el tiempo de desarrollo del equipo.

Ventajas

Las ventajas más destacables de Scrum son [13]:

- Permite establecer metas cuantificables.
- La flexibilidad del método no da lugar a conflictos entre el cliente y el equipo Scrum.
- Muestra un panorama completo del proyecto.
- Evita caer en el perfeccionismo que estanca y causa retardos.
- Los resultados son rápidos en periodos muy cortos.
- El método brinda transparencia en todas las etapas del proyecto.

Desventajas

Las desventajas encontradas en Scrum son [13]:

- Es fundamental que todos los miembros del equipo conozcan sobre el método.
- El Scrum Master juega un rol importante a lo largo del proyecto.
- Si existe algún retardo con algún sprint inevitablemente estancará el avance de los siguientes ya que pueden tener una relación lógica entre los incrementos.
- La metodología Scrum se puede complicar en proyectos complejos.

CAPÍTULO 3

ANÁLISIS DE LA SITUACIÓN ACTUAL

3.1. Definición de la situación actual

El curso MOOC HCD se oferta tres veces al año de febrero a abril, de mayo a septiembre y de octubre a enero, a estudiantes de pregrado y admisiones. Para cada uno de los cursos, el administrador descarga las listas de registro en formato .xlsx del Sistema Académico de profesores, distribuida de la siguiente forma:

Estudiantes Admisiones

- Lista de estudiantes MOOC HCD completo
- Lista de estudiantes MOOC HCD repetidores

Estudiantes ESPOL

- Lista de estudiantes MOOC HCD completo
- Lista de estudiantes MOOC HCD repetidores

El administrador registra a los estudiantes de admisiones y ESPOL en la plataforma edX, en los cursos MOOCs respectivos:

Curso MOOC HCD completo

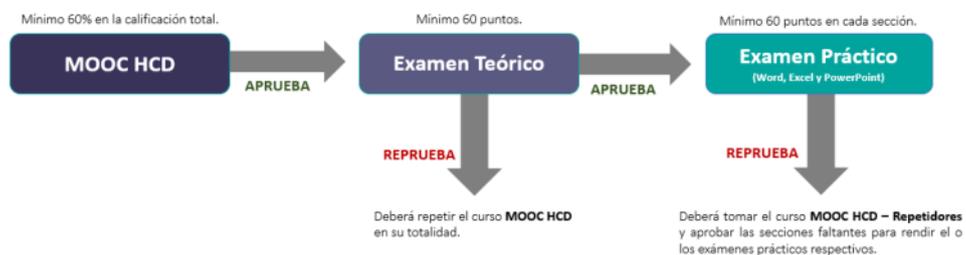
- Estudiantes ESPOL teórico práctico
- Estudiantes ADMISIONES teórico práctico

Curso MOOC HCD repetidores

- Estudiantes ESPOL práctico
- Estudiantes ADMISIONES práctico

Los estudiantes para obtener el requisito de ofimática pasan por diferentes etapas como se ilustra en la Figura 3.1.

PROCESO MOOC HCD - COMPLETO



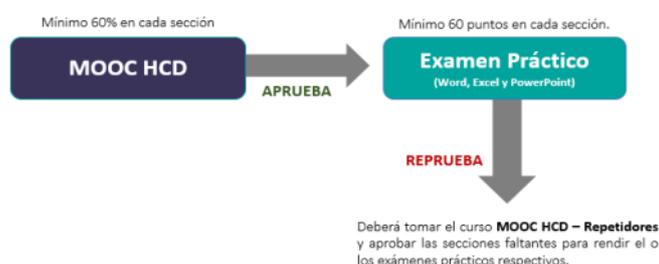
NOTA: Una vez aprobado el curso MOOC HCD, tendrá convalidado el requisito de Ofimática y se le habilitará la materia "Fundamentos de Programación".

Figura 3.1: Proceso de aprobación MOOC HCD completo

Fuente: <http://url.ie/1tzo4>

Los estudiantes que no aprueban todos o alguno de los componentes prácticos deben tomar el curso MOOC HCD Repetidores y completar el proceso como se visualiza en la Figura 3.2.

PROCESO MOOC HCD - REPETIDORES



NOTA: Una vez aprobado el curso MOOC HCD, tendrá convalidado el requisito de Ofimática y se le habilitará la materia "Fundamentos de Programación".

Figura 3.2: Proceso de aprobación MOOC HCD repetidores
Fuente: <http://url.ie/1tzo4>

El curso MOOC HCD está compuesto por seis módulos [14], cada uno de ellos tiene asignado un porcentaje que sumado entre ellos da 100, como se muestra en la Tabla 1.

Tabla 1: Ponderación de los módulos MOOC HCD

MOOC HCD	
Módulos	Ponderación
Módulo 1 - El Computador	5
Módulo 2 - Internet y sus Herramientas Básicas	5
Módulo 3: ¿Cómo crear presentaciones profesionales en PowerPoint?	30
Módulo 4 - Sacándole el jugo a los datos en Excel	25
Módulo 5 - ¿Cómo crear documentos profesionales en Word?	30
Módulo 6 -Trabajo Colaborativo en la Web	5
TOTAL	100

Una vez finalizado el curso MOOC HCD, el administrador descarga de la plataforma edX el reporte de calificaciones que se genera en formato .csv. Con este reporte filtra a los estudiantes que aprobaron el curso, es decir, obtuvieron una nota mayor o igual a 60 y les activa el examen teórico.

Este examen es evaluado en la plataforma edX, está compuesto por 50 preguntas de opciones múltiples y es calificado sobre 100 puntos, distribuidos como se muestra en la Tabla 2.

Tabla 2: Ponderación del examen teórico

Examen Teórico		
Módulos	N. Preguntas	Ponderación
Módulo 1 - El Computador	12	24
Módulo 2 - Internet y sus Herramientas Básicas	12	24
Módulo 3: ¿Cómo crear presentaciones profesionales en PowerPoint?	7	14
Módulo 4 - Sacándole el jugo a los datos en Excel	7	14
Módulo 5 - ¿Cómo crear documentos profesionales en Word?	7	14
Módulo 6 - Trabajo Colaborativo en la Web	5	10
TOTAL	50	100

El administrador nuevamente descarga de la plataforma edX el reporte de calificaciones (.csv), para extraer la nota del examen teórico. Aquellos estudiantes que obtuvieron una nota mayor igual a 60, serán registrados en la plataforma Sidweb, donde rendirán los exámenes

prácticos Word, PowerPoint y Excel, los cuales son calificados en archivos de Excel de siguiente manera:

- Word: Es calificado por un calificador automático por un software, el cual genera como resultado una rúbrica con el detalle de la calificación por tema.
- PowerPoint: Es calificado por los docentes mediante una rúbrica en xlsx clasificado por temas.
- Excel: Es de tipo evaluación con preguntas de opciones múltiples, que se califica de manera automática y es configurada directamente en Sidweb.

Las rúbricas de los componentes Word, PowerPoint y Excel se detallan en la Tabla 3, Tabla 4 y

Tabla 5 respectivamente. Estas rúbricas indican los temas y subtemas que serán evaluados en cada uno de los componentes.

Tabla 3: Rúbrica Word

Examen Word		
Temas	Subtemas	Puntos
Tabla de Contenido	Genera de Tabla de Contenido	5
	Aplica Estilos	10
	Asocia Multinivel	10
Formato del Documento	Interlineado	5

Examen Word		
Temas	Subtemas	Puntos
	Espaciado	5
	Fuente	4
	Tamaño	3
	Color	3
Bordes	Color	3
	Contorno	4
	Grosor	3
Nota al pie		5
Pie de pagina		8
Letra Capital		7
Columnas		10
Saltos		5
Viñetas		10
TOTAL		100

Tabla 4: Rúbrica PowerPoint

Examen PPT		
Temas	Subtemas	Puntos
Distribución de Contenidos	Una idea principal por diapositiva	10
	Los títulos adecuados	10
	Diapositivas no sobrecargadas	10
Representación adecuada de la información con (Viñetas, Tablas o SmarArt)	Inserta viñetas, tablas o smarArt al momento de presentar información	6
	Buenas prácticas en el uso de formatos	4
Número de Diapositivas		5
Imágenes	Hace uso de todas las imágenes	5

Examen PPT		
Temas	Subtemas	Puntos
	Buenas prácticas en el uso de formatos	5
Portada	Tipo de Fuente	2
	Tamaño	2
	Formato de fuente(N/K/S)	1
	Color de Relleno de Texto	2
	Contorno y Efecto de Texto	3
Títulos	Tipo de Fuente	3
	Tamaño	3
	Formato de fuente(N/K/S)	1
	Color de Relleno de Texto	4
	Contorno y Efecto de Texto	4
Texto	Tipo de Fuente	4
	Tamaño/Color	3
	Interlineado/Alineación	3
Fondo	Trama	5
	Color	5
TOTAL		100

Tabla 5: Rúbrica Excel

Examen Excel		
Subtemas	N. Preguntas	Puntos
Funciones Básicas	5	20
Ordenar – Primer Nivel	1	9
Ordenar – Segundo Nivel	1	11
Filtrar – Primer Nivel	1	9

Examen Excel		
Subtemas	N. Preguntas	Puntos
Ordenar – Segundo Nivel	1	11
Tabla de Contenido	4	25
SI Sencillo	1	5
SI Combinado	1	10
TOTAL	15	100

Adicionalmente, el administrador realiza tres encuestas dentro del curso MOOC HCD, dos de ellas al inicio y otra al finalizar el curso. Las encuestas abarcan los siguientes temas:

Encuesta 1.1: Demográfica

Encuesta 1.2: Diagnóstico

Encuesta 2.1: Satisfacción

Las dos primeras encuestas tienen la finalidad de conocer y evaluar las condiciones y conocimientos de los estudiantes previos al curso.

La última encuesta busca recabar la opinión del estudiante al finalizar el curso, con respecto a la distribución, actividades y contenido impartido.

Los resultados de estas encuestas son descargados de la plataforma edX en formato .csv.

Finalmente, el equipo del curso, conformado por tres miembros, un administrador y dos asistentes, descargan los archivos de tres fuentes Sistema Académico, edX y Sidweb, para obtener los datos de los estudiantes y los resultados de cada etapa.

El equipo integra la información de los archivos de forma manual para generar un conjunto de reportes que les permita analizar la información y toma de decisiones.

El equipo realiza un informe general con los resultados obtenidos en los cursos MOOC HCD completo y MOOC HCD repetidores, donde incluyen cuadros estadísticos con los resultados de cada etapa y de las encuestas realizadas, con el fin de detectar puntos críticos para la aplicación de mejoras continuas. Esta tarea es desarrollada haciendo uso de Excel y de funciones avanzadas como: tabla dinámica, filtros, funciones lógicas, referencias absolutas, relativas y mixtas.

Debido a la extensión del trabajo y a otras tareas asignadas en paralelo, la elaboración de los reportes es compartido entre los tres miembros del grupo, tomándole aproximadamente cuatro semanas para su desarrollo.

Adicionalmente, se envía un reporte con las calificaciones del curso MOOC HCD completo y otro con las calificaciones del MOOC HCD

repetidores a GTSI, para que este departamento asiente las notas en el sistema académico de profesores y de esta manera los estudiantes que aprobaron todos los componentes se le convalide el requisito de ofimática y se le habilite el registro a la materia Fundamentos de Programación. A los estudiantes que no aprobaron alguno o todos los componentes, se les asigna tomar el curso MOOC HCD nuevamente según el o los componentes reprobados.

En la Figura 3.3 se ilustra la intervención de los tres sistemas para administrar el curso MOOC completo y MOOC repetidores con los respectivos archivos de salida de cada componente.

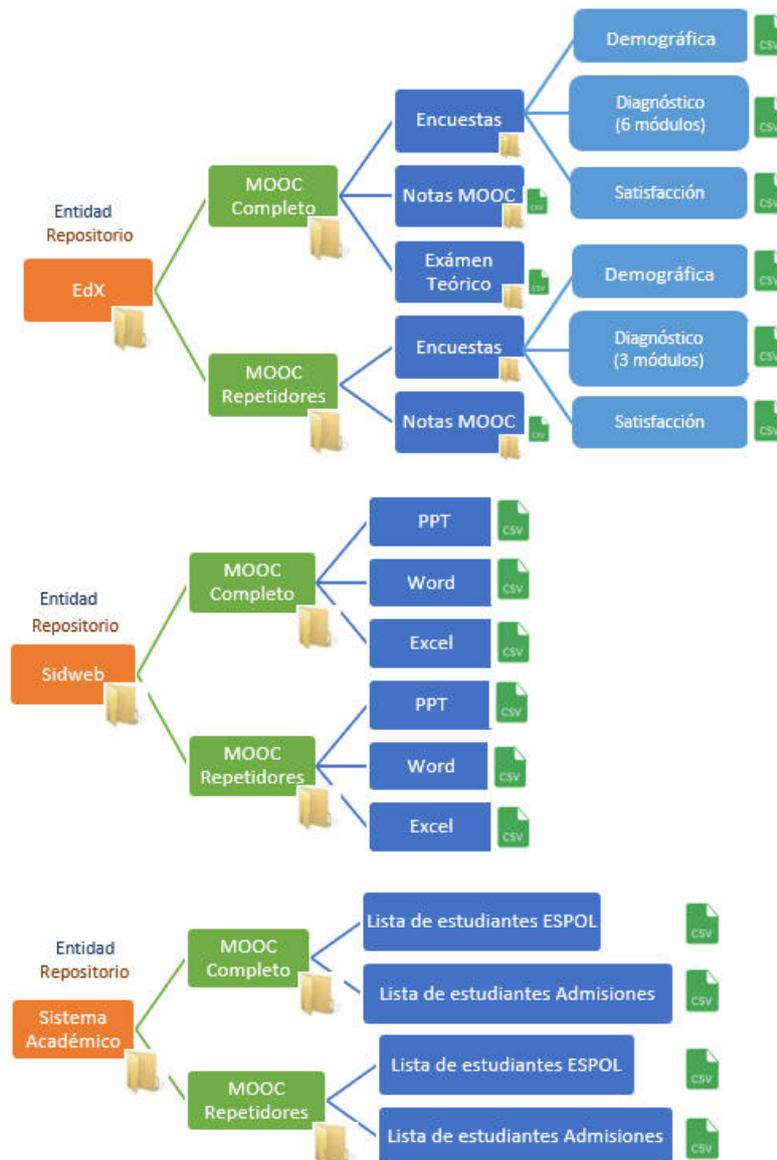


Figura 3.3: Proceso curso MOOC completo y MOOC Repetidores

3.2. Levamiento de requerimientos y expectativa funcional

Para el levantamiento de la información se entrevistó a los tres miembros del grupo MOOC HCD, en donde explicaron todo el proceso

que se genera desde el inicio del curso hasta su cierre, con la generación de los reportes.

El equipo indicó que el tiempo entre que un curso cierra y otro inicia es muy corto, por lo que, su mayor complicación radica en el desarrollo de los reportes de resultados, porque deben estar en este proceso y a la par en los preparativos para el inicio del nuevo curso.

Adicionalmente, una vez que los miembros del grupo generan los reportes parciales con los resultados por etapa, hacen una revisión cruzada de las tareas, con el fin de minimizar errores, por lo que destacó que el tiempo de dedicación para esta tarea les conlleva una sobrecarga de trabajo.

Los reportes de resultados que generan son importantes para la aplicación de mejoras, pero dado que este reporte general lo finalizan una vez que el nuevo curso inicia, las mejoras identificadas del curso anterior las deben aplicar para el próximo curso que ocurre 4 meses después.

Para la aplicación de mejoras en el curso, deben revisar en varias ocasiones reportes de cursos anteriores, porque se deben revisar archivos históricos para obtener la información deseada.

Por todo lo antes expuesto, el equipo tiene la necesidad de tener una única fuente de información que le permita integrar los datos, mantener un histórico, generar los reportes de forma inmediata y visualizar la información mediante gráficos estadísticos.

El equipo del curso proporcionó los archivos de las tres fuentes de datos y los reportes parciales y generales, para iniciar con el desarrollo del software.

Se acordó reuniones periódicas para esclarecer dudas al momento de la limpieza de los datos y para el mejor entendimiento del modelo de negocio.

3.2.1. Fuente de datos

En la Tabla 6 se presenta el Sistema o Plataforma de donde se extraen los datos, los diferentes archivos que se descargan de cada una de ellas y la descripción de los datos que contienen.

Tabla 6: Fuentes de datos MOOC HCD

Fuentes de datos		
Sistema/ Plataforma/ Repositorio	Archivos	Descripción
Sistema Académico	Lista de registro estudiantes de admisiones (MOOC completo)	Datos demográficos, el tipo de colegio, la carrera, número de veces que ha tomado el MOOC HCD
	Lista de registro estudiantes de admisiones (MOOC repetidores)	Datos demográficos, el tipo de colegio, la carrera, número de veces que ha tomado el MOOC HCD y las notas de los exámenes obtenidos en el curso anterior
	Lista de registro estudiantes de ESPOL (MOOC completo)	Datos demográficos, el tipo de colegio, la carrera, número de veces que ha tomado el MOOC HCD
	Lista de registro estudiantes de ESPOL (MOOC repetidores)	Datos demográficos, el tipo de colegio, la carrera, número de veces que ha tomado el MOOC HCD y las notas de los exámenes obtenidos en el curso anterior
Plataforma edX	Reporte de calificaciones general	Datos del promedio general del curso, promedio por módulo, nota del examen teórico, resultado de la encuesta 1.2 - Prueba de diagnóstico.
	Reporte de calificaciones detallado	Datos de puntos obtenidos por cada pregunta del examen teórico.
	Reporte de respuestas por ejercicio	Datos de las encuestas 1.1 – Demográfica

Fuentes de datos		
Sistema/ Plataforma/ Repositorio	Archivos	Descripción
		2.1 – Satisfacción
Sistema Sidweb	Archivos de los exámenes de Excel	Datos de puntos obtenidos por cada pregunta del examen.
Repositorio Onedrive	Rúbricas Examen Word	Datos de puntos obtenidos por cada tema del examen.
	Rúbricas Examen PowerPoint	Datos de puntos obtenidos por cada tema del examen.

3.2.2. Estructura de reportes

En base a los reportes parciales proporcionados por los miembros del equipo MOOC, se identificaron 21 reportes de interés a elaborar a nivel de sistema, los cuales se detallan en la Tabla 7.

Adicional a los reportes identificados, se debe generar gráficos estadísticos correspondientes a la información presente en cada uno de los ellos.

Tabla 7: Reportes MOOC HCD

MOOC HCD	
Reportes	Grupo de estudiantes
Reporte detallado de examen teórico	MOOC completo
Reporte detallado de curso MOOC edX	

MOOC HCD	
Reportes	Grupo de estudiantes
Reporte detallado de la encuesta demográfica	Estudiantes ESPOL y Admisiones
Reporte detallado de la prueba de diagnóstico	
Reporte detallado de la encuesta Satisfacción	
Reporte detallado de las calificaciones examen práctico Word	
Reporte detallado de las calificaciones examen práctico PPT	
Reporte detallado de las calificaciones examen práctico Excel	
Reporte de las calificaciones del examen práctico con el incremento de los puntos de la encuesta	
Reporte consolidado de calificaciones	
Reporte nivel de conocimiento detallado	
Reporte detallado de curso MOOC edX	
Reporte detallado de la encuesta demográfica	
Reporte detallado de la prueba de diagnóstico	
Reporte detallado de la encuesta Satisfacción	
Reporte detallado de las calificaciones examen práctico Word	
Reporte detallado de las calificaciones examen práctico PPT	
Reporte detallado de las calificaciones examen práctico Excel	
Reporte de las calificaciones del examen práctico con el incremento de los puntos de la encuesta	
Reporte consolidado de calificaciones	
Reporte nivel de conocimiento detallado	

3.3. Definición de infraestructura TI

Para la implementación del proyecto se hará uso del siguiente recurso técnico detallado por áreas como se ilustra en la Figura 3.4.

Front - End: Es la interfaz web, la parte visible con la que el usuario interactúa, comprendida por contenido en tablas, un menú y enlaces y que le permiten al usuario navegar por las diferentes páginas.

Para la creación de la interfaz de este proyecto usaremos:

- Bootstrap: Es un framework CSS de código abierto, con una gama de estilos predefinidos adaptables para el uso de diferentes dispositivos.
- HTML5: Es un lenguaje basado en etiquetas que permite estructurar el contenido de una página web, para incorporar múltiples objetos. Los principales de interés para este proyecto son tablas, gráficos, menús, descargas en diferentes formatos tanto para los reportes como para los gráficos.
- CSS3: Es un lenguaje de diseño gráfico que va de la mano con HTML5 para incorporar estilos y formatos al contenido web.

Back-End: Enfocado en el correcto funcionamiento de la página web, con respecto a la información a mostrar y validaciones a considerar de acuerdo con el modelo de negocio.

- PHP: Lenguaje de programación para entorno web. Compatible para conectar el servidor o base de datos con la interfaz.
- MYSQL Workbench: Es un entorno visual de base de datos. Está disponible de forma gratuita. Entre sus funciones está el diseño, desarrollo, creación, mantenimiento y optimización de la base de datos.

Herramientas de desarrollo: Son el complemento necesario para llevar a cabo el desarrollo de la aplicación web.

- Git con Github: Facilita el control de versionamiento del proyecto, de tal forma que, varios usuarios manipulen el código y puedan registrar el avance de forma individual siendo esto visible para todos los miembros del equipo.
- XAMP: Es un servidor de código libre. Incluye servidor Web Apache, servidor de base de datos MYSQL con su respectivo gestor phpAdmin e incluye el intérprete PHP.

- Highchart: Librería desarrollada en javascript para la creación de gráficos estadísticos. Es de uso gratuito para páginas no comerciales.
- Visual Studio Code: Es un editor de código (IDE) de Microsoft, soporta varios lenguajes de programación, entre ellos PHP, el cual usaremos en este proyecto.
- API desarrollado en PHP: La Interfaz de programación de aplicaciones conocida como API, tiene la funcionalidad de permitir la comunicación entre dos o más programas de software, de tal forma que, se simplifica la administración, diseño y uso de aplicaciones.
- Postman: Es una herramienta que permite testear, consumir y depurar API, permitiendo la monitorización de varios parámetros, entre ellos el tiempo que tarda en procesar las peticiones HTTP al servidor y los posibles errores de ejecución de los servicios del API, para corregirlos antes de mandar a producción.



Figura 3.4: Infraestructura TI

3.4. Definición de actores y casos de uso

Se desarrolla el diagrama de casos de uso para la generación y visualización de reportes como se visualiza en la Tabla 8, con el fin de identificar los procesos relacionados con cada uno de los actores involucrados.

Tabla 8: Actores y Casos de uso

Actores	Casos de uso
Coordinadora MOOCs	<ul style="list-style-type: none"> • Generar Reportes • Visualización de gráficas
Equipo MOOCs	<ul style="list-style-type: none"> • Gestión de archivos fuente para el proceso ETL • Publicación de reportes • Publicación de gráficas

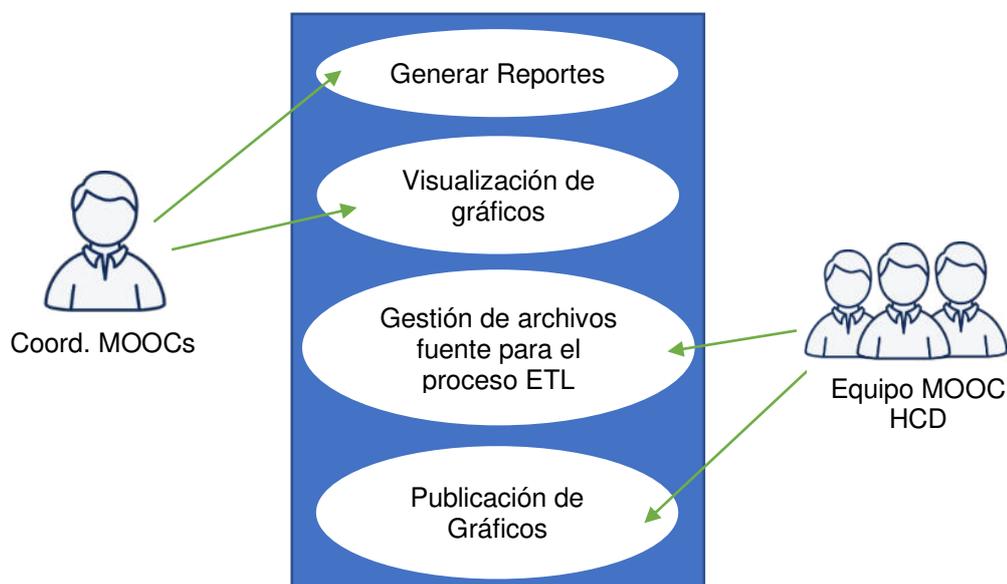


Figura 3.5: Caso de uso general

Tabla 9: Caso de uso 1 - Generar Reportes

Identificación	Casos de uso
Caso de Uso 1	Generar Reportes
Actor	Coordinadora MOOCs
Propósito	Seleccionar el reporte de acuerdo con los requerimientos
Descripción	El caso de uso inicia cuando la coordinadora MOOCs selecciona de una lista de reportes, el reporte de interés, el cual se visualiza para su descarga junto con las gráficas correspondientes.

Tabla 10: Caso de uso 2 - Gestión de archivos fuente para el proceso ETL

Identificación	Casos de uso
Caso de Uso 2	Gestión de archivos fuente para el proceso ETL
Actor	Equipo MOOC HCD
Propósito	Ubicar los archivos fuente en el repositorio central
Descripción	El caso de uso inicia cuando el equipo MOOC HCD ubica los archivos fuente en el repositorio central y el proceso de integración de datos se activa de forma automática. De esta forma se ejecuta el proceso ETL y se almacena la información en la

Identificación	Casos de uso
	base de datos, actualizando la lista de reportes y gráficos.

3.5. Planificación del uso de Scrum

Se estima desarrollar el proyecto, haciendo uso de la metodología ágil Scrum en 18 semanas, como se ilustra en la Figura 3.6. Los miembros del equipo están distribuidos como se muestra en la Tabla 11.

Tabla 11: Miembros Scrum

Usuarios	Roles
Otilia Alejandro	Scrum Master
Dolores Pesantez	Product Owner
Stiven Padro	Develoment Teams

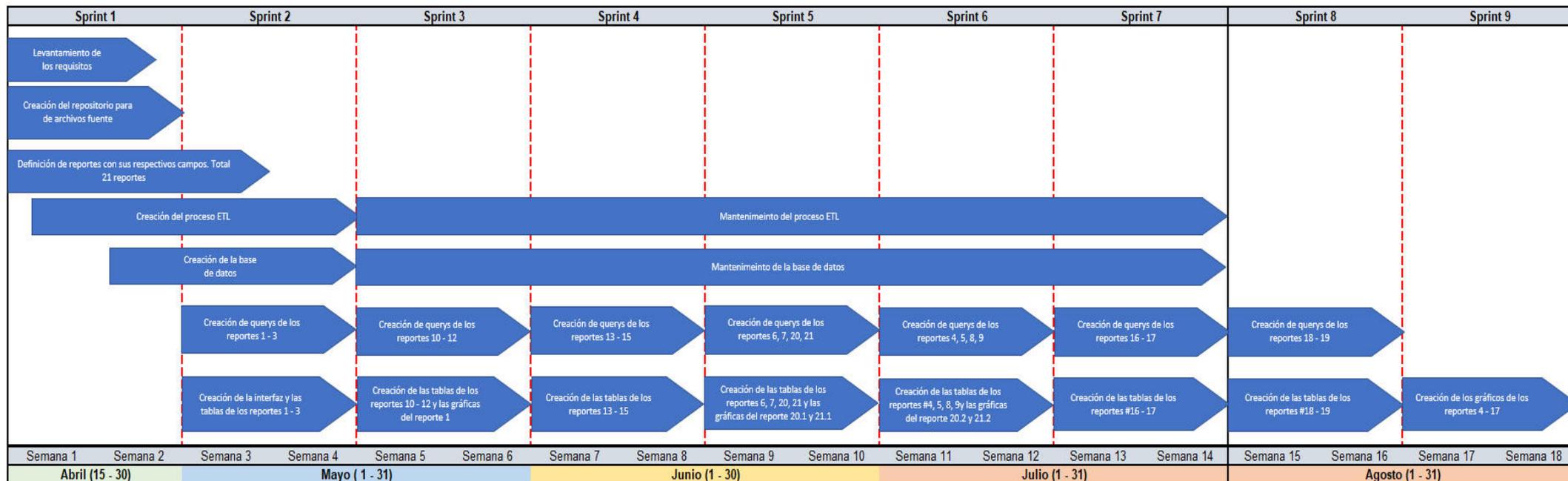


Figura 3.6: Planificación de entregables por cada Sprint

CAPÍTULO 4

ANÁLISIS, DISEÑO Y DESARROLLO

4.1. Análisis de requerimientos del proceso de desarrollo

Los requerimientos del sistema se clasifican en funcionales y no funcionales. Los requerimientos funcionales son los que están ligados a las necesidades del usuario y los no funcionales son los que están vinculados a las restricciones del sistema como capacidad de almacenamiento, tiempo de respuesta, requisitos de hardware y software, entre otros.

4.1.1. Requerimientos funcionales

Los requerimientos funcionales fueron levantados mediante la técnica de la entrevista a los administradores del curso MOOC HCD.

Durante la entrevista se identificó la necesidad de redefinir la escala de conocimiento que manejaba el administrador para analizar la trayectoria del estudiante antes, durante y después del curso MOOC HCD, debido a que, estaba mal planteada y no abarcaba adecuadamente el campo de análisis requerido para las calificaciones, según se visualizó en el análisis de evaluaciones realizadas en cursos MOOC anteriores. La escala propuesta se la ilustra en la Tabla 12.

Tabla 12: Nivel de Conocimiento

Escala	Nivel de Conocimiento
0 - 29	Inicial
30 - 59	Regular
60 - 75	En Desarrollo
76 - 89	Desarrollado
90 - 100	Excelente

Se definió otra escala de conocimiento para el detalle por tema de los exámenes prácticos (PPT, Word y Excel), como se visualiza en la Tabla 13, Tabla 14 y Tabla 15 respectivamente.

Tabla 13: Rúbrica examen PowerPoint

Examen PPT			
Tema	Subtema	Escala	N.C. por subtema
Distribución de contenido	Una idea principal por diapositiva	0 - 3	Inicial
		4 - 7	En Desarrollo
		8 - 10	Excelente

Examen PPT				
Tema	Subtema	Escala	N.C. por subtema	
0 – 9 Inicial 10 – 21 En Desarrollo 22 – 30 Excelente	Títulos adecuados	0 - 3	Inicial	
		4 - 7	En Desarrollo	
		8 - 10	Excelente	
	Diapositivas no sobrecargadas	0 - 3	Inicial	
		4 - 7	En Desarrollo	
		8 - 10	Excelente	
Representación adecuada de la información con (viñetas, tablas o smartArt)	Inserta viñetas, tablas o SmartArt al momento de presentar información	0 - 1	Inicial	
		2 - 4	En Desarrollo	
		5 - 6	Excelente	
	0 – 3 Inicial 4 – 7 En Desarrollo 8 – 10 Excelente	Buenas prácticas en el uso de formatos	0 - 1	Inicial
			2	En Desarrollo
			3 - 4	Excelente
Número de diapositivas	Número de diapositivas	0 - 1	Inicial	
		2 - 3	En Desarrollo	
		4 - 5	Excelente	
Imágenes	Hace uso de todas las imágenes	0 - 1	Inicial	
		2 - 3	En Desarrollo	
		4 - 5	Excelente	
	0 – 3 Inicial 4 – 7 En Desarrollo 8 – 10 Excelente	Buenas prácticas en el uso de formatos	0 - 1	Inicial
			2 - 3	En Desarrollo
			4 - 5	Excelente
Portada	Tipo de fuente	0	Inicial	
		1	En Desarrollo	

Examen PPT			
Tema	Subtema	Escala	N.C. por subtema
0 – 3 Inicial 4 – 7 En Desarrollo 8 – 10 Excelente	Tamaño	2	Excelente
		0	Inicial
		1	En Desarrollo
		2	Excelente
	Formato de fuentes (n/k/s)	0	Inicial
		1	Excelente
		2	Excelente
	Color de relleno de texto	0	Inicial
		1	En Desarrollo
		2	Excelente
	Contorno y efecto de texto	0	Inicial
		1 - 2	En Desarrollo
3		Excelente	
Título 0 – 4 Inicial 5 – 10 En Desarrollo 11 – 15 Excelente	Tipo de fuente	0	Inicial
		1 - 2	En Desarrollo
		3	Excelente
	Tamaño	0	Inicial
		1 - 2	En Desarrollo
		3	Excelente
	Formato de fuentes (n/k/s)	0	Inicial
		1	Excelente
	Color de relleno de texto	0 - 1	Inicial
		2	En Desarrollo
		3- 4	Excelente
	Contorno y efecto de texto	0 - 1	Inicial
		2	En Desarrollo
		3 - 4	Excelente
	Texto	Tipo de fuente	0 - 1
2			En Desarrollo

Examen PPT			
Tema	Subtema	Escala	N.C. por subtema
0 – 3 Inicial		3 - 4	Excelente
		0	Inicial
4 – 7 En Desarrollo	Tamaño/Color	1 - 2	En Desarrollo
		3	Excelente
8 – 10 Excelente	Interlineado/ alineación	0	Inicial
		1 - 2	En Desarrollo
		3	Excelente
Fondo 0 – 3 Inicial	Trama	0 - 1	Inicial
		2 - 3	En Desarrollo
		4 - 5	Excelente
4 – 7 En Desarrollo	Color	0 - 1	Inicial
		2 - 3	En Desarrollo
		4 - 5	Excelente
8 – 10 Excelente			Excelente

Tabla 14: Rúbrica examen Word

Examen WORD			
Tema	Subtema	Escala	N.C. por subtema
Tabla de contenido 0 – 7 Inicial	Genera Tabla de Contenido	0 - 1	Inicial
		2 - 3	En Desarrollo
		4 - 5	Excelente
8 – 17 En Desarrollo	Aplica Estilos	0 - 3	Inicial
		4 - 7	En Desarrollo
		8 - 10	Excelente
18 – 25 Excelente	Asocia Multinivel	0 - 3	Inicial
		4 - 7	En Desarrollo
		8 - 10	Excelente
	Interlineado	0 - 1	Inicial

Examen WORD				
Tema	Subtema	Escala	N.C. por subtema	
Formato del documento		2 - 3	En Desarrollo	
		4 - 5	Excelente	
	Espaciado	0 - 1	Inicial	
		2 - 3	En Desarrollo	
		4 - 5	Excelente	
	0 – 6 Inicial	Fuente	0 - 1	Inicial
			2	En Desarrollo
			3 - 4	Excelente
	7 – 14 En Desarrollo	Tamaño	0	Inicial
			1 - 2	En Desarrollo
	15 – 20 Excelente		3	Excelente
		Color	0	Inicial
			1 - 2	En Desarrollo
			3	Excelente
	Bordes de párrafo	Color	0 - 1	Inicial
2 - 3			En Desarrollo	
4 - 5			Excelente	
0 – 3 Inicial		Contorno	0 - 1	Inicial
			2	En Desarrollo
			3 - 4	Excelente
4 – 7 En Desarrollo		Grosor	0	Inicial
			1 - 2	En Desarrollo
8 – 10 Excelente			3	Excelente
Nota al Pie	Nota al Pie	0 - 1	Inicial	
		2 - 3	En Desarrollo	
		4 - 5	Excelente	
Pie de Página	Pie de Página	0 - 2	Inicial	

Examen WORD			
Tema	Subtema	Escala	N.C. por subtema
		3 - 5	En Desarrollo
		6 - 8	Excelente
Letra Capital	Letra Capital	0 - 2	Inicial
		3 - 4	En Desarrollo
		5 - 7	Excelente
Columnas	Columnas	0 - 3	Inicial
		4 - 7	En Desarrollo
		8 - 10	Excelente
Saltos de Página	Saltos de Página	0 - 1	Inicial
		2 - 3	En Desarrollo
		4 - 5	Excelente
Viñetas	Viñetas	0 - 3	Inicial
		4 - 7	En Desarrollo
		8 - 10	Excelente

Tabla 15: Rúbrica examen Excel

Examen Excel			
Tema/	Preguntas	Puntos	Nivel de Conocimiento
Funciones Básicas	Pregunta 1	4	4 Inicial
	Pregunta 2	4	
	Pregunta 3	4	12 En Desarrollo
	Pregunta 4	4	
	Pregunta 5	4	
Ordenar (1 Nivel)	Pregunta 6	0	Inicial
		9	Excelente
Ordenar (2 Nivel)	Pregunta 7	0	Inicial
		11	Excelente

Examen Excel			
Tema/	Preguntas	Puntos	Nivel de Conocimiento
Filtrar (1 Nivel)	Pregunta 8	0	Inicial
		9	Excelente
Filtrar (2 Nivel)	Pregunta 9	0	Inicial
		11	Excelente
Tabla Dinámica	Pregunta 10	0	5 Inicial 12 En Desarrollo 25 Excelente
		5	
	Pregunta 11	0	
		6	
	Pregunta 12	0	
		6	
	Pregunta 13	0	
		8	
Función SI (sencilla)	Pregunta 14	0	Inicial
		5	Excelente
Función SI (combinada)	Pregunta 15	0	Inicial
		10	Excelente

Durante el levantamiento de información se revisaron los archivos fuente del sistema académico, edX y Sidweb, y se pudo identificar la necesidad de información detallada en los siguientes reportes.

Tabla 16: Detalle de Reportes MOOC HCD

N.	Reportes	Detalle
1	Reporte detallado de examen teórico	<ul style="list-style-type: none"> • Datos demográficos • Nota obtenida por pregunta (50 preguntas) • Promedio obtenido por el grupo de preguntas de cada módulo (6 módulos) • Número de preguntas evaluadas, sin responder, correctas e incorrectas por módulo • Nivel de conocimiento por módulo • Calificación Total
2	Reporte detallado del curso MOOC edX	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Promedio obtenido por módulo (6 módulos) • Nivel de conocimiento por módulo • Promedio General
3	Reporte detallado de la encuesta demográfica	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Conteo de respuestas obtenidas por cada opción de cada pregunta • Número total de estudiantes registrados. • Número total de estudiantes que realizaron la prueba. • Número total de estudiantes que no realizaron la prueba.
4	Reporte detallado de la prueba de diagnóstico	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Conteo de respuestas obtenidas por cada opción de cada módulo • Número total de estudiantes registrados.

N.	Reportes	Detalle
		<ul style="list-style-type: none"> • Número total de estudiantes que realizaron la prueba. • Número total de estudiantes que no realizaron la prueba.
5	Reporte detallado de la encuesta Satisfacción	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Conteo de respuestas obtenidas por cada opción de cada pregunta • Número total de estudiantes registrados. • Número total de estudiantes que realizaron la prueba. • Número total de estudiantes que no realizaron la prueba.
6	Reporte detallado de las calificaciones examen práctico Word	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Franja, Paralelo, Nombre del docente • Nota obtenida por tema • Nivel de conocimiento por tema • Calificación Total • Nivel de conocimiento por calificación total
7	Reporte detallado de las calificaciones examen práctico PPT	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Franja, Paralelo, Nombre del docente • Nota obtenida por tema • Nivel de conocimiento por tema • Calificación Total • Nivel de conocimiento por calificación total

N.	Reportes	Detalle
8	Reporte detallado de las calificaciones examen práctico Excel	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Franja, Paralelo, Nombre del docente • Nota obtenida por tema • Nivel de conocimiento por tema • Calificación Total • Nivel de conocimiento por calificación total
9	Reporte de las calificaciones del examen práctico con el incremento de los puntos de la encuesta	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Franja, Paralelo, Nombre del docente • Calificación Total Word • Calificación Total PPT • Calificación Total Excel • Calificación Total Word con incremento • Calificación Total PPT con incremento • Calificación Total Excel con incremento
10	Reporte consolidado de calificaciones	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Franja, Paralelo, Fecha inicio curso MOOC • Calificación Examen Teórico • Calificación Total Word con incremento • Calificación Total PPT con incremento • Calificación Total Excel con incremento

N.	Reportes	Detalle
11	Reporte nivel de conocimiento detallado	<ul style="list-style-type: none"> • Datos demográficos • Tipo de estudiante: Admisiones o ESPOL • Reporte Encuesta Prueba de Diagnóstico. Campos: Niveles de conocimiento del MOD 1 al 6 • Reporte MOOC HCD edX. Campos: Niveles de conocimiento del MOD 1 al 6 • Reporte Examen Teórico. Campos: Niveles de conocimiento del MOD 1 al 6 • Reporte examen PPT. Campo: Nivel de conocimiento por calificación total • Reporte examen Word. Campo: Nivel de conocimiento por calificación total • Reporte examen Excel. Campo: Nivel de conocimiento por calificación total

4.1.2. Requerimientos no funcionales

Los requerimientos no funcionales son los que están enfocados en el rendimiento del sistema, tales como:

Diseño: Debe tener una interfaz amigable y un menú que agrupe la información de acuerdo con el contenido de cada reporte, de tal forma que facilite la navegación del usuario.

La gama de colores a usar estará vinculada al componente y categoría que se ilustre en las gráficas, como se define en la

Tabla 18 y Tabla 17.

Tabla 17: Gama de colores por categoría

Tipo de estudiante		Género	
ESPOL	Admisiones	Hombre	Mujer
#7A6248	#AC9597	#34b5aa	#e94f53

Tabla 18: Gama de colores por componente

Módulo	Componentes MOOC HCD			
	Diagnóstico	MOOC	Ex. Teórico	Ex. Práctico
Computador	#FFD966	#B08600	#7F6000	N.A.
Internet	#47CFFF	#00A1DA	#00729A	N.A.
PowerPoint	#F8CBAD	#F4B183	#C55A11	#843C0C
Word	#BDD7EE	#9DC3E6	#2E75B6	#1F4E79
Excel	#97E597	#4AD24A	#009900	#006600
Herramientas Colaborativas	#C7A0E4	#9A57CD	#7030A0	N.A.

Rendimiento: La generación de reportes y gráficas debe realizarse en lapsos de tiempos cortos imperceptibles al usuario, para ello, la distribución detallada de los datos en tablas y los queries de las consultas, deben estar segmentados en relación con los requerimientos para un mejor aprovechamiento de los recursos.

Seguridad: El sistema tendrá una página de inicio de sesión para que los administradores puedan autenticarse y brindar seguridad a los datos.

Requerimiento de Hardware: Para el desarrollo del sistema se requiere una computadora de alto rendimiento para el procesamiento de los datos con procesador i7, 1 tera de SSD, 500 GB de disco duro y 32GB RAM.

Requerimiento de Software: Se requieren las aplicaciones Visual Studio Code V1.58, MySQL Workbench 8.0.25, Xampp V3.3.0, Git con GitHub 2.9.0, HighCharts, Bootstrap, HTML5 y CSS3, Lenguaje PHP, APP Postman.

4.2. Diseño del modelo de datos

Debido a la gran cantidad de datos que genera el curso MOOC HCD cada semestre, la administración del curso busca integrar los datos distribuidos en tres sistemas para crear información, que le permita la toma de decisiones en torno a mejoras y distribución de contenido.

4.2.1. Definición de la arquitectura

Para una mejor comprensión de la aplicación, la estructura del sistema se aprecia en la Figura 4.1.

Sistema de fuente de datos: Contienen los datos demográficos y calificaciones de los estudiantes que cursan el MOOC HCD.

Proceso ETL: Consiste en realizar los procesos de extracción, transformación y carga de los archivos alojados en el repositorio OneDrive a la base de datos integrada.

Base de datos integrada: Contiene las entidades en general y consultas para la generación de los reportes integrados.

Interfaz de usuario: Aplicación donde los administradores podrán visualizar y descargar la información de los estudiantes.

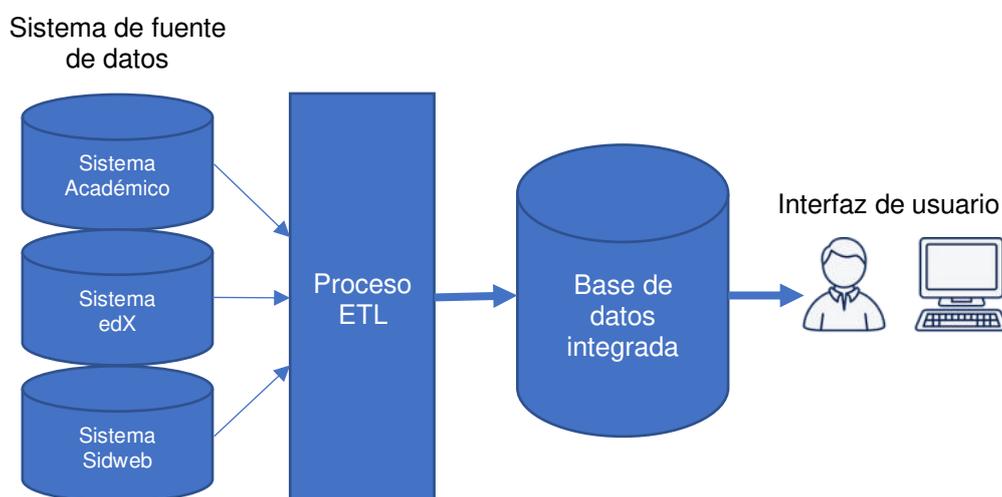


Figura 4.1: Estructura del Sistema

4.2.2. Modelo de datos fuente

Los archivos fuentes como se detalla en la Tabla 19, son extraídos de los sistemas y colocados en el repositorio central por el administrador. Este proceso es fundamental para el cumplimiento del proceso ETL y posterior integración de los datos.

Tabla 19: Fuente de datos

Sistema	Componente	Nombre de Archivo	Detalle
Académico		Listado_Admisiones_Practico Listado_Espol_Practico Listado_Admisiones_Teorico_Practico Listado_Espol_Teorico_Practico	Contienen los datos demográficos de los estudiantes. En el caso de la lista de estudiantes prácticos, adicional contienen las notas obtenidas en el curso anterior.
edX	Encuesta Demográfica	Pregunta 1. FIEC_HCDEA03_2020-2S-P Pregunta 2. FIEC_HCDEA03_2020-2S-P Pregunta 3. FIEC_HCDEA03_2020-2S-P Pregunta 4. FIEC_HCDEA03_2020-2S-P Pregunta 5. FIEC_HCDEA03_2020-2S-P Pregunta 6. FIEC_HCDEA03_2020-2S-P Pregunta 7. FIEC_HCDEA03_2020-2S-P	Reportes de respuestas por cada pregunta del formulario
	Encuesta Diagnóstico	FIEC_HCDEA03_2020-2S-P_problem_grade_report	Reporte detallado por problema del curso MOOC
	Encuesta Satisfacción	Pregunta 1. FIEC_HCDEA03_2020-2S-P Pregunta 2. FIEC_HCDEA03_2020-2S-P Pregunta 3. FIEC_HCDEA03_2020-2S-P Pregunta 4. FIEC_HCDEA03_2020-2S-P Pregunta 5. FIEC_HCDEA03_2020-2S-P Pregunta 6. FIEC_HCDEA03_2020-2S-P	Reportes de respuestas por cada pregunta del formulario
	Notas MOOC	FIEC_HCDEA03_2020-2S-P_grade_report	Reporte general del curso MOOC
Sidweb	Franjas N	Evaluaciones Examen Práctico Excel	Reporte general por paralelo con los resultados de la evaluación de Excel (15 preguntas)
Usuario Administrador	Encuestas	Códigos_encuestas_demográficas Códigos_encuestas_satisfacción	Extracción de los códigos de las opciones de respuestas de cada encuesta para la correcta interpretación de los resultados
	Franjas N	Rúbricas Examen Práctico Word Rúbricas Examen Práctico PPT	Rúbricas calificadas por los profesores asignados

4.3. Proceso ETL

El proceso ETL es ejecutado por un API, el cual se ilustra en la Figura 4.2, este contiene varios archivos .php que permiten completar cada fase del proceso. El tiempo de ejecución del proceso ETL para toda la data es de 80 minutos y 53 segundos, este resultado fue extraído de la herramienta Postman.

```

carga_data_academico_practico.php
carga_data_academico_teorico_practico.php
carga_data_ec_demografica_practico.php
carga_data_ec_demografica_respuestas.php
carga_data_ec_demografica_teorico_practico.php
carga_data_ec_diagnostico_practico.php
carga_data_ec_diagnostico_teorico_practico.php
carga_data_ec_satisfaccion_p_respuestas.php
carga_data_ec_satisfaccion_practico.php
carga_data_ec_satisfaccion_teorico_practico.php
carga_data_ec_satisfaccion_tp_respuestas.php
carga_data_edx_practico.php
carga_data_edx_teorico_practico.php
carga_data_examenTeorico_computador.php
carga_data_examenTeorico_excel.php
carga_data_examenTeorico_herramientas.php
carga_data_examenTeorico_internet.php
carga_data_examenTeorico_ppt.php
carga_data_examenTeorico_word.php
carga_data_puntos_extras_practico.php
carga_data_puntos_extras_teorico_practico.php
carga_data_rubrica_excel_practico.php
carga_data_rubrica_excel_teorico_practico.php
carga_data_rubrica_ppt_practico.php
carga_data_rubrica_ppt_teorico_practico.php
carga_data_rubrica_word_practico.php
carga_data_rubrica_word_teorico_practico.php
convierte_archivos_csv_ruta.php
procesa_archivos_ruta.php

```

Figura 4.2: Proceso de carga de datos

En el proceso de extracción, se descargan los archivos del repositorio OneDrive, almacenados en tres carpetas principales: Sistema Académico, EdX y Sidweb. Los primeros archivos por extraer son de la carpeta Sistema Académico como se ilustra en la Figura 4.3, en

donde constan las listas de los estudiantes MOOC completo y MOOC repetidores.

```

<?php
include("lib/funciones.php");
require_once "../src/SimpleXML.php";
ini_set('MAX_EXECUTION_TIME', 0);
set_time_limit(0);

try
{
    $conexion = Conectarse();
    $query = "SELECT idRuta, ruta, termino, sistema, tipo,
        ifnull(tipoEstudiante,'') as tipoEstudiante,
        ifnull(tipoDetalle,'') as tipoDetalle,
        ifnull(procesado,0) as procesado,
        ifnull(convertido,0) as convertido
        FROM rutasarchivosdata where procesado = 0;";
    $result = mysqli_query($conexion, $query);

    if (mysqli_num_rows($result) > 0)
    {
        while($fila = mysqli_fetch_assoc($result))
        {
            $idRuta      = $fila["idRuta"];
            $ruta        = $fila["ruta"];
            $termino     = $fila["termino"];
            $anio        = substr($termino,3,4);
            $sistema     = $fila["sistema"];
            $tipo        = $fila["tipo"];
            $tipoDetalle = $fila["tipoDetalle"];
            $tipoEstudiante = $fila["tipoEstudiante"];
            $datos       = listarArchivos($ruta);

            foreach ($datos as $data)
            {
                ////////////////////////////////////////////////////////////////////academico
                if ($sistema == 'ACA')
                {
                    if ($tipo == 'P')
                    {
                        $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_academico_practico.php";
                        eliminarDatosAcademicoPractico($termino,$anio,$sistema,$tipoEstudiante);
                        postdataAcademicoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);
                    }

                    if ($tipo == 'TP')
                    {
                        $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_academico_teorico_practico.php";
                        eliminarDatosAcademicoTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante);
                        postdataAcademicoTeoricoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);
                    }
                }
            }
        }
    }
}

```

Figura 4.3: Proceso de extracción Sistema Académico

La siguiente carpeta del repositorio es EdX, en donde primero se descargan los archivos de las notas de los cursos MOOC completo y MOOC repetidores, como se muestra en la Figura 4.4.

```

//////////////////////////////////////////////////////////////////edx
if ($sistema == 'EDX')
{
    ////////////////////////////////////////////////////////////////////datos mooc
    if ($tipoDetalle == 'M' && $tipo == 'P')
    {
        $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_edx_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosEdxPractico($termino,$anio,$sistema);}
        if ($tipoEstudiante != '') {eliminarDatosEdxPractico($termino,$anio,$sistema,$tipoEstudiante);}
        postdataEdxPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);
    }

    if ($tipoDetalle == 'M' && $tipo == 'TP')
    {
        $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_edx_teorico_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosEdxTeoricoPractico($termino,$anio,$sistema);}
        if ($tipoEstudiante != '') {eliminarDatosEdxTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante);}
        postdataEdxTeoricoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);
    }
}

```

Figura 4.4: Proceso de extracción notas MOOC

Los siguientes archivos para extraer son las encuestas: demográfica, satisfacción y diagnóstico de los estudiantes MOOC completo y MOOC repetidores, como se ilustra en la Figura 4.5, Figura 4.6, Figura 4.7 y Figura 4.8 respectivamente.

```

////////////////////////////////////encuesta demografica
if ($tipoDetalle == 'D' )
{
    $pregunta = substr($data,1,1);
    $respuesta = substr($data,0,1);

    if ($tipo == 'P')
    {
        ////////////////////////////////////PARA INSERTAR SOLO PREGUNTAS
        if ($respuesta != 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_demografica_practico.php";
            if ($tipoEstudiante == '') {eliminarDatosEncuestaDemograficoPractico($pregunta,$termino,$anio,$sistema);}
            if ($tipoEstudiante != '') {eliminarDatosEncuestaDemograficoPractico($pregunta,$termino,$anio,$sistema,$tipoEstudiante);}
            postDataEncuestaDemograficaPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$pregunta);
        }

        ////////////////////////////////////PARA INSERTAR SOLO RESPUESTAS
        if ($respuesta == 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_demografica_respuestas.php";
            if ($tipoEstudiante == '') {eliminarRespuestasEncuestaDemografico($termino,$anio,$sistema,$tipo);}
            if ($tipoEstudiante != '') {eliminarRespuestasEncuestaDemografico($termino,$anio,$sistema,$tipo,$tipoEstudiante);}
            postRespuestasEncuestaDemografica($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$tipo);
        }
    }

    if ($tipo == 'TP')
    {
        ////////////////////////////////////PARA INSERTAR SOLO PREGUNTAS
        if ($respuesta != 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_demografica_teorico_practico.php";
            if ($tipoEstudiante == '') {eliminarDatosEncuestaDemograficoTeoricoPractico($pregunta,$termino,$anio,$sistema);}
            if ($tipoEstudiante != '') {eliminarDatosEncuestaDemograficoTeoricoPractico($pregunta,$termino,$anio,$sistema,$tipoEstudiante);}
            postDataEncuestaDemograficaTeoricoPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$pregunta);
        }

        ////////////////////////////////////PARA INSERTAR SOLO RESPUESTAS
        if ($respuesta == 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_demografica_respuestas.php";
            if ($tipoEstudiante == '') {eliminarRespuestasEncuestaDemografico($termino,$anio,$sistema,$tipo);}
            if ($tipoEstudiante != '') {eliminarRespuestasEncuestaDemografico($termino,$anio,$sistema,$tipo,$tipoEstudiante);}
            postRespuestasEncuestaDemografica($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$tipo);
        }
    }
}
}

```

Figura 4.5: Proceso de extracción encuesta demográfica

```

////////////////////////////////////encuesta satisfaccion
if ($tipoDetalle == 'S' )
{
    $pregunta = substr($data,1,1);
    $respuesta = substr($data,0,1);

    if ($tipo == 'P')
    {
        ////////////////////////////////////PARA INSERTAR SOLO PREGUNTAS
        if ($respuesta != 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_satisfaccion_practico.php";
            if ($tipoEstudiante == '') {eliminarDatosEncuestaSatisfaccionPracticoSte($pregunta,$termino,$anio,$sistema);}
            if ($tipoEstudiante != '') {eliminarDatosEncuestaSatisfaccionPractico($pregunta,$termino,$anio,$sistema,$tipoEstudiante);}
            postDataEncuestaSatisfaccionPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$pregunta);
        }

        ////////////////////////////////////PARA INSERTAR SOLO RESPUESTAS
        if ($respuesta == 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_satisfaccion_r_respuestas.php";
            if ($tipoEstudiante == '') {eliminarRespuestasEncuestaSatisfaccionSte($termino,$anio,$sistema,$tipo);}
            if ($tipoEstudiante != '') {eliminarRespuestasEncuestaSatisfaccion($termino,$anio,$sistema,$tipo,$tipoEstudiante);}
            postRespuestasEncuestaSatisfaccionP($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$tipo);
        }
    }

    if ($tipo == 'TP')
    {
        ////////////////////////////////////PARA INSERTAR SOLO PREGUNTAS
        if ($respuesta != 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_satisfaccion_teorico_practico.php";
            if ($tipoEstudiante == '') {eliminarDatosEncuestaSatisfaccionTeoricoPracticoSte($pregunta,$termino,$anio,$sistema);}
            if ($tipoEstudiante != '') {eliminarDatosEncuestaSatisfaccionTeoricoPractico($pregunta,$termino,$anio,$sistema,$tipoEstudiante);}
            postDataEncuestaSatisfaccionTeoricoPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$pregunta);
        }

        ////////////////////////////////////PARA INSERTAR SOLO RESPUESTAS
        if ($respuesta == 'r')
        {
            $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_satisfaccion_tp_respuestas.php";
            if ($tipoEstudiante == '') {eliminarRespuestasEncuestaSatisfaccionSte($termino,$anio,$sistema,$tipo);}
            if ($tipoEstudiante != '') {eliminarRespuestasEncuestaSatisfaccion($termino,$anio,$sistema,$tipo,$tipoEstudiante);}
            postRespuestasEncuestaSatisfaccionTP($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$tipo);
        }
    }
}

```

Figura 4.6: Proceso de extracción encuesta de satisfacción

```

////////////////////////////////////encuesta diagnostico practico
if ($tipoDetalle == '0' && $tipo == 'P')
{
    if ($tipoEstudiante == '')
    {
        eliminarDatosEncuestaDiagnosticoPracticoSte($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoPPTPracticoSte($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoExcelPracticoSte($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoWordPracticoSte($termino,$anio,$sistema);
    }
    if ($tipoEstudiante != '')
    {
        eliminarDatosEncuestaDiagnosticoPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoPPTPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoExcelPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoWordPractico($termino,$anio,$sistema,$tipoEstudiante);
    }

    $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_ec_diagnostico_practico.php";
    postDataEncuestaDiagnosticoPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante);

    $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_examenTeorico ppt practico.php";
    postExamenTeoricoPPT_Practico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante);

    $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_examenTeorico excel practico.php";
    postExamenTeoricoExcel_Practico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante);

    $url = "http://localhost:8080/sistemaprocesanotasmooc/cargaData/carga_data_examenTeorico word practico.php";
    postExamenTeoricoWord_Practico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante);
}

```

Figura 4.7: Proceso de extracción encuesta diagnóstico estudiantes MOOC repetidores

```

////////////////////////////////////encuesta diagnostico teorico practico
if ($tipoDetalle == '0' && $tipo == 'TP')
{
    if ($tipoEstudiante == '')
    {
        eliminarDatosEncuestaDiagnosticoTeoricoPractico($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoComputador($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoInternet($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoPPTTeoricoPractico($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoExcelTeoricoPractico($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoWordTeoricoPractico($termino,$anio,$sistema);
        eliminarDatosExamenTeoricoHerramientas($termino,$anio,$sistema);
    }
    if ($tipoEstudiante != '')
    {
        eliminarDatosEncuestaDiagnosticoTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoComputador($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoInternet($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoPPTTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoExcelTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoWordTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante);
        eliminarDatosExamenTeoricoHerramientas($termino,$anio,$sistema,$tipoEstudiante);
    }
}

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_encuesta_diagnostico_teorico_practico.php";
postDataEncuestaDiagnosticoTeoricoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_examenTeorico_computador.php";
postExamenTeoricoComputador($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_examenTeorico_internet.php";
postExamenTeoricoInternet($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_examenTeorico_ppt_teorico_practico.php";
postExamenTeoricoPPT_TeoricoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_examenTeorico_excel_teorico_practico.php";
postExamenTeoricoExcel_TeoricoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_examenTeorico_word_teorico_practico.php";
postExamenTeoricoWord_TeoricoPractico($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

$url = "http://localhost:8080/sistemasprocesanotasmooc/cargaData/carga_data_examenTeorico_herramientas.php";
postExamenTeoricoHerramientas($url, $ruta, $data, $termino, $sistema, $tipoEstudiante);

```

Figura 4.8: Proceso de extracción encuesta diagnóstico estudiantes MOOC completo

Finalmente, los archivos a extraer son los de la carpeta Sidweb, donde constan las rúbricas de los exámenes prácticos Excel, Word y PPT como se muestra en la Figura 4.9.

```

//////////////////////////////////////sidweb
if ($sistema == 'SID')
{
    $franja = substr($data,3,1);
    $paralelo = substr($data,6,1);
    $docente = str_replace(' ','',strtoupper(ltrim(trim(quitar_tildes(substr($data,9,-5))))));

    if (substr($data,7,1) != '_')
    {
        $paralelo = substr($data,6,2);
        $docente = str_replace(' ','',strtoupper(ltrim(trim(quitar_tildes(substr($data,9,-5))))));
    }

    ////////////////////////////////////////rubricas excel
    if ($tipo == 'P' || $tipoDetalle == 'E')
    {
        $url = "http://localhost:8080/sistemaprosesanotasmooc/cargaData/carga_data_rubrica_excel_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosRubricaExcelPracticoSte($termino,$anio,$sistema,$franja,$paralelo);}
        if ($tipoEstudiante != '') {eliminarDatosRubricaExcelPractico($termino,$anio,$sistema,$tipoEstudiante,$franja,$paralelo);}
        postdataRubricaExcelPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$franja,$paralelo,$docente);
    }

    if ($tipo == 'TP' || $tipoDetalle == 'E')
    {
        $url = "http://localhost:8080/sistemaprosesanotasmooc/cargaData/carga_data_rubrica_excel_teorico_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosRubricaExcelTeoricoPracticoSte($termino,$anio,$sistema,$franja,$paralelo);}
        if ($tipoEstudiante != '') {eliminarDatosRubricaExcelTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante,$franja,$paralelo);}
        postdataRubricaExcelTeoricoPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$franja,$paralelo,$docente);
    }

    ////////////////////////////////////////rubricas ppt
    if ($tipo == 'P' || $tipoDetalle == 'P')
    {
        $url = "http://localhost:8080/sistemaprosesanotasmooc/cargaData/carga_data_rubrica_ppt_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosRubricaPPTPracticoSte($termino,$anio,$sistema,$franja,$paralelo);}
        if ($tipoEstudiante != '') {eliminarDatosRubricaPPTPractico($termino,$anio,$sistema,$tipoEstudiante,$franja,$paralelo);}
        postdataRubricaPPTPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$franja,$paralelo,$docente);
    }

    if ($tipo == 'TP' || $tipoDetalle == 'P')
    {
        $url = "http://localhost:8080/sistemaprosesanotasmooc/cargaData/carga_data_rubrica_ppt_teorico_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosRubricaPPTTeoricoPracticoSte($termino,$anio,$sistema,$franja,$paralelo);}
        if ($tipoEstudiante != '') {eliminarDatosRubricaPPTTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante,$franja,$paralelo);}
        postdataRubricaPPTTeoricoPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$franja,$paralelo,$docente);
    }

    ////////////////////////////////////////rubricas word
    if ($tipo == 'P' || $tipoDetalle == 'W')
    {
        $url = "http://localhost:8080/sistemaprosesanotasmooc/cargaData/carga_data_rubrica_word_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosRubricaWordPracticoSte($termino,$anio,$sistema,$franja,$paralelo);}
        if ($tipoEstudiante != '') {eliminarDatosRubricaWordPractico($termino,$anio,$sistema,$tipoEstudiante,$franja,$paralelo);}
        postdataRubricaWordPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$franja,$paralelo,$docente);
    }

    if ($tipo == 'TP' || $tipoDetalle == 'W')
    {
        $url = "http://localhost:8080/sistemaprosesanotasmooc/cargaData/carga_data_rubrica_word_teorico_practico.php";
        if ($tipoEstudiante == '') {eliminarDatosRubricaWordTeoricoPracticoSte($termino,$anio,$sistema,$franja,$paralelo);}
        if ($tipoEstudiante != '') {eliminarDatosRubricaWordTeoricoPractico($termino,$anio,$sistema,$tipoEstudiante,$franja,$paralelo);}
        postdataRubricaWordTeoricoPractico($url,$ruta,$data,$termino,$sistema,$tipoEstudiante,$franja,$paralelo,$docente);
    }
}

```

Figura 4.9: Proceso extracción archivos Sidweb

Los archivos fuentes descargados de los sistemas son de tipo semiestructurados, debido a que, están en formato csv y xlsx, por lo que, mantienen diferentes estructuras. El proceso de transformación convertirá los archivos a un mismo formato, en este caso .xlsx, para

una correcta depuración de los datos, como se muestra en la Figura 4.10.

```

<?php
include("lib/funciones.php");
require '../vendor/autoload.php';
require_once '../src/SimpleXLSX.php';

ini_set('memory_limit', '-1');
set_time_limit(1500);

use PhpOffice\PhpSpreadsheet\Spreadsheet;
use PhpOffice\PhpSpreadsheet\Writer\Xlsx;

try
{
    $conexion = Conectarse();
    $query = "SELECT idRuta, ruta, termino , sistema, tipo,
                ifnull(tipoEstudiante,'') as tipoEstudiante,
                ifnull(tipoDetalle,'') as tipoDetalle,
                ifnull(procesado,0) as procesado,
                ifnull(convertido,0) as convertido
            FROM rutasarchivosdata where convertido = 0;";
    $result = mysqli_query($conexion, $query);

    if (mysqli_num_rows($result) > 0)
    {
        while($fila = mysqli_fetch_assoc($result))
        {
            $idRuta      = $fila["idRuta"];
            $ruta        = $fila["ruta"];
            $datos       = listarArchivos($ruta);

            foreach ($datos as $data)
            {
                $pos = strpos($data, '.csv');
                $path = $ruta.$data;

                if ($pos !== false)
                {
                    $spreadsheet = new Spreadsheet();
                    $reader = new \PhpOffice\PhpSpreadsheet\Reader\Csv();

                    /* Set CSV parsing options */
                    $encoding = \PhpOffice\PhpSpreadsheet\Reader\Csv::guessEncoding($path, ".csv");
                    $reader->setInputEncoding($encoding);
                    $reader->setEnclosure('');
                    $reader->setSheetIndex(0);
                    $reader->setEscapeCharacter((version_compare(PHP_VERSION, '7.4') < 0) ? "\x0" : '');

                    /* Load a CSV file and save as a XLS */
                    $newfileName = basename($path, ".csv");
                    $spreadsheet = $reader->load($path);
                    $writer = new Xlsx($spreadsheet);
                    $writer->setPreCalculateFormulas(false);
                    $writer->save($ruta.$newfileName.'.xlsx');

                    $spreadsheet->disconnectWorksheets();
                    unset($spreadsheet);
                    unlink($path);
                }

                actualizaRutaConvertido($idRuta);
            }
        }
    }
    else {
        die("Error: No hay datos en la tabla seleccionada");
    }
} catch (RuntimeException $e) {
    echo $e->getMessage();
}

```

Figura 4.10: Proceso para convertir archivo csv a xlsx

Adicionalmente, el proceso de transformación depura los datos de tal forma que se cumpla con la estructura de reporte solicitada por los administradores, considerando todas las validaciones respectivas. En la Figura 4.2 se muestra la transformación y carga para el examen Excel práctico.

```

if ($? -eq 0) continue;
for ($i = 0; $i < $cols; $i++)
{
    if ($isset($r[0]))          $name      = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[0])))); else $name = '';
    if ($isset($r[1]))        $idEstudiante = utf8_encode($r[1]); else $idEstudiante = 0;
    if ($isset($r[2]))        $sis_id     = utf8_encode($r[2]); else $sis_id = '';
    if ($isset($r[3]))        $section    = utf8_encode($r[3]); else $section = '';
    if ($isset($r[4]))        $señaliado   = utf8_encode($r[4]); else $señaliado = '';
    if ($isset($r[7]))        $intento     = utf8_encode($r[7]); else $intento = 0;

    if ($isset($r[9]))        $p1 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[9])))); else $p1 = '';
    if ($isset($r[9]) && $r[9] != '') $c1 = utf8_encode($r[9]); else $c1 = 0;

    if ($isset($r[10]))       $p2 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[10])))); else $p2 = '';
    if ($isset($r[11]) && $r[11] != '') $c2 = utf8_encode($r[11]); else $c2 = 0;

    if ($isset($r[12]))       $p3 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[12])))); else $p3 = '';
    if ($isset($r[13]) && $r[13] != '') $c3 = utf8_encode($r[13]); else $c3 = 0;

    if ($isset($r[14]))       $p4 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[14])))); else $p4 = '';
    if ($isset($r[15]) && $r[15] != '') $c4 = utf8_encode($r[15]); else $c4 = 0;

    if ($isset($r[16]))       $p5 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[16])))); else $p5 = '';
    if ($isset($r[17]) && $r[17] != '') $c5 = utf8_encode($r[17]); else $c5 = 0;

    if ($isset($r[18]))       $p6 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[18])))); else $p6 = '';
    if ($isset($r[19]) && $r[19] != '') $c6 = utf8_encode($r[19]); else $c6 = 0;

    if ($isset($r[20]))       $p7 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[20])))); else $p7 = '';
    if ($isset($r[21]) && $r[21] != '') $c7 = utf8_encode($r[21]); else $c7 = 0;

    if ($isset($r[22]))       $p8 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[22])))); else $p8 = '';
    if ($isset($r[23]) && $r[23] != '') $c8 = utf8_encode($r[23]); else $c8 = 0;

    if ($isset($r[24]))       $p9 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[24])))); else $p9 = '';
    if ($isset($r[25]) && $r[25] != '') $c9 = utf8_encode($r[25]); else $c9 = 0;

    if ($isset($r[26]))       $p10 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[26])))); else $p10 = '';
    if ($isset($r[27]) && $r[27] != '') $c10 = utf8_encode($r[27]); else $c10 = 0;

    if ($isset($r[28]))       $p11 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[28])))); else $p11 = '';
    if ($isset($r[29]) && $r[29] != '') $c11 = utf8_encode($r[29]); else $c11 = 0;

    if ($isset($r[30]))       $p12 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[30])))); else $p12 = '';
    if ($isset($r[31]) && $r[31] != '') $c12 = utf8_encode($r[31]); else $c12 = 0;

    if ($isset($r[32]))       $p13 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[32])))); else $p13 = '';
    if ($isset($r[33]) && $r[33] != '') $c13 = utf8_encode($r[33]); else $c13 = 0;

    if ($isset($r[34]))       $p14 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[34])))); else $p14 = '';
    if ($isset($r[35]) && $r[35] != '') $c14 = utf8_encode($r[35]); else $c14 = 0;

    if ($isset($r[36]))       $p15 = str_replace(' ', ' ',strtoupper(trim(trim(quitartildes($r[36])))); else $p15 = '';
    if ($isset($r[37]) && $r[37] != '') $c15 = utf8_encode($r[37]); else $c15 = 0;

    if ($isset($r[38]) && $r[38] != '') $incorrecto = utf8_encode($r[38]); else $incorrecto = 0;
    if ($isset($r[39]) && $r[39] != '') $sinincorrecto = utf8_encode($r[39]); else $sinincorrecto = 0;
    if ($isset($r[40]) && $r[40] != '') $calificacion = utf8_encode($r[40]); else $calificacion = 0;
}

$termino = $POST["termino"];
$anio = substr($termino,2,4);
$sisistema = $POST["sisistema"];
$tipoEstudiante = $POST["tipoEstudiante"];
$ejemplar = $POST["ejemplar"];
$paralelo = $POST["paralelo"];
$docente = $POST["docente"];
$archivo = $file;

if ($isset($tipoEstudiante) && $tipoEstudiante != '') {
    $adminReport = $tipoEstudiante;
} else {
    if (strlen($sis_id) < 10) $adminReport = 'RESPOL';
    if (strlen($sis_id) > 9) $adminReport = 'ADMINISTRON';
}

guardarDatosRubricaExcelPractico($name,$idEstudiante,$sis_id,$section,$señaliado,$p1,$c1,$p2,$c2,$p3,$c3,$p4,$c4,$p5,$c5,$p6,$c6,$p7,$c7,$p8,$c8,$p9,$c9,$p10,$c10,$p11,$c11,$p12,$c12,$p13,$c13,$p14,$c14,$p15,$c15,$incorrecto,$sinincorrecto,$calificacion)
}

```

Figura 4.11: Proceso de Transformación y Carga

Una vez completado el proceso ETL, se generan las tablas con la información integrada y depurada en la base de datos, como se muestra en la Figura 4.12.



Figura 4.12: Base de datos integrada MOOC HCD

En la Tabla 20, se describe de forma simplificada el proceso ETL, donde se especifica el paso a paso de cada una de las etapas antes mencionadas.

Tabla 20: Proceso ETL MOOC HCD

Sistema	Proceso ETL		
	Extracción	Transformación	Crear tablas y cargar datos
Académico	Valida que el archivo se encuentre en el repositorio	<p>Valida que no se ingrese a la base los estudiantes de colegio extranjero</p> <p>No existan datos repetidos</p> <p>Ingresar a los estudiantes teórico-práctico y práctico en tablas diferentes</p>	<p><i>/*Datos demográficos de los estudiantes teórico práctico*/</i> academicoteoricopractico.sql</p> <p><i>/*Datos demográficos y notas de los estudiantes práctico */</i> academicopractico.sql</p>
edX	Valida que el archivo se encuentre en el repositorio	<p>Valida que los usuarios unenrolled (estudiantes desvinculados) y que no tengan cohorte (docentes) no sean agregados a la base</p> <p>Reconocer los siguientes campos y asignarles un valor numérico para mejorar el rendimiento en las consultas:</p> <p><u>Contenido no habilitado</u> not available = -1</p> <p><u>Actividad no desarrollada</u> not attempted = -2</p> <p>Pregunta de encuesta no contestada none = -3</p> <p>Ingresar a los estudiantes teórico-práctico y práctico en tablas diferentes</p> <p>Las calificaciones las convierte al 100% y se trabaja con 2 decimales</p>	<p><i>/*Notas MOOC*/</i> edxteoricopractico.sql edxpractico.sql</p> <p><i>/*Examen Teórico*/</i> et_computador.sql et_internet.sql et_word.sql et_ppt.sql et_excel.sql et_herramientas.sql</p> <p><i>/*Encuestas*/</i> ec_demografica_teorico_practico.sql ec_demografica_practico.sql</p> <p>ec_diagnostico_teorico_practico.sql ec_diagnostico_practico.sql</p> <p>ec_satisfaccion_teorico_practico.sql ec_satisfaccion_practico.sql</p>
Sidweb	Valida que el archivo se encuentre en el repositorio	Valida el nombre de cada archivo para que sean parte de los campos de la tabla:	<p><i>/*Rúbricas Examen Práctico*/</i> rubricaspptteoricopractico.sql rubricaswordteoricopractico.sql rubricasexcelteoricopractico.sql</p>

Sistema	Proceso ETL		
	Extracción	Transformación	Crear tablas y cargar datos
		NotasPPT- NombreApellidoProfesor- F#-P# NotasWord- NombreApellidoProfesor- F#-P# NotasEx- NombreApellidoProfesor- F#-P# Valida el puntaje definido por tema para cada rúbrica Ingresar a los estudiantes teórico-práctico y práctico en tablas diferentes	rubricaspptpractico.sql rubricaswordpractico.sql rubricasexcelpractico.sql
Puntos Extras		Validar el cumplimiento por parte del estudiante de las 3 encuestas para otorgar el incremento de 5 puntos extras a la nota del examen práctico que más requiera para aprobar el módulo	/*Rúbricas Examen Práctico con incremento de los puntos extras*/ puntos_extras_practico.sql puntos_extras_teorico_practico.sql

4.4. Proceso de Integración de datos

En este proyecto se aplica la combinación de dos técnicas de integración, la de aplicaciones y la de datos físicos, dado que se trabaja con una copia de los archivos fuente.

Una vez depurada y almacenada la información en la base de datos integrada, se crean los queries basados en cada uno de los reportes solicitados por los administradores como se presenta en la Figura 4.13.

Luego se crea la tabla respectiva que almacenará la información vinculada a cada uno de los reportes, con ello, se estima que los tiempos de respuesta de cada consulta, sean inmediatos.

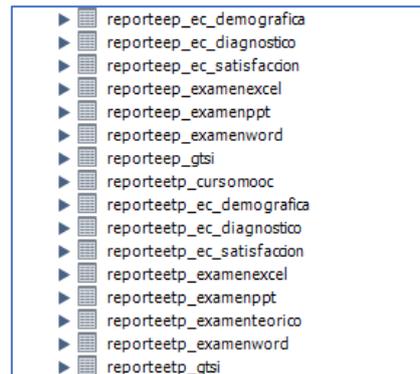


Figura 4.13: Reportes MOOC HCD

En la Figura 4.14 se plasma el paso a paso de este proceso, hasta presentar los reportes y gráficos estadísticos al usuario.



Figura 4.14: Proceso de creación de reportes integrados

4.5. Pruebas de usabilidad

Se solicitó al administrador HCD una reunión vía zoom, para que, por medio de las pruebas de usabilidad no moderadas, interactúe en la plataforma navegando por el menú, revisando los reportes y analizando las gráficas. Esto se realizó con el objetivo de percibir el

comportamiento del administrador y comprobar si el objetivo del proyecto fue cumplido. Para que la interacción se llevara a cabo se le dio permisos de anfitrión para navegar por la aplicación web.

El administrador aportó con varias observaciones con respecto a los reportes:

- Corregir el encabezado del reporte rúbrica examen Excel.
- Agregar al reporte del curso MOOC de la cohorte práctica, la columna niveles de conocimiento.
- Agregar las páginas de reporte (.html) al menú niveles de conocimiento, que en ese instante contenía solo las páginas de las gráficas estadísticas.
- En los reportes examen teórico y curso MOOC reemplazar la palabra Grade por Promedio General.

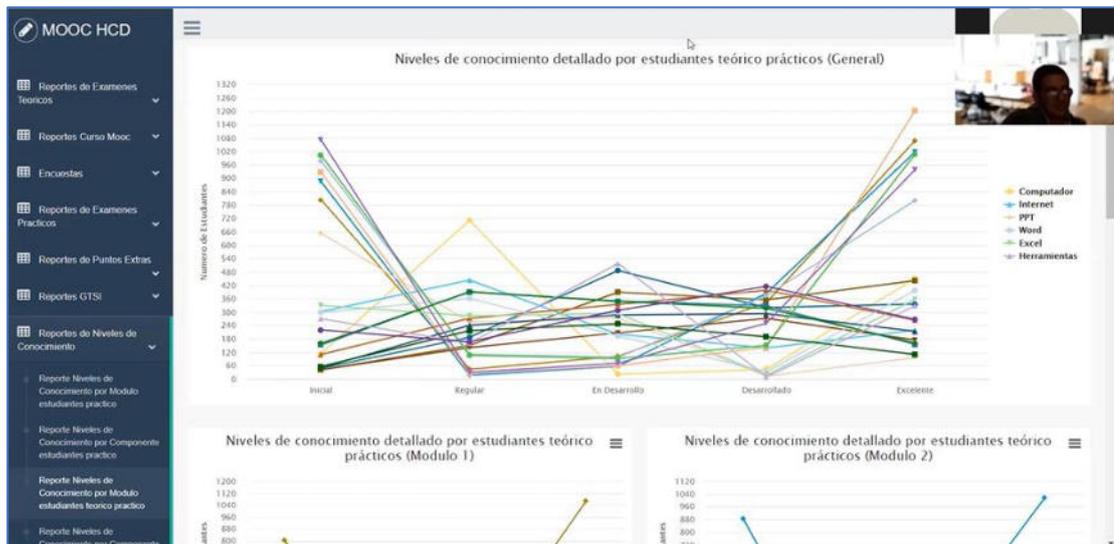


Figura 4.15: Menú MOOC HCD - Reporte Niveles de Conocimiento

ID	NOMBRES		SEXO	EMAIL	MATRICULA	CEDULA	CARRERA	GRADE	AVERIADES POR MODULO				PPT	Excel	Word
	1	2							PPT	EXCEL	WORD	PPT			
0322	PALMIRINO	RIUZ BARBERAN	MASCULINO	PALMIRINO_RIUZ@OUTLOOK.ES	0952208718	0952208718	ESTADISTICA	32	null	100	null	null	APROBADO	null	null
20308	ROY	WILSON	MASCULINO	ROYWILSON21@GMAIL.COM	0952967420	0952967420	ELECTRÓNICA Y AUTOMATIZACIÓN	79	100	84	83	APROBADO	APROBADO	APROBADO	
21523	DENIS	SAMUEL	MASCULINO	AVLESSAMUEL95@GMAIL.COM	120793648	120793648	ELECTRÓNICA Y AUTOMATIZACIÓN	0	0	null	0	REPROBADO	null	REPROBADO	
21828	ELIANA	ESTHER	FEMENINO	ELIANAEESTHER@GMAIL.COM	0932320891	0932320891	MECATRÓNICA	48	100	null	null	APROBADO	null	null	
21884	ERIKA	JAMILETH	FEMENINO	FLOREDERIKA34@GMAIL.COM	070702029	070702029	BIOLOGIA	0	0	0	0	REPROBADO	REPROBADO	REPROBADO	
21817	MACYLENE	STEPHANIE	FEMENINO	GARRIDOMACYLENESTEPHANIE@HOTMAIL.COM	0953201017	0953201017	MECÁNICA	30	null	null	100	null	null	APROBADO	
21912	OLIVER	JORDY	MASCULINO	YOLIVER09@HOTMAIL.COM	0970647034	0970647034	GEOLOGIA	83	100	52	100	APROBADO	REPROBADO	APROBADO	
22187	RODDY	MERCHAN	MASCULINO	RODDYMERCHAN@GMAIL.COM	0943321463	0943321463	COMPUTACIÓN	96	100	96	100	APROBADO	APROBADO	APROBADO	
22242	LINDA	LINDA	FEMENINO	TAPALINDA@HOTMAIL.COM	087321948	087321948	GEOLOGIA	95	100	100	100	APROBADO	APROBADO	APROBADO	
22705	JHONATAN	FLORES	MASCULINO	JHONATANFLORES2001@OUTLOOK.ES	0958402413	0958402413	ADMINISTRACIÓN DE EMPRESAS	0	0	0	0	REPROBADO	REPROBADO	REPROBADO	

Figura 4.16: Reporte Exámenes Prácticos - MOOC Repetidores

Al finalizar el ejercicio de navegación el administrador se sintió conforme con la distribución de la información, con la forma en la que se presentan las tablas y gráficos, con los colores, la interfaz amigable y con los tiempos de respuesta de cada reporte.

4.6. Planteamiento del uso de Scrum

Para el desarrollo de este proyecto se calendarizaron nueve sprint, cada uno de ellos tiene una duración de dos semanas; al cierre de cada sprint se tiene una sesión virtual de trabajo para revisar el cumplimiento y aplicación de mejoras.

A continuación, se muestra la evidencia de cuatro de los nueve sprint entregados, estos son los Sprint 1, 3, 5 y 7. Dado que son los más relevantes del proyecto, porque marcan la fase inicial del levantamiento de requerimientos, del proceso ETL, de la generación de reportes y de la ilustración de gráficos estadísticos respectivamente.

4.6.1. Sprint 1 Planificación Inicial

En base a la entrevista realizada con los administradores para el levantamiento de los requerimientos, se definieron las tareas a entregar en el sprint 1:

- Extracción de los archivos de los sistemas fuente.
- Creación del repositorio en OneDrive para el alojamiento de los archivos fuente como se ilustra en la Figura 1.2.

- Definición de reportes en base a toda la documentación recolectada como se muestra en la Tabla 7.

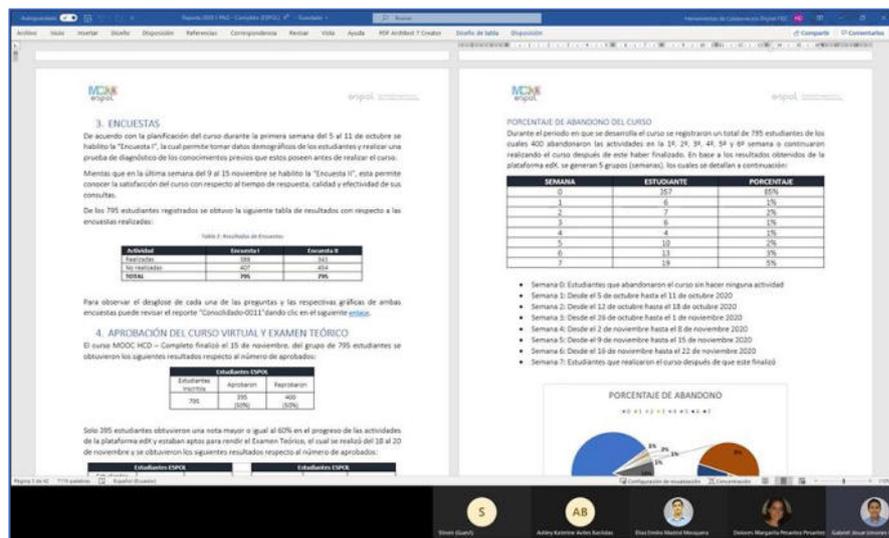


Figura 4.17: Reporte General elaborado por los administradores

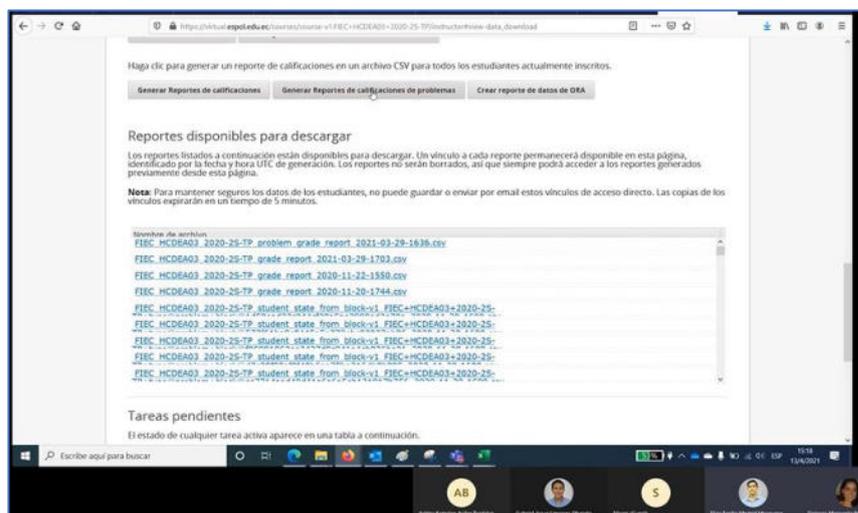


Figura 4.18: Archivos descargados plataforma edX

4.6.2. Sprint 3 – Datos integrados en la base

Una vez lleno el repositorio con los archivos fuente, las tareas asignadas para la entrega del sprint 3 son:

- Creación del proceso ETL para el tratamiento de los datos.
- Creación de la base de datos con sus respectivas entidades.

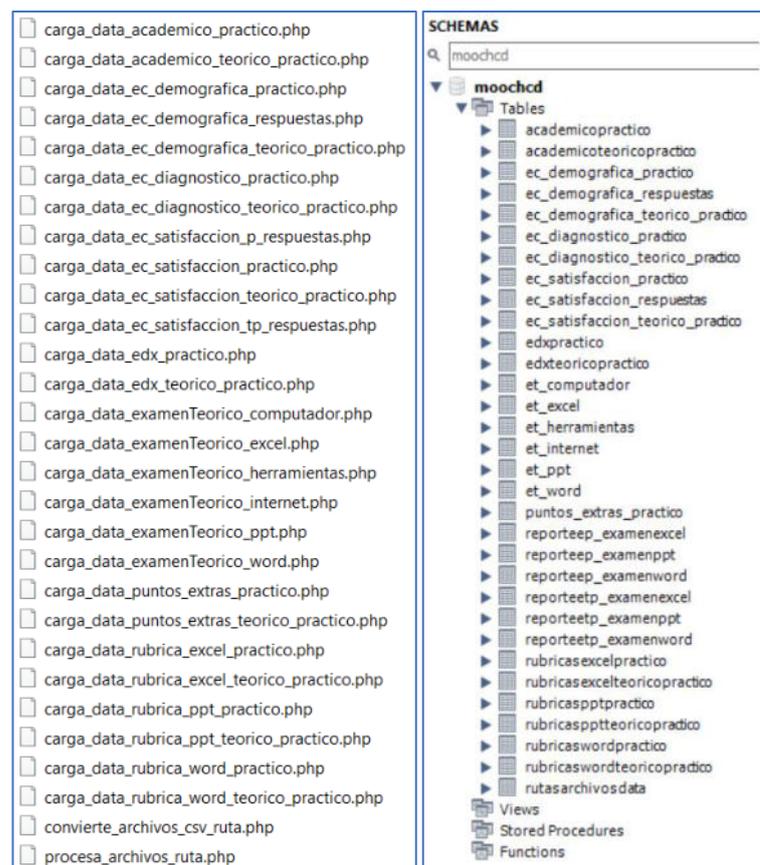


Figura 4.19: API Proceso ETL y Base de datos

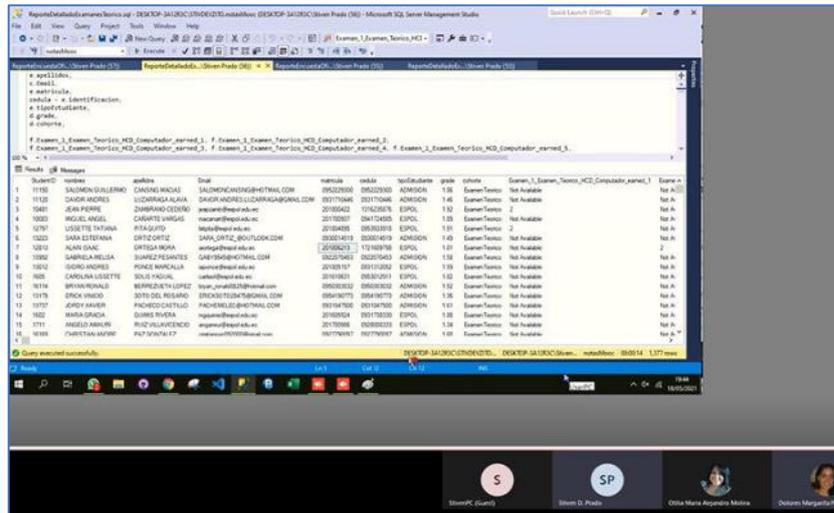


Figura 4.21: Querys MOOC HCD

TALLE DE CALIFICACIONES (ADMISIONES)

Search:

LA	¿HIZO EXAMEN?	GRADE	CARRERA	CALIFICACIONES						
				COMPUTADOR	INTERNET	PPT	EXCEL	WORD	HERRAMIENTAS	TOTAL
59328	SI	1.85	ECONOMIA	20.00	22.00	12.00	8.00	12.00	10.00	84.00
55155	SI	1.97	TURISMO	24.00	20.00	14.00	14.00	14.00	10.00	90.00
59232	SI	1.51	DISEÑO GRAFICO	22.00	24.00	10.00	8.00	8.00	0.00	66.00
10448	SI	1.45	ARQUEOLOGIA	18.00	14.00	4.00	8.00	4.00	6.00	54.00
59300	SI	1.88	ECONOMIA	22.00	22.00	12.00	14.00	10.00	10.00	90.00

Figura 4.22: Reporte examen teórico

4.6.4. Sprint 7 – Gráficas ilustrativas

Una vez completado un porcentaje de los querys, la nueva tarea asignada para el sprint 7, es iniciar las ilustraciones de los reportes. Para este sprint se entregaron 18 gráficas correspondientes a tres reportes.

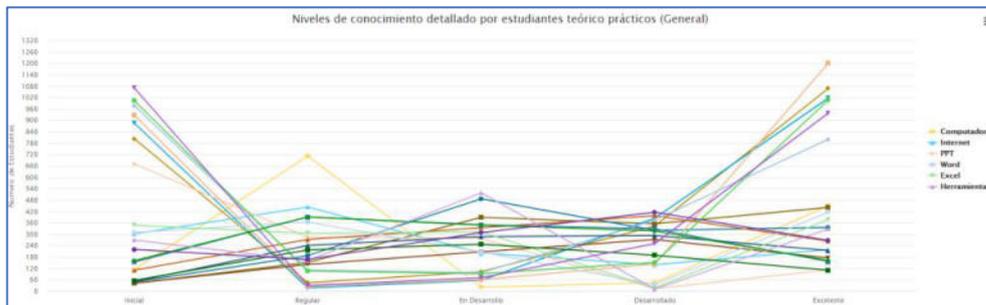


Figura 4.23: Gráfica de niveles de conocimiento estudiantes teórico-práctico



Figura 4.24: Graficas del rendimiento del examen teórico por módulo

CAPÍTULO 5

ANÁLISIS DE RESULTADOS

5.1. Implementación

En este proyecto se implementaron 21 reportes distribuidos en 7 secciones, a continuación, se detallan algunos de ellos.

Encuestas

En la Figura 5.1 se presenta el reporte de encuesta demográfica. Esta encuesta es aplicada a los estudiantes de ESPOL y Admisiones de los cursos MOOC Repetidores y MOOC Completo, donde se muestra información como: nombres, apellidos, email, matricula, cédula, carrera, sexo, edad, colegio, cohorte, ¿cuál es su ocupación?, usted como estudiante, ¿de qué forma prefiere recibir clases?, ¿en qué nivel consideras que están tus conocimientos de ofimática?, ¿por qué

medio se enteró que estaba registrado en el curso MOOC HCD?, ¿cuál es la materia que se habilita al momento de aprobar el MOOC HCD?, ¿con qué equipos tecnológicos cuenta en su hogar para realizar el curso?, ¿qué conexión de internet utiliza?.

NO	EDAD	COLEGIO	COHORTE	¿CÓMO ES SU OCUPACION?	¿CÓMO ES SU OCUPACION?	¿CÓMO ES SU OCUPACION?				
ENINO	18	UNIDAD EDUCATIVA LICEO CRISTIANO DE GUAYACUL	WORD - EXCEL	Estudio/Trabajo	Virtual	Intermedio	Cursos Electrónicos	Fundamentos de Programación	Computador/Laptop	Propio (Fijo)
ENINO	18	DE LA ASUNCION	PRACTICO	Estudio	Presencial	Intermedio	Cursos Electrónicos	Fundamentos de Programación	Computador/Laptop	Propio (Fijo)
ENINO	19	VICENTE ROCAFUERTE	PPT	Estudio	Presencial	Intermedio	Cursos Electrónicos	Fundamentos de Programación	Tablet	Prostado
ENINO	17	LICEO NIVAL	EXCEL	Estudio	Presencial	Intermedio	Cursos Electrónicos	Fundamentos de Programación	Computador/Laptop	
JULINO	18	AGUIRRE ABAD	WORD - EXCEL	Estudio	Presencial	Intermedio	Cursos Electrónicos	Fundamentos de Programación	Computador/Laptop	Propio (Fijo)

Figura 5.1: Reporte encuesta demográfica

En la Figura 5.2 se presenta el reporte de la prueba de diagnóstico. Esta encuesta es aplicada a los estudiantes de ESPOL y Admisiones del curso MOOC Completo y MOOC Repetidores.

En el caso del curso MOOC Completo se evalúa contenido básico de los seis módulos. Para el curso MOOC Repetidores se evalúa contenido de los módulos que el estudiante se encuentre cursando.

El reporte muestra información personal del estudiante, las notas obtenidas por módulo y el nivel de conocimiento en relación con los puntos obtenidos por cada módulo.

TOTAL PRUEBA DIAGNOSTICO					NIVELES DE CONOCIMIENTO					
INET	PPT	EXCEL	WORD	HERRAMIENTAS	COMPUTADOR	INTERNET	PPT	EXCEL	WORD	HERRAMIENTAS
0.00	20.00	30.00	20.00	100.00	EXCELENTE	EXCELENTE	INICIAL	REGULAR	INICIAL	EXCELENTE
1.00	30.00	100.00	100.00	20.00	EXCELENTE	DESARROLLADO	REGULAR	EXCELENTE	EXCELENTE	INICIAL
1.00	20.00	20.00	30.00	100.00	REGULAR	EN DESARROLLO	INICIAL	INICIAL	REGULAR	EXCELENTE
1.00	70.00	100.00	50.00	70.00	REGULAR	DESARROLLADO	EN DESARROLLO	EXCELENTE	REGULAR	EN DESARROLLO
0.00	20.00	100.00	50.00	20.00	EXCELENTE	EXCELENTE	INICIAL	EXCELENTE	REGULAR	INICIAL

Figura 5.2: Reporte prueba de diagnóstico

En la Figura 5.3 se presenta el reporte de la encuesta satisfacción de los estudiantes de ESPOL y Admisiones de los cursos MOOC Repetidores y MOOC Completo. En esta encuesta se mide la calidad de atención por parte de los administradores, la calidad de contenido informativo e intelectual y el tiempo de dedicación del estudiante, para ello, se muestra información personal del estudiante y se evalúa preguntas como: ¿cuál es el porcentaje de conocimiento que te aportó cada módulo?, ¿cuánto tiempo le dedicas al MOOC HCD por semana?, ¿tenías conocimiento de las fechas importantes del MOOC HCD?, ¿sabías acerca del proceso y políticas de aprobación del MOOC HCD?, ¿ingresaste a la pestaña “pregunta frecuente” del MOOC HCD?, ¿a través de qué medio de comunicación realizas con más frecuencias tus consultas o reportas tus problemas con la plataforma?, ¿La respuesta recibida por parte del equipo MOOC HCD fue clara y entendible?, ¿Su consulta o requerimiento fue solucionado eficazmente?, escoja el rango de tiempo en que sus consultas o problemas fueron atendidos.

PARTE PARA IDENTIFICAR EL MOOC	¿CUÁL ES EL PORCENTAJE DE CONOCIMIENTO QUE TE APORTÓ CADA MÓDULO?						¿A TRAVÉS DE QUE MEDIO DE COMUNICACIÓN REALIZAS MÁS FRECUENCIAS TUS CONSULTAS O REPORTAS TUS PROBLEMAS CON LA PLATAFORMA?	CLASIFIQUE LA SATISFACCIÓN RESPECTO A LA ATENCIÓN RECIBIDA		ESCOJA EL RANGO DE TIEMPO EN QUE SUS CONSULTAS O PROBLEMAS FUERON ATENDIDOS
	MÓDULO 1: EL COMPUTADOR	MÓDULO 2: INTERNET Y SUS HERRAMIENTAS BÁSICAS	MÓDULO 3: POWERPOINT	MÓDULO 4: EXCEL	MÓDULO 5: WORD	MÓDULO 6: HERRAMIENTAS COLABORATIVAS		¿LA RESPUESTA RECIBIDA POR PARTE DEL EQUIPO MOOC HCD FUE CLARA Y ENTENDIBLE?	¿SU CONSULTA O REQUERIMIENTO FUE SOLUCIONADO EFICAZMENTE?	
Si	1	1	0.75	0.75	0.75	0.75	Correo	Totalmente Satisfecho	Totalmente Satisfecho	15 a 30 minutos
No	0.5	0.5	0.5	0.5	0.5	0.5	Correo	Totalmente Satisfecho	Totalmente Satisfecho	15 a 30 minutos
No	1	1	1	1	1	1	Foro	Totalmente Satisfecho	Totalmente Satisfecho	1 a 2 horas
Si	1	0.75	1	0.75	0.5	0.5	Correo	Parcialmente Satisfecho	Parcialmente Satisfecho	1 a 2 horas

Figura 5.3: Reporte encuesta de satisfacción

Curso MOOC HCD

En la Figura 5.4 se presenta el reporte de las notas de los estudiantes de ESPOL y Admisiones de los cursos MOOC Repetidores y MOOC completo, donde se muestra información personal del estudiante, las notas obtenidas en el curso y el nivel de conocimiento en relación con los puntos obtenidos por cada módulo.

En el caso del curso MOOC Completo se muestran las notas obtenidas en los seis módulos y para el curso MOOC Repetidores muestra las notas de el o los módulos prácticos que el estudiante se encuentre cursando.

MÓDULO			NIVELES DE CONOCIMIENTO						ESTADO GENERAL	COHORTE
EXCEL	WORD	HERRAMIENTAS	COMPUTADOR	INTERNET	PPT	EXCEL	WORD	HERRAMIENTAS		
OBADO	REPROBADO	REPROBADO	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	REPROBADO	TeóricoPráctico
OBADO	REPROBADO	REPROBADO	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	REPROBADO	TeóricoPráctico
OBADO	REPROBADO	REPROBADO	EXCELENTE	EXCELENTE	INICIAL	INICIAL	INICIAL	INICIAL	REPROBADO	TeóricoPráctico
OBADO	REPROBADO	REPROBADO	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	REPROBADO	TeóricoPráctico
OBADO	REPROBADO	REPROBADO	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	INICIAL	REPROBADO	TeóricoPráctico

Figura 5.4: Reporte notas curso MOOC HCD

Examen Teórico

En la Figura 5.5 se presenta el reporte de las notas del examen teórico del curso MOOC HCD de los estudiantes de ESPOL y Admisiones del curso MOOC completo, donde se muestra información personal del estudiante y las notas obtenidas en los seis módulos, sumando un gran total sobre 100. Adicionalmente, se muestra el nivel de conocimiento en relación con los puntos obtenidos por cada módulo.

CALIFICACIONES					NIVELES DE CONOCIMIENTO					
ID	EXCEL	WORD	HERRAMIENTAS	TOTAL	COMPUTADOR	INTERNET	PPT	EXCEL	WORD	HERRAMIENTAS
00	12.00	10.00	8.00	78.00	REGULAR	DESARROLLADO	EXCELENTE	DESARROLLADO	EN DESARROLLO	DESARROLLADO
00	10.00	14.00	8.00	74.00	EN DESARROLLADO	DESARROLLADO	REGULAR	EN DESARROLLO	EXCELENTE	EN DESARROLLO
00	10.00	12.00	10.00	88.00	EN DESARROLLO	EXCELENTE	EXCELENTE	EN DESARROLLO	DESARROLLADO	EXCELENTE
00	12.00	10.00	8.00	82.00	DESARROLLADO	DESARROLLADO	DESARROLLADO	DESARROLLADO	EN DESARROLLO	DESARROLLADO
00	10.00	12.00	6.00	80.00	EXCELENTE	EN DESARROLLO	EXCELENTE	EN DESARROLLO	DESARROLLADO	EN DESARROLLO

Figura 5.5: Reporte examen teórico

De la Figura 5.6 a la Figura 5.11 se visualizan seis gráficos de barras que representan el porcentaje de aciertos y fallos de los estudiantes

de ESPOL y Admisiones en las preguntas evaluadas por cada módulo del examen teórico.

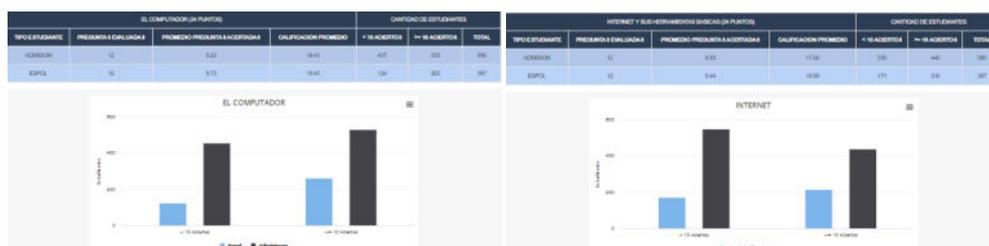


Figura 5.6: Reporte de aciertos y fallos examen teórico – módulo el computador

Figura 5.7: Reporte de aciertos y fallos examen teórico – módulo el internet



Figura 5.8: Reporte de aciertos y fallos examen teórico – módulo PowerPoint

Figura 5.9: Reporte de aciertos y fallos examen teórico – módulo Excel

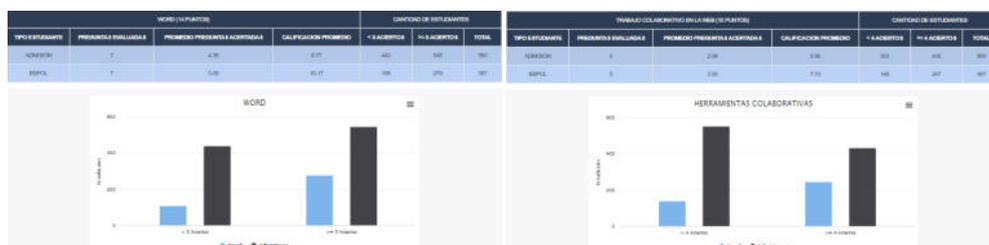


Figura 5.10: Reporte de aciertos y fallos examen teórico – módulo Word

Figura 5.11: Reporte de aciertos y fallos examen teórico – módulo herramientas colaborativas

Examen Práctico

Examen Excel

En la Figura 5.12 se presenta el reporte del examen práctico Excel de los estudiantes de ESPOL y Admisiones del curso MOOC repetidores

y MOOC completo, donde se muestra información personal del estudiante y los puntos obtenidos en cada pregunta de opciones múltiples. Adicionalmente, se muestra el nivel de conocimiento en relación con la calificación obtenida.

Nvl.	TABLA DINAMICA				FUNCION 81		TOTAL GENERAL	NIVEL CONOCIMIENTO	CARRERA	COHORTE
	P1	P2	P3	P4	SENCILLO	CONCATENADO				
5	6	0	8	19	5	10	94	EXCELENTE	DISEÑO GRÁFICO	Practico
5	0	0	8	13	5	0	72	EN DESARROLLO	INGENIERIA NAVAL	Excel
5	6	0	8	19	5	10	94	EXCELENTE	INGENIERIA CIVIL	Word-Excel
5	0	6	8	19	5	10	81	DESARROLLADO	ADMINISTRACIÓN DE EMPRESAS	Practico
5	6	0	0	11	5	10	72	EN DESARROLLO	NUTRICIÓN Y DIETETICA	Excel

Figura 5.12: Reporte rúbrica examen Excel

Examen Word

En la Figura 5.13 se presentan los reportes del examen práctico Word de los estudiantes de ESPOL y Admisiones del curso MOOC repetidores, donde se muestra información personal del estudiante y los puntos obtenidos por cada tema evaluado. Adicionalmente, se muestra el nivel de conocimiento en relación con la calificación obtenida.

BORDES			NOTA AL PIE (5)	PIE DE PAGINA (8)	LETRA CAPITAL (7)	COLUMNAS (10)	SALTOS (5)	VIÑETAS (10)	TOTAL GENERAL	NIVEL CONOCIMIENTO
OR (1)	CONTORNO (4)	GROSOR (3)								
0	0	0	0	5	7	0	0	7	32	REGULAR
1	1	1	5	5	7	10	5	10	90	EXCELENTE
3	4	3	0	5	7	7	5	10	87	DESARROLLADO
3	4	3	0	8	4	10	5	10	72	EN DESARROLLO

Figura 5.13: Reporte rúbrica examen Word

Examen PPT

En la Figura 5.14 se presenta el reporte del examen práctico PPT de los estudiantes de ESPOL y Admisiones del curso MOOC repetidores y MOOC completo, donde se muestra información personal del estudiante y los puntos obtenidos por cada tema evaluado. Adicionalmente, se muestra el nivel de conocimiento en relación con la calificación obtenida.

TITULOS (15)				TEXTO (10)			FONDO (10)		TOTAL GENERAL	NIVEL CONOCIMIENTO
WIO (1)	FORMATO DE FUENTE (1)	COLOR DE RELLENO (4)	CONTORNO Y EFECTO DE TEXTO (4)	TIPO DE FUENTE (4)	TAMANO / COLOR (3)	INTERLINEADO / ALINEACION (3)	TRAMA (5)	COLOR (5)		
3	1	4	4	0	0	3	5	5	87	DESARROLLADO
2	1	4	4	4	3	1.5	5	5	86.5	DESARROLLADO
3	1	1	4	4	3	3	5	5	88	DESARROLLADO
3	1	4	3	4	3	3	5	4	88	DESARROLLADO
3	1	4	2	4	3	3	5	5	92	EXCELENTE

Figura 5.14: Reporte rúbrica examen PPT

Puntos extras

En la Figura 5.15 se presenta el reporte de los cinco puntos extras obtenidos por realizar las tres encuestas efectuadas a los estudiantes de ESPOL y Admisiones del curso MOOC repetidores y MOOC completo, donde se muestra información personal del estudiante, las notas de los exámenes prácticos, los puntos extras y las notas de los exámenes prácticos con el incremento correspondiente.

PARTE	EXAMEN			ENCUESTA			PUNTOS EXTRAS	EXAMEN + PUNTOS EXTRAS		
	EXCEL	PPT	WORD	DEMOGRAFICA	DIAGNOSTICO	SATISFACCION		EXCEL	PPT	WORD
PRÁCTICO	81	0	87	0.00	1.25	2.50	0	81	0	87
PRÁCTICO	86	0	80	0.00	0.00	2.50	0	86	0	80
PRÁCTICO	67	50	96	1.25	1.25	2.50	5	67	93	100
PRÁCTICO	100	89	80	1.25	1.25	2.50	5	100	94	80
PRÁCTICO	91	0	100	0.00	0.00	2.50	0	91	0	100

Figura 5.15: Reporte puntos extras encuestas vs exámenes prácticos

Nivel de conocimiento

En la Figura 5.16 se presenta el reporte del nivel de conocimiento por cada componente del curso MOOC HCD, tal como: prueba de diagnóstico, notas MOOC, examen teórico y exámenes prácticos de los estudiantes de ESPOL y Admisiones del curso MOOC repetidores y MOOC completo, donde se muestra información personal del estudiante y los niveles de conocimiento en relación con los puntos obtenidos en cada componente.

	EXAMENES TEORICOS						EXAMENES PRACTICOS			
	HERRAMIENTAS	COMPUTADOR	INTERNET	POWER POINT	EXCEL	WORD	HERRAMIENTAS	POWER POINT	EXCEL	WORD
INICIAL	INICIAL	null	null	null	null	null	null	null	null	null
DESARROLLADO	DESARROLLADO	EXCELENTE	EXCELENTE	EN DESARROLLO	EXCELENTE	DESARROLLADO	EXCELENTE	DESARROLLADO	REGULAR	DESARROLLADO
EN DESARROLLO	EXCELENTE	EN DESARROLLADO	DESARROLLADO	EXCELENTE	DESARROLLADO	DESARROLLADO	EN DESARROLLADO	DESARROLLADO	REGULAR	DESARROLLADO
COMPLETO	DESARROLLADO	EXCELENTE	EXCELENTE	EN DESARROLLO	REGULAR	REGULAR	INICIAL	EN DESARROLLADO	EN DESARROLLADO	EN DESARROLLADO
REPETIDORES	null	null	null	null	null	null	null	null	null	null

Figura 5.16: Reporte nivel de conocimiento vs componentes MOOC HCD

De la Figura 5.17 a la Figura 5.23 se presenta la comparativa del progreso del nivel de conocimiento de cada módulo por cada componente del curso MOOC HCD, tal como: prueba de diagnóstico, notas MOOC, examen teórico y exámenes prácticos de los estudiantes de ESPOL y Admisiones del curso MOOC repetidores y MOOC completo.

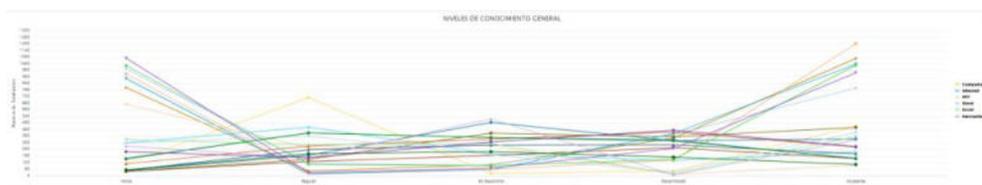


Figura 5.17: Reporte nivel de conocimiento de los módulos vs cada uno de los componentes MOOC HCD

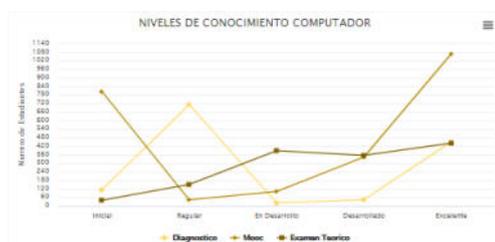


Figura 5.18: Reporte nivel de conocimiento El Computador

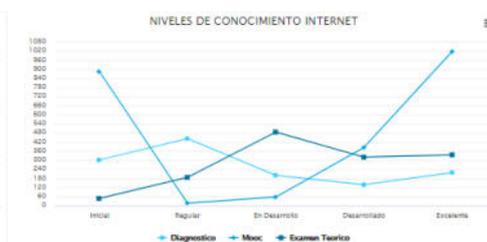


Figura 5.19: Reporte nivel de conocimiento El Internet

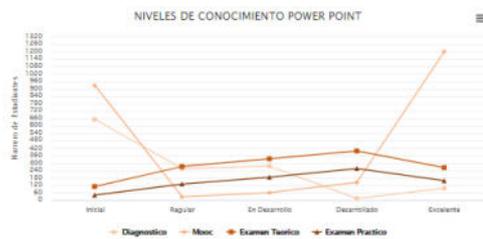


Figura 5.20: Reporte nivel de conocimiento PowerPoint

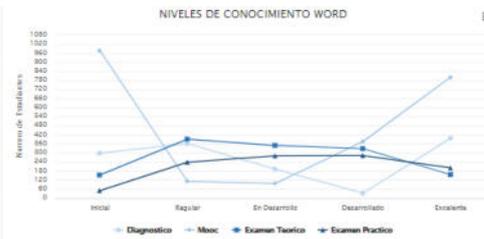


Figura 5.21: Reporte nivel de conocimiento Word

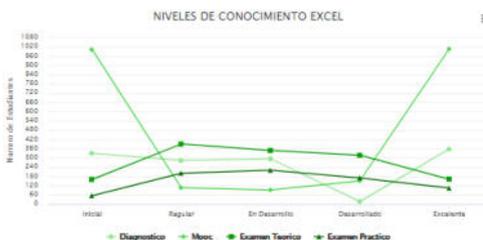


Figura 5.22: Reporte nivel de conocimiento Excel

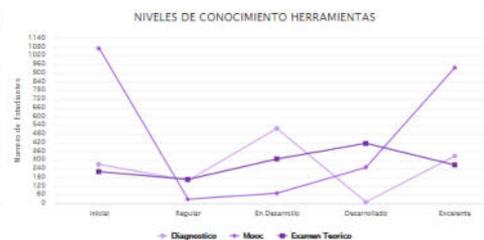


Figura 5.23: Reporte nivel de conocimiento Herramientas Colaborativas

De la Figura 5.24 a la Figura 5.28 muestra el progreso del nivel de conocimiento de cada componente del curso MOOC HCD, tal como: prueba de diagnóstico, notas MOOC, examen teórico y exámenes prácticos de los estudiantes de ESPOL y Admisiones del curso MOOC repetidores.

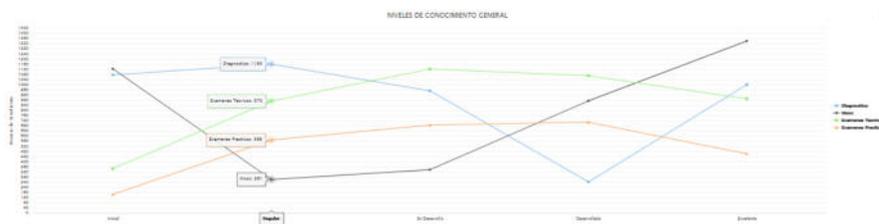


Figura 5.24: Nivel de conocimiento por componente MOOC HCD

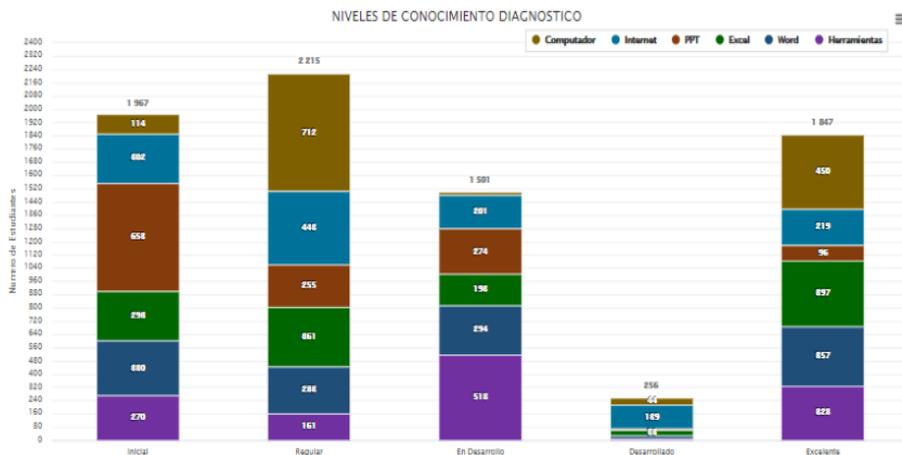


Figura 5.25: Nivel de conocimiento prueba de Diagnóstico MOOC HCD

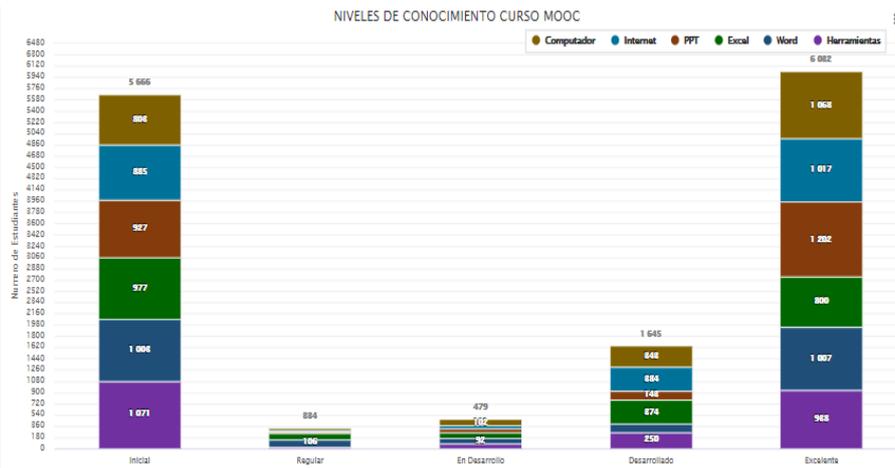


Figura 5.26: Nivel de conocimiento curso edX MOOC HCD

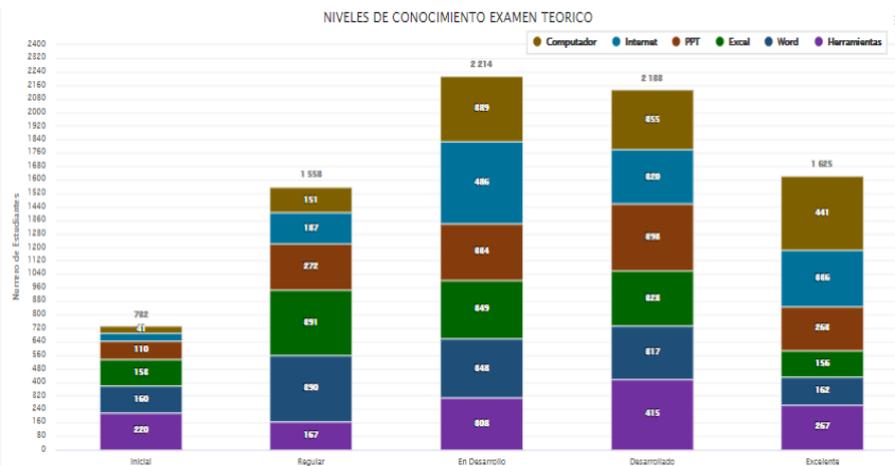


Figura 5.27: Nivel de conocimiento examen teórico MOOC HCD

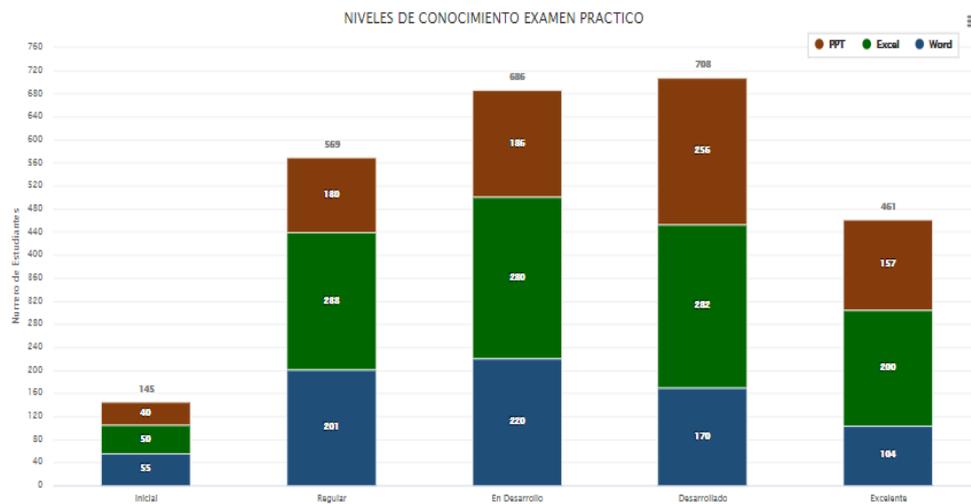


Figura 5.28: Nivel de conocimiento examen práctico MOOC HCD

En la Figura 5.29 se presenta el reporte GTSI donde se muestra información personal del estudiante de ESPOL y Admisiones del curso MOOC repetidores y MOOC completo, franja y paralelo donde rindieron el examen práctico, fecha de inicio del curso MOOC, nota del examen teórico y exámenes prácticos con el incremento de los puntos extras.

ID	EMAIL	FRANJA	PARALELO	FECHA DE EXAMEN	EXAMEN TEORICO	EXAMENES PRACTICO			APROBACION	COHORTE
						WORD	PPT	EXCEL		
23425	adriana@espol.edu.ec	4	8	2020-06-10	82.00	79	0	88	REPROBADO	PRACTICO
23607	adriana@espol.edu.ec	-	-	2020-06-10	0.00	0	0	0	REPROBADO	PRACTICO
23646	adriana@espol.edu.ec	1	8	2020-06-10	84.00	84	0	88	REPROBADO	PRACTICO
23679	adriana@espol.edu.ec	2	1	2020-06-10	76.00	88	66	83	APROBADO	PRACTICO
23820	adriana@espol.edu.ec	-	-	2020-06-10	0.00	0	0	0	REPROBADO	PRACTICO

Figura 5.29: Reporte GTSI

5.2. Análisis de resultados

En este proyecto se cumplió con lo solicitado por parte de los administradores, lo cual fue, la integración de datos, la sistematización

de los reportes del curso MOOC HCD y la generación de gráficas estadísticas, esto se pudo comprobar en las pruebas de usabilidad.

La generación de 21 reportes distribuidos en 7 menús, como se ilustra en Figura 5.30, la que antes conllevaba un trabajo aproximado de 2 semanas realizado por tres personas, ahora el sistema lo realiza en 81 minutos. Esto permite al personal administrativo centrarse en otros objetivos relevantes como la actualización de contenido, la elaboración del banco de exámenes prácticos y teórico, y el seguimiento del progreso de los estudiantes.



Figura 5.30: Aplicación Web MOOC HCD

Las gráficas les permiten a los administradores identificar los puntos críticos donde los estudiantes bajan su rendimiento, de tal forma que, puedan reforzar con más actividades formativas y sumativas.

Dado que el curso MOOC HCD pasa por diferentes etapas de evaluación, el administrador podrá agregar los archivos resultantes de forma paulatina en el repositorio y con ello generar el proceso ETL, de esta forma va obteniendo reportes inmediatos por etapa, que le

ayudarán a la toma de decisiones preventivas. Esto también permite que el proceso ETL se realice en tiempos más cortos, dado que se va generando por etapas.

CONCLUSIONES Y RECOMENDACIONES

1. Los reportes se los genera de forma inmediata, una vez hayan finalizado las etapas de evaluación del curso MOOC.
2. Se usó Highchars para la visualización de reportes y gráficos estadísticos.
3. En MySQL se desarrolló la base de datos integrada que contiene la información proveniente de tres fuentes de datos y en la cual se lleva a cabo el proceso ETL.
4. Los tiempos de elaboración de reportes fueron optimizados de días a horas, permitiendo que los administradores se enfoquen en otros objetivos prioritarios.
5. Este proyecto se enfocó en la data de los cursos MOOC completo y MOOC repetidores del 2020-2T, en el cual los sistemas fuentes fueron el Sistemas Académico, Sidweb y EdX.
6. La pandemia por el COVID19 provocó un cambio en el proceso habitual del curso MOOC HCD, dado que, en este término abrieron dos cursos MOOC completo, uno para estudiantes de admisiones y otro para estudiantes ESPOL, lo cual no es habitual, debido a que, en un mismo curso se registran a los dos tipos de estudiantes.

7. El administrador creará el repositorio y agregará los archivos fuente en base a la estructura definida en este proyecto, para que se ejecuten los procesos.

BIBLIOGRAFÍA

- [1] F. Bolivar, «Synergixs,» 13 Julio 2019. [En línea]. Available: <https://synergixs.com/que-es-integracion-de-datos-etl/>. [Último acceso: 10 Mayo 2021].
- [2] T. Naeem, «Enabling Data Driven Innovation,» Astera, 20 noviembre 2020. [En línea]. Available: <https://www.astera.com/es/tipo/blog/integraci%C3%B3n-de-datos-del-cliente/>. [Último acceso: 10 mayo 2021].
- [3] «Talend Data Fabric,» [En línea]. Available: <https://www.talend.com/es/resources/what-is-data-integration/>. [Último acceso: 10 mayo 2021].
- [4] POWERDATA SOLUTIONS, S.L., Procesos ETL, Buenos Aires: POWERDATA SOLUTIONS, S.L., 2013.
- [5] «Tecnologías Información,» [En línea]. Available: <https://www.tecnologias-informacion.com/integracion.html>. [Último acceso: 12 mayo 2021].
- [6] Josep Curto y Jordi Conesa, «El conocimiento imprescindible,» [En línea]. Available: <http://reader.digitalbooks.pro/content/preview/books/43005/book/OEBPS/chapter02.xhtml>.

- [7] M. Tena, «Grupo BBVA,» [En línea]. Available:
<https://www.bbva.com/es/metodologia-agile-la-revolucion-las-formas-trabajo/>.
[Último acceso: 14 Mayo 2021].
- [8] «Enciclopedia Económica,» 2018. [En línea]. Available:
<https://enciclopediaeconomica.com/metodologias-agiles/>. [Último acceso: 14 mayo 2021].
- [9] V. R. Villán, «IEBS, Innovation & Entrepreneurship Business School,» 15 marzo 2019.
[En línea]. Available: <https://www.iebschool.com/blog/que-son-metodologias-agiles-agile-scrum/>. [Último acceso: 14 mayo 2021].
- [10] «Sinnaps,» [En línea]. Available: <https://www.sinnaps.com/blog-gestion-proyectos/tipos-de-metodologias-agiles>. [Último acceso: 14 mayo 2021].
- [11] «Tecnologías para los negocios,» Camara Valencia, [En línea]. Available:
<https://ticnegocios.camaravalencia.com/servicios/tendencias/desarrollo-proyectos-agiles-scrum/#1499947769140-cbc0643e-52d2>. [Último acceso: 16 Mayo 2021].
- [12] K. Schwaber y J. Sutherland, La Guía Definitiva de Scrum, 2017.
- [13] L. SALISMEY, «Compara Software,» 17 julio 2020. [En línea]. Available:
<https://blog.comparasoftware.com/scrum-ventajas-y-desventajas/#:~:text=%20Scrum%3A%20Ventajas%20y%20Desventajas%20%201%20>

Ventajas,m%C3%A1ximo%20el%20tiempo%20y%20los%20recursos%2C...%20More%
20. [Último acceso: 16 mayo 2021].

- [14] O. M. Alejandro Molina, J. V. Pincay Nieves y D. M. Pesantez Pesantez, «Creación de MOOC en una Universidad Latina - Una experiencia de aciertos y errores,» EMOOCs, Italia, 2017.