

MUYSCA: sistema de enriquecimiento del perfil de una persona en el marco de un campus
inteligente

Viky Julieta Arias Delgado

PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN
MAESTRÍA EN ANALÍTICA PARA LA INTELIGENCIA DE NEGOCIOS
BOGOTÁ, D.C.
2020

MUYSCA: sistema de enriquecimiento del perfil de una persona en el
marco de un campus inteligente

Autor:

Viky Julieta Arias Delgado

MEMORIA DEL TRABAJO DE GRADO REALIZADO PARA CUMPLIR UNO
DE LOS REQUISITOS PARA OPTAR AL TÍTULO DE
MAGÍSTER EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

Director

Ingeniero Enrique González Guerrero, PhD

Comité de Evaluación del Trabajo de Grado

Ingeniera Anabel Montero Posada, MS

Ingeniero Ricardo Fernando Otero Caicedo, MS

Página web del Trabajo de Grado

<https://livejaverianaedu.sharepoint.com/sites/Ingsis/TGMISC/203003>

PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN INGENIERIA DE SISTEMAS Y COMPUTACIÓN
BOGOTÁ, D.C.
Diciembre, 2020

MUYSCA: sistema de enriquecimiento del perfil de una persona en el
marco de un campus inteligente

Autor:

Viky Julieta Arias Delgado

MEMORIA DEL TRABAJO DE GRADO REALIZADO PARA CUMPLIR UNO
DE LOS REQUISITOS PARA OPTAR AL TÍTULO DE
MAGISTRA EN ANALÍTICA PARA LA INTELIGENCIA DE NEGOCIOS

Director

Ingeniero Enrique González Guerrero, PhD

Comité de Evaluación del Trabajo de Grado

Ingeniera Anabel Montero Posada, MS

Ingeniero Ricardo Fernando Otero Caicedo, MS

Página web del Trabajo de Grado

<https://livejaverianaedu.sharepoint.com/sites/Ingsis/TGMISC/203003>

PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN ANALÍTICA PARA LA INTELIGENCIA DE NEGOCIOS
BOGOTÁ, D.C.
Diciembre, 2020

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**

Rector Magnífico

Jorge Humberto Peláez, S.J.

Decano Facultad de Ingeniería

Ingeniero Lope Hugo Barrero Solano

Director Maestría en Ingeniería de Sistemas y Computación

Ingeniera Angela Carrillo Ramos

Director Departamento de Ingeniería de Sistemas

Ingeniero Efraín Ortíz Pabón

**PONTIFICIA UNIVERSIDAD JAVERIANA
FACULTAD DE INGENIERIA
MAESTRÍA EN ANALÍTICA PARA LA INTELIGENCIA DE NEGOCIOS**

Rector Magnífico

Jorge Humberto Peláez, S.J.

Decano Facultad de Ingeniería

Ingeniero Lope Hugo Barrero Solano

Director Maestría en Analítica para la Inteligencia de Negocios

Ingeniero Jorge Andrés Alvarado Valencia

Director Departamento de Ingeniería Industrial

Ingeniero Carlos Eduardo Fúquene Retamoso

Artículo 23 de la Resolución No. 1 de Junio de 1946

“La Universidad no se hace responsable de los conceptos emitidos por sus alumnos en sus proyectos de grado. Sólo velará porque no se publique nada contrario al dogma y la moral católica y porque no contengan ataques o polémicas puramente personales. Antes bien, que se vean en ellos el anhelo de buscar la verdad y la Justicia”

AGRADECIMIENTOS

A mi director de trabajo de grado, ingeniero Enrique González Guerrero, PhD, por sus valores humanos, su experiencia profesional, su paciencia y su motivación que me impulsaron a finalizar con éxito esta investigación.

A Ronald, compañero de vida, por su amor, su ejemplo y su fortaleza que me han inspirado durante el maravilloso camino recorrido en las maestrías: un sueño y retos construidos en familia.

A mi familia, especialmente a German, Aydeé, Luz, Astrid y Yamile, por el bienestar y los pilares brindados en valores, esfuerzo contante y disciplina que han direccionado mi vida.

A la directora de la Maestría en Ingeniería de Sistemas y Computación, ingeniera Ángela Cristina Carrillo Ramos, PhD por su valiosa disposición a servir, consejos y acompañamiento en el camino recorrido.

Al director de la Maestría en Analítica para la Inteligencia de Negocios, ingeniero Jorge Andrés Alvarado Valencia, PhD por su valioso acompañamiento y atención de inquietudes del programa.

Al equipo de la Dirección de Tecnologías de Información, por su apoyo en la investigación mediante el acceso a datos asociados a la investigación.

Al equipo Smart-UJ, por su apoyo en la investigación mediante los aportes constantes, sobre diferentes enfoques asociados al trabajo de grado.

Al equipo del Sistema de Alertas Tempranas de Intervención y Seguimiento, por su acompañamiento constante en el proceso de investigación.

Al equipo de la Biblioteca Alfonso Borrero Cabal, S.J. por los espacios de asesoría especializada y pruebas funcionales realizadas en el proceso de validación.

A los expertos evaluadores de los servicios de Muysca, por sus valiosos aportes y nuevas ideas que permitieron la evolución y mejora de la investigación.

A todos los profesores de las maestrías, por sus enseñanzas y su entrega en la noble labor de enseñar, inspirar y crear nuevas soluciones a las problemáticas de la sociedad.

CONTENIDO

INTRODUCCIÓN	15
1. DESCRIPCIÓN DEL PROYECTO	16
1.1 PROBLEMÁTICA.....	16
1.2 OPORTUNIDAD.....	16
1.3 OBJETIVO GENERAL	18
1.4 OBJETIVOS ESPECÍFICOS	18
1.5 FASES DE DESARROLLO	18
1.5.1 Fase I: Analizar.....	19
1.5.2 Fase II: Diseñar	20
1.5.3 Fase III: Validar.....	21
2. MARCO TEÓRICO / ESTADO DEL ARTE	22
2.1 MARCO TEÓRICO.....	22
2.2 ESTADO DEL ARTE	27
2.2.1 Caracterización de Servicios Potenciales en un Campus Inteligente	27
2.2.2 Servicios Basados en Localización	28
2.2.3 Perfil de Usuario y Contexto.....	29
2.2.4 Inteligencia Artificial para Detección de Comportamientos y Hábitos	30
2.2.5 Inteligencia Artificial para Servicios Inteligentes en un CamIn.....	31
2.2.6 Deserción Estudiantil.....	31
2.2.7 Selección del Caso de Referencia.....	32
2.2.8 Consideraciones de Seguridad y Ética.....	33
3. TRABAJOS RELACIONADOS PUJ.....	34
3.1 MODELO AMI-IOT <i>QUYSQUA</i>	34
3.2 PROYECTO <i>SMART-UJ</i>	35
3.3 SPACE ANALYTICS	39
3.4 IAM: MODELO INTEGRADO DE ADAPTACIÓN	40
3.5 ESCENARIOS DE DESERCIÓN ESTUDIANTIL EN LA PUJ	41
3.6 FORTALEZAS Y APORTES DE TRABAJOS RELACIONADOS	42
4. MODELO MUYSCA	44
4.1 DESCRIPCIÓN GENERAL DEL MODELO MUYSCA.....	44
4.2 PERFILES DE ENRIQUECIMIENTO <i>MUYSCA</i>	45
4.2.1 Perfil de Usuario.....	46

4.2.2 Perfil de Contexto.....	46
4.3 SERVICIOS MUYSKA ASOCIADOS AL CASO DE REFERENCIA	48
4.4 PROCESO DE DETECCIÓN DE HÁBITOS EN MUYSKA	50
5. VALIDACIÓN DEL MODELO MUYSKA	51
5.1 METODOLOGÍA ANALÍTICA.....	51
5.1.1 Comprensión del Negocio	51
5.1.2 Comprensión de los Datos.....	52
5.1.3 Preparación de los Datos.....	55
5.1.4 Modelado.....	56
5.1.5 Evaluación.....	57
5.1.6 Implementación	60
5.2 PROTOTIPO.....	62
5.3 VALIDACIÓN DE RESULTADOS.....	65
5.3.1 Caso de Referencia.....	65
5.3.2 Validación con Expertos.....	67
6. CONCLUSIONES.....	70
7. ANEXOS	71
REFERENCIAS	72

LISTADO DE FIGURAS

Figura 1. Fases de Desarrollo Muysca.....	19
Figura 2. Análisis causa-efecto a partir del análisis del estado del arte.....	33
Figura 3. Trabajo Relacionado P.U.J. Modelo Quysqua.....	34
Figura 4. Modelo de agentes para crear el perfil de una persona.....	36
Figura 5. Arquitectura de alto nivel <i>Quysqua Smart-UJ</i>	38
Figura 6. Arquitectura trabajo relacionado Space Analytics.....	40
Figura 7. Modelo de trabajo relacionado IAM.....	41
Figura 8. Aplicación del modelo <i>Quysqua Smart-UJ</i> en Muysca.....	44
Figura 9. Componentes del perfil enriquecido de una persona en Muysca.....	45
Figura 10. Elementos para enriquecer servicios de Muysca.....	46
Figura 11. Perfil de usuario del modelo Muysca.....	46
Figura 12. Perfil de contexto del modelo Muysca.....	47
Figura 13. Representación de lugares en el campus P.U.J.....	48
Figura 14. Descripción servicio de asistencia a clase.....	49
Figura 15. Descripción servicio de actividades frecuentes.....	49
Figura 16. Descripción servicio gustos y preferencias temáticos.....	50
Figura 17. Recopilación inicial de los datos.....	52
Figura 18. Exploración de datos y selección de estudiantes para análisis exploratorio.....	54
Figura 19. Análisis de resultados de clustering para el estudiante 105.....	58
Figura 20. Descripción implementación componentes segmento DS.....	60
Figura 21. Validación de resultado - caso de referencia Muysca.....	61
Figura 22. Diagrama de despliegue <i>Quysqua Smart-UJ</i> aplicado a Muysca.....	63
Figura 23. Diagrama de clases sistema Muysca.....	64
Figura 24. Representación gráfica del caso de referencia Muysca.....	65
Figura 25. Análisis de resultados de la primera sección de la encuesta.....	68
Figura 26. Análisis evaluación de expertos primera sección.....	69

LISTADO DE TABLAS

Tabla 1. Definiciones de Campus Inteligente (CamIn).....	23
Tabla 2. Análisis de algunos algoritmos de clustering.....	25
Tabla 3. Criterios de Evaluación Sistemática de la Bibliografía	27
Tabla 4. Comparación con el estado del arte	29
Tabla 5. Análisis del Estado del Arte para Perfil de Usuario y Contexto	29
Tabla 6. Análisis del Estado del Arte sobre Inteligencia Artificial en un CamIn	32
Tabla 7. Fortaleza de trabajos relacionados	42
Tabla 8. Mapeo de lugares de permanencia para el modelo Muysca.....	47
Tabla 9. Descripción inicial de los datos	53
Tabla 10. Estudiantes seleccionados para análisis exploratorio.....	53
Tabla 11. Formato de los datos para análisis exploratorio.....	56
Tabla 12. Evaluación de modelos ejecutados	59
Tabla 13. Validación de arquitectura <i>Quysqua Smart-UJ</i> aplicada a los servicios Muysca....	66

ABSTRACT

Muysca develops feature extraction mechanisms based on location and transactional data to allow people's profile enrichment. By performing an analysis of the state of the art, we characterized different key services within the framework of an intelligent campus. We perform implementations of some components of the reference architecture: *Quysqua Smart-UJ*. *Muysca* was designed using the Engineering Basic Design Cycle for the reference case, associated with the services: class attendance, frequent activities, and thematic preferences and tastes.

RESUMEN

Muysca desarrolla mecanismos de extracción de características basados en la ubicación y los datos transaccionales para permitir el enriquecimiento del perfil de las personas. Al realizar un análisis del estado del arte, caracterizamos diferentes servicios clave en el marco de un campus inteligente. Realizamos implementaciones de algunos componentes de la arquitectura de referencia: *Quysqua Smart-UJ*. *Muysca* se diseñó utilizando el Ciclo de Diseño Básico de Ingeniería para el caso de referencia, asociado a los servicios: asistencia a clase, actividades frecuentes y preferencias y gustos temáticos.

RESUMEN EJECUTIVO

Un campus inteligente busca brindar servicios a la comunidad universitaria, que fortalezcan los aspectos de conocimiento, aprendizaje, bienestar, medio ambiente, operación y sociedad. A partir de información tanto transaccional como adquirida con sensores, se puedan realizar análisis que permitan prestar mejores servicios adaptados al contexto. Por ejemplo, si el sistema pudiese reconocer aspectos del “estilo de vida” de una persona, podría realizar recomendaciones o acompañamiento a diferentes tipos de usuario según sus necesidades, mediante el uso de tecnologías emergentes.

Las tecnologías emergentes, han transformado la forma de procesar la información en tiempo real captando la información de los usuarios desde diferentes tipos de sensores, como por ejemplo el internet de las cosas, la inteligencia ambiental, entre otros. La integración de estas tecnologías permite prestar servicios que generen acciones a corto plazo como reglas reactivas, en caso de que un evento ocurra, o impacto a largo plazo mediante el uso de técnicas y estrategias de analítica, a partir del comportamiento de las personas, que aporten a problemas de las universidades como la deserción estudiantil.

La deserción estudiantil es una problemática que preocupa a instituciones privadas y públicas en el ámbito nacional e internacional. El Ministerio de Educación de Colombia ha investigado durante los últimos años diferentes factores determinantes que influyen en la deserción como características individuales, características socioeconómicas, características académicas y características institucionales. Frente a este panorama, la Pontificia Universidad Javeriana ha creado el proyecto Sistema de Alertas Tempranas, Intervención y Seguimiento, facilitando el monitoreo académico de los estudiantes, el acompañamiento constante desde la formación integral y la *cura personalis*, que permita aportar en el éxito académico, la prevención de la deserción y la graduación oportuna de los estudiantes.

El Sistema de Alertas Tempranas, Intervención y Seguimiento busca identificar las condiciones de alerta de deserción, que permitan determinar el nivel de riesgo de los estudiantes, establecer las acciones concretas que pueden desarrollarse para mitigar dicho riesgo, y definir un protocolo de intervención y seguimiento. También, se generan indicadores de alertas e intervenciones que permiten realimentar el sistema y fortalecer la cultura de seguimiento de acompañamiento al estudiante y mejora de los procesos institucionales.

Surge la oportunidad de investigación, para enriquecer el perfil de un estudiante y diseñar servicios adaptados al contexto a partir de su localización en el campus. Este enriquecimiento aporta a la problemática de deserción estudiantil, mediante el consumo de nuevas características y aportan información valiosa al proyecto de Alertas Tempranas, Intervención y Seguimiento y al consejero académico. La pregunta motivadora del trabajo de grado *Muysca* plantea cómo se puede enriquecer de forma *non-intrusive* el perfil de una persona, en el marco de un campus inteligente, brindando mejores servicios adaptados a su contexto. El perfil enriquecido incluye información comportamental, detectada mediante hábitos de conexión a redes *Wifi* del campus, *RFID*, datos transaccionales y desde cualquier tipo de dispositivo. Información del perfil y del contexto, así como el análisis de hábitos puede aportar al fenómeno de la deserción, facilitando nuevas características al perfil de un estudiante.

Para responder a la pregunta de investigación, este trabajo aplicó inicialmente la arquitectura *Quysqua Smart-UJ* centrado en servicios basados en los hábitos y la localización de una **persona**, enriqueciendo el perfil y profundizando en los perfiles de usuario y de contexto. Para lograrlo, se investigó en el primer nivel la parte sensorial, que permite explorar información pertinente, de forma *non-intrusive* y continua sobre las personas. Con esta información, se define en un segundo nivel, la detección del comportamiento de las personas (acciones, actividades y lugares). Finalmente, en el tercer nivel, con una visión de largo plazo, se detectan hábitos (AD) con información sobre los gustos, preferencias y elementos del “estilo de vida”. Estos elementos son útiles para tener un perfil más completo de las personas, con el fin de prestar servicios adaptados al contexto.

El desarrollo del sistema *Muysca* se inspiró en el Ciclo Básico de Diseño de Ingeniería y se desarrolló en tres fases: análisis, diseño y validación. En la primera fase se aplicó una metodología exploratoria para caracterizar los servicios potenciales relacionados con la problemática, a partir del estado del arte. A partir del análisis del estado del arte se obtuvo el caso de referencia asociado a la deserción estudiantil en relación con los servicios de asistencia a clase, actividades frecuentes y gustos y preferencias temáticas de acuerdo con la localización.

En la segunda fase, se ajustó el caso de referencia de la investigación y se definieron los componentes a profundizar en la arquitectura de referencia *Quysqua Smart-UJ*. Posteriormente, en el marco de la investigación *Muysca* se desarrollaron algunos componentes de la arquitectura de referencia. En la capa *Adaptation Services*, se realizó la construcción del perfil de usuario y de contexto de una persona a partir de la localización. En los componentes *Node Data Manager* y *Network Data Consumer* se construyó el mapeo de lugares del campus. En la capa *Analytic Tools* se realizó un análisis exploratorio, siguiendo la metodología *CRISP-DM* y aplicando mecanismos de clustering para conocer hábitos. Finalmente, se validó el consumo de los servicios de *Muysca* en la capa *User Experience Tools* del segmento AS, considerando aspectos de seguridad y acceso mediante *LDAP*.

Finalmente, en la tercera fase, se validaron los servicios propuestos con expertos tomando como inspiración la metodología *TAM*. Mediante la aplicación de las tareas de modelado, evaluación e implementación de la metodología analítica *CRISP-DM* se obtuvo información de los datos asociados a *Wifi*, *RFID* e información transaccional relacionada con el caso de referencia de deserción estudiantil, en cuanto a las características que enriquecen el perfil de una persona a partir de tecnologías emergentes.

Como trabajo futuro se plantean nuevas integraciones de información proveniente de sensores que permita realizar otros análisis y modelos con enfoque predictivo, implementación de los servicios *Muysca* en producción en el marco del proyecto *Smart-UJ*, continuidad en la investigación sobre características sociales del perfil de una persona para validar grupos de interés y recomendaciones a partir de los servicios propuestos enfocados al bienestar, la investigación y el desarrollo integral en el Campus Javeriano, considerando como base los datos de localización de forma física de un estudiante en el campus.

INTRODUCCIÓN

Del concepto de ciudad inteligente, se deriva el concepto de campus inteligente, el cual plantea objetivos de mejora en la calidad de vida de las personas, sostenibilidad, administración eficiente de recursos, diseño y planeación de los servicios y bienestar e interacción de las personas [1]. En un campus inteligente universitario, se propone el diseño de servicios adaptados al contexto para lograr los objetivos planteados. Por ejemplo, monitoreo del internet de banda ancha, monitoreo en el consumo de recursos impulsado por las tecnologías de información y comunicación (TIC), renovación de espacios de clase y monitoreo de salud de las personas [2].

Los servicios adaptados al contexto se basan en el concepto conciencia del contexto, siendo es un componente esencial en el diseño de inteligencia ambiental y computación inteligente. De esta manera, es posible reconocer las situaciones en las que la persona espera que se provean ciertos servicios adaptados, mejorando la usabilidad y la interacción cotidiana, facilitada a partir de tecnologías emergentes [3]. El servicio adaptado también debe considerar el perfil de la persona, detectando mediante tecnologías emergentes necesidades, preferencias o hábitos [4]. La información del perfil puede ser obtenida de fuentes transaccionales o mediante sensores. Por ejemplo, el modelo de adaptación basado en el perfil de la persona debe permitir adaptar y personalizar los productos/servicios de portales web, de acuerdo con la información básica, preferencias, gustos, intereses y contexto propios de la persona que realiza la navegación en el sitio [5].

El modelo *Muysca* plantea el diseño de servicios adaptados al contexto a partir de la localización de una persona en un campus inteligente. Los servicios *Muysca* se asocian a hábitos de asistencia a clase, lugares frecuentes de conexión *Wifi* en el campus de forma presencial y hábitos de uso de la biblioteca, de acuerdo con los préstamos de recursos realizados por un estudiante. Se aplica la arquitectura *Quysqua Smart-UJ* para el diseño de los servicios *Muysca* y se enriquecen algunos componentes como la capa *Adaptation Services*, perteneciente al segmento *AS* en cuando a la construcción del perfil de usuario y de contexto de una persona a partir de la localización; componente *Node Data Manager* y *Network Data Consumer* de segmento *DS*; el mapeo de lugares del campus; el consumo de los servicios de *Muysca* en la capa *User Experience Tools* del segmento *AS*, gestionando el acceso a la información mediante niveles de seguridad y autenticación con *Lightweight Director Access Protocol (LDAP)* y análisis exploratorio en la capa *Analytic Tools* del segmento *Data Analytics (DA)*, mediante clustering *Density-based spatial clustering of applications with noise (DBSCAN)* por día de la semana y hora del día asociado a la conexión de un *AP Wifi* por parte de una persona.

El presente documento se estructura en seis capítulos. En el primer capítulo, se realiza la descripción del proyecto, incluyendo la problemática, oportunidad de investigación, objetivos y fases de desarrollo. En el segundo capítulo, se presenta el marco teórico y el estado del arte. En el tercer capítulo, se analizan los trabajos realizados en la Pontificia Universidad Javeriana relacionados con la investigación. En el cuarto capítulo, se describe el modelo *Muysca*, mediante la aplicación de la arquitectura *Quysqua Smart-UJ* y la construcción de perfiles de usuario y de contexto asociado a los servicios propuestos. En el quinto capítulo, se detalla la validación del modelo *Muysca*. Finalmente, en el sexto capítulo se presentan las conclusiones y trabajo futuro.

1. DESCRIPCIÓN DEL PROYECTO

En este capítulo se realiza la descripción del trabajo de grado *Muysca*, en cuanto a la problemática y la oportunidad de investigación, el objetivo general, los objetivos específicos del trabajo de grado y las fases de desarrollo fundamentadas en el ciclo básico de ingeniería.

1.1 PROBLEMÁTICA

Un *Campus Inteligente* (en adelante *CamIn*) busca prestar servicios inteligentes a la comunidad universitaria, que aporten a problemáticas de conocimiento, aprendizaje, el bienestar, medio ambiente, operación y sociedad. La mayoría de estos servicios, requieren el conocimiento del comportamiento de una persona. Algunas formas de conocer el comportamiento son el reconocimiento de patrones [6], servicios basados en agentes inteligentes, la detección de hábitos [7], o lugares visitados [8], entre otros. Las universidades enfrentan retos frente a aspectos como la infraestructura, la calidad académica o la deserción estudiantil [9].

La deserción estudiantil es una problemática significativa para las universidades [10]. Algunas investigaciones internacionales, resaltan factores de deserción como niveles de satisfacción, percepciones de apoyo institucional y adaptación cultural [11]. Según el Ministerio de Educación, para el año 2015 el 46,1% de los estudiantes que ingresan en un periodo académico abandonan sus estudios [12]. Por otra parte, en la Pontificia Universidad Javeriana (en adelante PUJ) el índice de deserción en el periodo 2017-1 fue de 33,3 %, en comparación con el periodo 2016-1 que fue de 32,7 % [13]. En este sentido, la Universidad tiene el reto de desarrollar estrategias de monitoreo académico de los estudiantes, que permita identificar y hacer seguimiento a los estudiantes según sus riesgos de deserción.

El programa de prevención de la deserción en la PUJ considera factores académicos, individuales, socio económicos e institucionales, que permiten apoyar a un estudiante la formación integral y *cura personalis* [14]. Sin embargo, en cuanto a los factores académicos e individuales, no se conocen los patrones de movilidad y mecanismos de enriquecimiento del perfil en el campus a partir de la localización de un estudiante. Estos patrones están asociados a lugares de permanencia como ambientes de estudio individual o grupal, aulas de clase, servicios de alimentación, espacios deportivos, entre otros. Indicadores como conocer la asistencia a clase o conocer los lugares de interés de un estudiante de forma no intrusiva, podrían apoyar actividades de consejería y apoyo por parte del consejero académico, en el contexto de un *CamIn*.

1.2 OPORTUNIDAD

La problemática descrita en el apartado anterior ha tenido diferentes aproximaciones y soluciones. Servicios inteligentes soportados en el uso e integración de tecnologías emergentes como el *internet de las cosas* (en adelante *IoT*), la *inteligencia ambiental* (en adelante *AmI*) y la *analítica de datos* (en adelante *DA*) [15], buscan mejorar la calidad de vida y la sostenibilidad de la comunidad universitaria [16]. En general, se plantea que, a partir de información tanto transaccional o adquirida con sensores, se puedan realizar análisis de datos que permitan realizar recomendaciones adaptadas basadas en el usuario, el contexto o el aprendizaje, a partir del “estilo de vida” de una persona en el campus [17].

Las tecnologías emergentes, han transformado la forma de procesar la información en tiempo real [18]. Mediante *IoT* se pueden realizar el análisis del flujo de personas basado en redes inalámbricas [3]. Por su parte, aplicaciones soportadas en *AmI* facilitan la personalización de servicios, a partir del contexto de una persona [19]. Finalmente, con *DA* se pueden encontrar relaciones entre las variables en un estudio con horizontes espacio temporales más amplios; por ejemplo, realizar la detección de anomalías en una tienda, a partir del comportamiento de las personas [20].

Dentro de las formas de estudio del comportamiento de una persona, se encuentran el reconocimiento de patrones [21], la detección de hábitos [22], el análisis de lugares visitados [23], entre otros. La exploración del contexto facilita la oferta de servicios de forma reactiva, de acuerdo con la ubicación [24]. Otros modelos, se centran en acotar el espacio personal, con el uso de herramientas como simuladores [25]. También existen servicios basados en agentes inteligentes, que toman la información de entrada de sensores de redes inalámbricas [26].

Frente al problema de identificación de patrones de movilidad mediante la observación del comportamiento de una persona se han propuesto soluciones a través de trabajos de grado y proyectos de planeación universitaria de la PUJ. En 2018, Luis Barreto [27] propuso el modelo *Quysqua*, el cual contiene cuatro segmentos: *Aplicación*, que provee servicios adaptados y herramientas para el desarrollo de aplicaciones; *Analítica de Datos*, que permite tomar decisiones; *Inteligencia Ambiental*, que permite interpretar los datos de los sensores y cambiar el estado de la vivienda a partir de los actuadores e *Internet de las Cosas*, que permite configurar las redes de nodos para supervisar y controlar la vivienda. Posteriormente, en 2019, se creó el proyecto *SmartUJ* [28], desde el enfoque de una comunidad inteligente que ofrece servicios soportados en tecnologías avanzadas, con el fin de fortalecer o enriquecer un campus universitario en las dimensiones de *Conocimiento*, *Aprendizaje*, *Medioambiente*, *Operación* y *Estrategia*, *Sociedad* y *Bienestar*. Este enfoque requiere la creación de equipos de trabajo con visión multidisciplinar que, frente a una necesidad planteada, ofrezcan una mirada desde el punto de vista institucional, funcional y tecnológico.

Frente al desarrollo de mecanismos de enriquecimiento del perfil de una persona, se han propuesto soluciones mediante investigaciones y proyectos de planeación universitaria de la PUJ. En 2009, Ángela Carrillo, et. al. [5] desarrollaron *IAM: Modelo Integrado de Adaptación*, incluyendo componentes de un modelo de adaptación integrado mediante cuatro componentes: *Módulo de Presentación* para desplegar la información al usuario en el dispositivo de acceso, *Módulo Contextual* para representar la interacción del usuario con el sistema, *Módulo de Conexión inalámbrica* para definir la tecnología de conexión más apropiada para el usuario a través de su dispositivo y *Módulo de Contenido* para establecer las características de un individuo y/o comunidad. El perfil individual de una persona considera las preferencias, gustos, intereses e información básica. Posteriormente, En 2018, la Vicerrectoría Académica creó el proyecto *Sistema de Alertas Tempranas de Intervención y Seguimiento* (en adelante *SATIS*) [29], una iniciativa colaborativa aplicado inicialmente en la Facultad de Ingeniería de la PUJ, que ha buscado la identificación de situaciones que ponen en riesgo el éxito estudiantil, acompañando a los estudiantes para que logren la graduación oportuna. Las alertas tempranas de *SATIS* incorporan el perfil de un estudiante desde fuentes transaccionales, asociados a factores académicos, individuales, socio económicos e institucionales.

Surge la oportunidad de desarrollar mecanismos de extracción de características que permitan enriquecer el perfil de un estudiante, a partir de la localización en el campus. A partir de los datos captados cada cinco minutos mediante *Wireless Fidelity* (en adelante *Wifi*), se hace posible construir trayectorias de lugares de permanencia por parte de un estudiante. Luego, mediante técnicas de analítica es viable detectar hábitos y patrones de movilidad en el campus. Estos elementos serán útiles para tener un perfil más completo de las personas a partir de información de comportamiento, con el fin de prestar servicios adaptados al contexto.

Tomando como referencia el modelo *Quysqua Smart-UJ*, se pueden enriquecer componentes del segmento *Application Service*: en la capa *Information Enrichment* integrando los componentes de perfil de usuario y perfil de contexto de un estudiante a partir de *IAM*, aspectos de seguridad, acceso entre los segmentos y en la capa *Application Tools* aplicando el consumo y la visualización de los datos. Posteriormente, es viable la prestación de servicios adaptados al contexto de un *CamIn*, como patrones de movilidad, detección de anomalías o detección de asistencia a clase y comportamiento en el uso de la biblioteca en cuanto a préstamos de libros, equipos y espacios que brinde información sobre el perfil enriquecido de un estudiante, aportando información valiosa en el acompañamiento que realiza el consejero académico en el marco *SATIS*.

1.3 OBJETIVO GENERAL

Desarrollar mecanismos de extracción de características que permitan enriquecer el perfil de una persona, mediante tecnologías emergentes, a partir de la observación y comprensión de su comportamiento en forma *non-intrusive*, con el propósito de prestar servicios adaptados al contexto, en el marco de un campus inteligente.

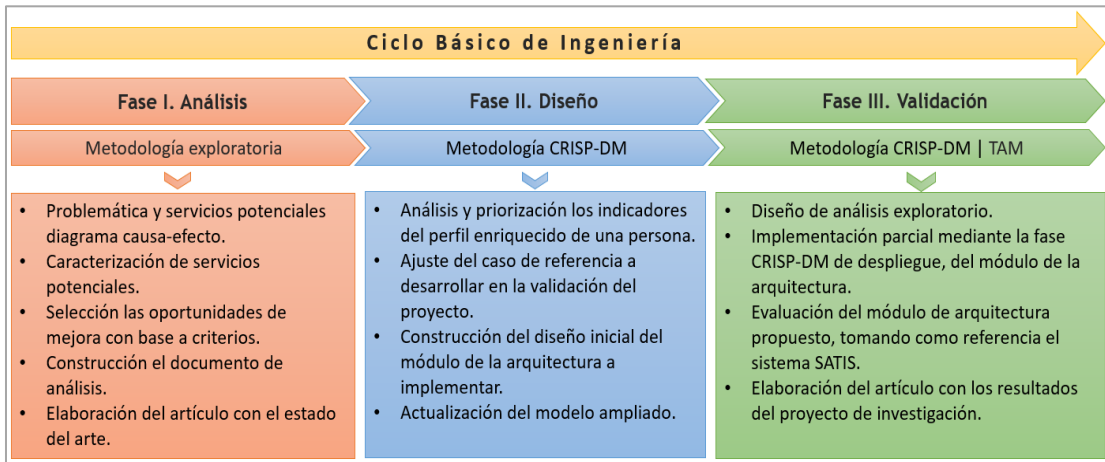
1.4 OBJETIVOS ESPECÍFICOS

1. Seleccionar un servicio en el marco de un campus inteligente, que permita el desarrollo de un modelo, a partir del análisis de tecnologías emergentes, para enriquecer el perfil de una persona.
2. Diseñar un módulo de la arquitectura, que amplíe los indicadores incluidos en el perfil de adaptación, a partir del análisis del comportamiento de una persona de forma *non-intrusive*, en el marco de un campus inteligente.
3. Validar el módulo de la arquitectura orientada al enriquecimiento del perfil, mediante un análisis exploratorio que integre los componentes y tecnologías inteligentes seleccionados, en el caso de aplicación deserción académica en la Pontificia Universidad Javeriana.

1.5 FASES DE DESARROLLO

La metodología se fundamenta en el ciclo básico de ingeniería, compuesto por análisis, diseño y validación, como se detalla en la figura 1. Cada fase se relaciona de forma directa con los objetivos específicos planteados. Se propone un caso de referencia transversales a todas las fases. El caso de referencia es definido a partir del análisis del estado del arte.

Figura 1. Fases de Desarrollo Muysca



Fuente. Elaboración propia.

1.5.1 Fase I: Analizar

Para desarrollar la fase de análisis se sigue la metodología de investigación exploratoria, basada en la revisión del estado del arte y en la revisión con expertos en los temas relacionados con el trabajo de grado: *CamIn*, tecnologías emergentes, oportunidades de enriquecimiento del perfil del estudiante a partir de su comportamiento, sistemas adaptados al contexto e información del caso de referencia con aplicaciones a corto plazo (sistemas reactivos) y a largo plazo (herramientas *AD*).

Posteriormente, se realiza un cuadro comparativo, en el cual se han caracterizado los aportes de cada trabajo revisado. Se establece un proceso de evaluación sistemática de la bibliografía, a partir de criterios definidos en esta fase. Dentro de los criterios más relevantes, se encuentran las metodologías, herramientas, aspectos abordados, entre otros. Por cada una de las publicaciones o trabajos, se efectúa la correspondiente profundización y establecerá su aplicabilidad para el proyecto, realizando una valoración cuantitativa.

Adicionalmente, en esta fase se toma como apoyo para validar la investigación de las fuentes de datos y mecanismos relacionados con el perfil de un estudiante, que permitan complementar la información existente (datos transaccionales, demográficos y académicos) para aportar a los servicios de consejería académica del proyecto institucional *SATIS* de la PUJ. La presente fase, tiene el siguiente conjunto de actividades y entregables.

- Identificación y análisis de la problemática y servicios potenciales en el estado del arte, con la herramienta diagrama causa-efecto.
- Caracterización de servicios potenciales en el marco de un *CamIn*, facilitados a partir del perfil de una persona, mediante tecnologías emergentes.
- Selección las oportunidades de mejora de la propuesta de investigación, con base a criterios que se definan en la primera fase.
- Construcción del documento de análisis a partir de las oportunidades de mejora seleccionadas y caracterizadas.

E. Elaboración del artículo con el estado del arte, en relación con la propuesta de investigación. El artículo será sometido a una revista de carácter científico.

Entregables de la fase de análisis:

- A. Documento de análisis del estado del arte.
- B. Documento de análisis y caracterización de oportunidades y servicios identificados.
- C. Artículo que presenta el estado del arte.

1.5.2 Fase II: Diseñar

La fase de diseño se desarrolla mediante la metodología *CRISP-DM* con el fin de abarcar el problema de investigación; a través de las fases entendimiento del negocio, comprensión de los datos, preparación de los datos y modelado. Los datos utilizados en el trabajo de grado son facilitados por la Dirección de Tecnologías de Información (*DTI*) y por la Biblioteca Alfonso Borrero Cabal, S.J. de la PUJ, teniendo en cuenta que los datos personales o sensibles están anonimizados.

Los datos considerados provienen de tecnologías *Wifi*, radiofrecuencia (*RFID*) y datos transaccionales. Los datos de tipo transaccional se consultan de las bases de datos de la biblioteca y del sistema académico. Por otro lado, se valida la coherencia y aplicación de las soluciones identificadas, en relación con el caso con criterios de evaluación que se definan en esta fase, y en relación con el caso de referencia seleccionado, que permita seleccionar una solución o varias soluciones integradas al problema de investigación planteado.

El entendimiento del contexto y la exploración iterativa sobre el conjunto de datos permite conocer los datos, establecer un conjunto de pasos para la preparación y limpieza de los datos para la conformación de modelos que permiten clasificar o establecer patrones de comportamiento para el enriquecimiento del perfil. Posteriormente, se listan tanto los descriptores para la obtención de características, como los modelos que pueden incluirse en la arquitectura. La validación iterativa del diseño del módulo y los nuevos indicadores a partir del enriquecimiento del perfil se realiza en contraste con los indicadores actuales del sistema *SATIS*. La fase de diseño tiene el siguiente conjunto de actividades y entregables.

- A. Análisis y priorización los indicadores del perfil enriquecido de una persona, a partir de la primera fase.
- B. Ajuste del caso de referencia a desarrollar en la validación del proyecto.
- C. Construcción del diseño inicial del módulo de la arquitectura, para ampliar los indicadores incluidos en el perfil enriquecido, a partir del análisis *non-intrusive* del comportamiento de una persona.
- D. Actualización del modelo ampliado.

Entregables de la fase de diseño:

- A. Documento de indicadores del perfil enriquecido de una persona.
- B. Documento de diseño del módulo de la arquitectura.

1.5.3 Fase III: Validar

En la fase validación, se realiza una evaluación de los modelos dentro del módulo, tomando como referencia la etapa de validación y despliegue de la metodología *CRISP-DM*. Adicionalmente, la evaluación se apoya mediante encuestas con expertos, inspirado en la metodología Modelo de Aceptación de Tecnología (*TAM*).

Para realizar la validación de esta fase, se define y aplica un análisis exploratorio a partir de los datos de *Wifi*, Biblioteca (*RFID* y sistema transaccional) y sistema académico (horario de clase). Los resultados y servicios propuestos son validados con expertos. Se toma un curso del programa con mayor deserción de la Facultad de Ingeniería y otro curso con menor deserción de la Universidad. La fase de validación tiene el siguiente conjunto de actividades y entregables.

- A. Diseño del análisis exploratorio, que integre la evaluación del módulo de arquitectura propuesto.
- B. Implementación parcial mediante la fase *CRISP-DM* de despliegue, del módulo de la arquitectura orientada al enriquecimiento del perfil de una persona.
- C. Evaluación del módulo de arquitectura propuesto, tomando como referencia el sistema *SATIS* de la Pontificia Universidad Javeriana, inspirado la metodología *TAM*.
- D. Elaboración del artículo con el estado del arte, en relación con el diseño, la validación y los resultados del proyecto de investigación. El artículo será sometido a una revista de carácter científico.

Entregable fase de validación:

- A. Documento del análisis exploratorio (*CRISP-DM*).
- B. Módulo de la arquitectura (código fuente, instaladores, plan de las pruebas de funcionamiento).
- C. Documento de análisis de resultados de validación con expertos para el caso de referencia.
- D. Artículo que presenta el diseño, la validación y los resultados del proyecto.

2. MARCO TEÓRICO / ESTADO DEL ARTE

Este capítulo implementa la primera fase de desarrollo propuesta en el numeral 1.5.1 del presente documento, asociado a la revisión y análisis del estado del arte. En la primera fase, se lleva a cabo la metodología exploratoria para el marco teórico y para el estado del arte. Se considera información científica de bases de datos bibliográficas, así como información institucional, casos de aplicación, modelos, sistemas y metodologías existentes, para entender y definir marcos de trabajo que sean útiles para el proyecto. También se realiza un análisis de ventajas y desventajas de la información seleccionada, con base a criterios definidos en esta fase, que permiten obtener y documentar los elementos claves para el trabajo de grado, a partir de la evaluación del estado del arte.

2.1 MARCO TEÓRICO

Hernández, Fernández y Baptista [30] plantean que en el marco teórico “la función más importante de una teoría es explicar: decir por qué, cómo y cuándo ocurre un fenómeno”. En este sentido, se revisan conceptos asociados con el trabajo de grado, como *CamIn*, servicios soportados en localización, servicios adaptados al contexto, comportamiento o perfiles de una persona o contexto.

Campus Inteligente (CamIn)

El término *Campus Inteligente* (también llamado *Smart Campus*, *Smart University Campus*, *Universidad Inteligente* o *Smart University*) se deriva del término *ciudad inteligente*. Un *CamIn* plantea objetivos de mejora en la calidad de vida de las personas, sostenibilidad, administración eficiente de recursos, diseño y planeación de los servicios y bienestar e interacción de las personas [1]. En un *CamIn*, se propone el diseño de servicios adaptados al contexto para lograr los objetivos planteados. Por ejemplo, monitoreo en el consumo de recursos impulsado por las tecnologías de información y comunicación (TIC) [31], renovación de espacios de clase de acuerdo con el aforo [32], monitoreo de salud de las personas [33], entre otros.

Dentro de los conceptos relacionados se encuentra el término *inteligente*, el cual hace referencia en este trabajo al modelo de sistema inteligente, que incorpora sistemas de automatización, facilitando un sofisticado monitoreo y control sobre las funciones de un componente, a través de dispositivos electrónicos [34]. Este paradigma se relaciona con el término *sistemas inteligentes*, compuesto por materiales y tecnologías integrados, que permiten superar los requisitos de diseño tradicional. Los sistemas inteligentes no sólo realizan acciones programadas, también puedes aprender sobre el entorno y los personas, adaptándose a los cambios del contexto [35].

En la tabla 1 se describen algunos conceptos de *CamIn*. Para esta investigación se utilizará el enfoque de Hirsch & Ng, dada la orientación a servicios inteligentes de la universidad, considerando servicios personalizados para los usuarios, en diferentes dimensiones como aprendizaje, salud, administración de recursos, medio ambiente y gobierno.

Tabla 1. Definiciones de Campus Inteligente (CamIn)

Autor	Año	Concepto CamIn
Kaneko, Sugino, Suzuki & Ishijima [35]	2000	Un <i>smart campus</i> integra la sociedad, las universidades y la industria, en un entorno web, un lugar donde la cooperación entre las universidades y la industria puede ayudar al establecimiento de nuevos negocios. También acumula y sistematiza el conocimiento adquirido durante este proceso para desarrollar y multiplicar las necesidades de recursos humanos.
Hirsch & Ng [36]	2011	Una <i>universidad inteligente</i> es un lugar, donde se comparte el conocimiento de forma continua a través de tecnologías emergentes, con características ambientales, robusto, personalizado, responsable, interactivo y adaptativo, accesible desde cualquier momento y dispositivo, adaptándose a las características situacionales. Los pilares de un campus inteligente son <i>iLearning, iManagement, iSocial, iHealth, iGovernance</i> y <i>iGreen</i>
Uskov, Bakken, Howlett, & Jain [37]	2018	<i>Smart University</i> representa una integración de tecnologías inteligentes, funciones inteligentes, sistemas de software y hardware inteligentes, pedagogía inteligente, planes de estudio inteligentes, aprendizaje inteligente y análisis académico, y diversas ramas de la informática y la ingeniería informática.
Min-Allah & Alrashed [38]	2020	Un <i>campus inteligente</i> se considera como la integración de la informática en la nube y el <i>IoT</i> , que ayuda en la gestión, docencia, investigación y otras actividades de las universidades. Un <i>campus inteligente</i> se adhiere a los conceptos de ciudades inteligentes y hace frente a los mismos desafíos.

Tecnologías Emergentes

Los servicios diseñados dentro del contexto de un *campus inteligente* se soportan en *tecnologías emergentes*, las cuales permiten la adquisición de datos en tiempo real, con el fin de controlar, optimizar y predecir diferentes acciones en diferentes aplicaciones interconectadas [36]. Tecnologías como *IoT*, las redes inalámbricas de sensores, los paradigmas de computación en la nube, en el borde o en la niebla y el big data, permiten la conexión de objetos y personas en los cambios y los servicios futuros. Estas tecnologías permiten cumplir objetivos transversales planteados en la visión de un *CamIn*.

Algunas tecnologías consideradas en un *CamIn* incluyen *RFID, IoT*, computación en la nube, tecnología de visualización 3D (realidad aumentada), tecnología de sensores (movimiento, temperatura, luz, humedad, entre otros), tecnología móvil (incluye *NFC*, código *QR* o *GPS*) y servicios web [37]. Sin embargo, se requiere de otras tecnologías que permitan conectar dispositivos *IoT* como *Bluetooth, LoRa, SigFox, Wifi, Zigbee* o *4GLTE*, con diferentes requisitos de consumo de energía y velocidades de datos.

Servicios Basados en Localización

La localización es un elemento fundamental en un *CamIn*. Conocer la ubicación de los seres humanos y los objetos es una parte importante del contexto. Especialmente la necesidad de información de ubicación de humanos, sistemas de ayuda móviles y objetos de uso diario en áreas habitables [38]. Los servicios basados en la localización (*LBS*) son aplicaciones informáticas (específicamente, aplicaciones informáticas móviles) que proporcionan información dependiendo de la ubicación del dispositivo y de la persona, principalmente a través de dispositivos portátiles móviles (por ejemplo, teléfonos inteligentes) y redes móviles [39].

Existen tecnologías que permiten localizar a una persona en entornos exteriores e interiores. Hay varios sistemas de localización disponibles para entornos exteriores como *GSM/UMTS* o *GPS* [40]. En cuanto a sistemas de localización en interiores, se utilizan tecnologías como sensores inalámbricos, Bluetooth, cámaras, acelerómetro del sistema microelectromecánico (*MEMS*), giroscopio *MEMS*, radio identificación de frecuencia (*RFID*), banda ultra ancha (*UWB*) o detección de luz y rango (*LiDAR*) [41] y puntos de acceso (*AP*) *Wifi* como sistemas no intrusivos, ubicuos y baratos que no requieren la instalación de costosos equipos de hardware [42]. *Wifi* es adecuado para inferir la ubicación analizando y midiendo las propiedades de la señal, por ejemplo, utilizando *RSSI* (indicador de intensidad de señal recibida).

Servicios Adaptados al Contexto

Los servicios adaptados al contexto (*context-aware services*) se derivan del concepto *conciencia del contexto*, siendo un componente esencial en el diseño de inteligencia ambiental y computación inteligente. De esta manera, es posible reconocer las situaciones en las que la persona espera que se provean ciertos servicios adaptados, mejorando la usabilidad y la interacción cotidiana, facilitada a partir de tecnologías emergentes [43]. El servicio adaptado al contexto puede considerar el perfil de la persona o el contexto gustos, preferencias o hábitos [44]. Los servicios adaptados basado en otros aspectos incluye características como la percepción del servicio por parte del persona y demás características de los sistemas adaptativos complejos, que se analizan mediante tecnologías emergentes [45] y considera preocupaciones sobre cuestiones éticas y legales de las tecnologías de hogares inteligentes con respecto a la privacidad, la seguridad y la confidencialidad debido a la naturaleza altamente identificable de los datos [46].

Comportamiento y Hábitos

Según la Real Academia Española [47], el comportamiento denota “la manera de comportarse”, actuando de una manera determinada. El comportamiento se puede ejecutar deliberadamente para lograr objetivos específicos. Con la repetición, tal comportamiento puede volverse habitual y no depender de acciones-resultados generando hábitos [48]. Un hábito es un modo especial de proceder o conducirse adquirido por repetición de actos iguales o semejantes, u originado por tendencias instintivas [49]. Las expresiones de los estados humanos pueden ser involuntarias (por ejemplo, actividades, emociones) o intencionales (por ejemplo, comandos verbales y gestos con la mano / cuerpo) y pueden ser captadas mediante tecnologías emergentes.

Muysca considera el término *no intrusivo* (*non-intrusive*) como la capacidad de observar y captar una acción de una persona de forma transparente, mediante tecnologías emergentes [50]. A continuación, se enuncian algunas técnicas empleadas para la detección de hábitos de una persona, a partir de la localización [51].

- Clustering. La agrupación de datos tiene como entrada datos sin etiquetar. Cada algoritmo de agrupación viene en dos variantes: una clase, que implementa el método *fit* para aprender las agrupaciones en los datos entrenamiento, y una función, que, dados los datos de entrenamiento, devuelve una matriz de etiquetas enteras correspondientes a los diferentes grupos. Para la clase, las etiquetas sobre los datos de entrenamiento se pueden encontrar en el atributo *labels_*.

- DBSCAN [52]. Los clústeres se observan como áreas de alta densidad separadas por áreas de baja densidad. Debido a esta vista bastante genérica, los grupos encontrados por *DBSCAN* pueden tener cualquier forma, a diferencia de *k-means*, que asume que los grupos tienen forma convexa. El componente central del DBSCAN es el concepto de muestras centrales, que son muestras que se encuentran en áreas de alta densidad. Por lo tanto, un grupo es un conjunto de muestras centrales, cada una cercana entre sí (medida por alguna medida de distancia) y un conjunto de muestras no centrales que están cerca de una muestra central. Hay dos parámetros para el algoritmo, *min_samples* y *eps*, que definen formalmente a qué nos referimos cuando decimos denso. Más alto *min_samples* o más bajo *eps* indican mayor densidad necesaria para formar un grupo. Un clúster es un conjunto de muestras centrales que se pueden construir tomando de forma recursiva una muestra central
- KMEANS. Agrupa los datos tratando de separar las muestras en *n* grupos de igual varianza, minimizando un criterio conocido como inercia o suma de cuadrados dentro del grupo (ver más abajo). Este algoritmo requiere que se especifique el número de clústeres. En la tabla 2 se analizan algunos aspectos de los algoritmos de clustering más usados en la actualidad.

Tabla 2. Análisis de algunos algoritmos de clustering

Nombre del método	Parámetros	Escalabilidad	Caso de uso	Geometría (métrica utilizada)
K-Means	número de clústeres	Muy grande (<i>n_samples</i>), mediano (<i>n_clusters</i>) con código MiniBatch	Uso general, tamaño de clúster uniforme, geometría plana, no tantos clústeres	Distancias entre puntos
Affinity propagation	damping, sample preference	No escalable con <i>n_samples</i>	Muchos grupos, tamaño de grupo desigual, geometría no plana	Distancia del gráfico (por ejemplo, gráfico del vecino más cercano)
Mean-shift	bandwidth	No escalable con <i>n_samples</i>	Muchos grupos, tamaño de grupo desigual, geometría no plana	Distancias entre puntos
Spectral clustering	número de clústeres	Mediano (<i>n_samples</i>), pequeño (<i>n_clusters</i>)	Pocos grupos, incluso tamaño de grupo, geometría no plana	Distancia del gráfico (por ejemplo, gráfico del vecino más cercano)
Ward hierarchical clustering	número de clústeres o umbral de distancia	Grande (<i>n_samples</i> , <i>n_clusters</i>)	Muchos clústeres, posiblemente limitaciones de conectividad	Distancias entre puntos
Agglomerative clustering	número de clústeres o umbral de distancia, linkage type, distancia	Grande (<i>n_samples</i> , <i>n_clusters</i>)	Muchos conglomerados, posiblemente limitaciones de conectividad, distancias no euclidianas	Cualquier distancia por pares
DBSCAN	neighborhood size	Muy grande (<i>n_samples</i>), mediano (<i>n_clusters</i>)	Geometría no plana, tamaños de grupos desiguales	Distancias entre puntos más cercanos
OPTICS	minimum cluster membership	Muy grande (<i>n_samples</i>), grande (<i>n_clusters</i>)	Geometría no plana, tamaños de grupos desiguales, densidad de grupos variable	Distancias entre puntos
Mezclas gaussianas	muchos	No escalable	Geometría plana, buena para estimar la densidad	Distancia de Mahalanobis a los centros
Birch	factor de ramificación branching factor), umbral, optional global clusterer	Grande (<i>n_clusters</i> , <i>n_samples</i>)	Gran conjunto de datos, eliminación de valores atípicos, reducción de datos.	Distancia euclidiana entre puntos

Fuente. *Clustering methods sklearn* [51].

Perfil de una Persona

Wood, & Paine describen el perfil como un conjunto de características específicas que describen a una persona [53]. A continuación, describen algunos aspectos asociados al perfil de una persona.

- Datos básicos. Asociados a los datos que no cambian frecuentemente [5]. Los datos como nombre, lugar de residencia, estado civil hacen parte de este aspecto.
- Gustos. Corresponde a los sentidos que le permiten a una persona percibir el mundo, mediante el gusto, la vista, el olfato, el oído y el tacto [54].
- Preferencias. Corresponden a las actividades que se decide realizar en un momento determinado, facilitando la personalización de los servicios [55].
- Intereses. Hacen referencia al aspecto del perfil de una persona, que varía constantemente, donde se aprecia un cambio constante de los intereses, dependiendo de las condiciones del entorno [56].
- Características socio culturales. Asociadas a las creencias, valores y significados de los símbolos que tiene una persona, como parte de un grupo o comunidad. La cultura puede ser considerada como marco de referencia que influye en la forma de pensar, percibir, interpretar cosas en un entorno social [57].
- Personalidad. Define el aspecto del perfil en el que una persona tiene aptitud a nuevas experiencias, realiza acciones bajo su responsabilidad, desarrolla acciones amables con otras personas y mantiene la estabilidad emocional en las acciones cotidianas [58].
- Estilos de aprendizaje. Se refiere a las características de una persona para adquirir nuevo conocimiento, de forma activa, reflexiva, teórico o pragmático [59].
- Inteligencias múltiples. Forma en que una persona puede expresar sus habilidades de forma lingüístico verbal, lógico matemática, espacial, corporal kinestésica, musical, interpersonal, intrapersonal o naturalista [60].
- Aspecto psicológico. Define la dimensión emocional, estados de ánimo y ajuste interpersonal de la persona [61].

Perfil de Contexto

El contexto del negocio hace referencia a la representación del contexto, en la cual se puede realizar la adaptación de acuerdo con la variabilidad del entorno. La estrategia de adaptación se puede apoyar en modelos arquitectónicos, favoreciendo la confiabilidad, el rendimiento y la seguridad [62]. Para diseñar sistemas que incluyan características situacionales implícitas, el concepto contexto también debe ser analizado. El contexto es cualquier información que se pueda utilizar para caracterizar la situación de una entidad. Una entidad es una persona, lugar, servicio de software u objeto que se considera relevante para la interacción entre una persona y una aplicación [63]. Por otro lado, Dey [64] definió la conciencia del contexto se define como el uso del contexto para proporcionar información y/o servicios relevantes a la persona, donde la relevancia depende de la tarea de la persona.

2.2 ESTADO DEL ARTE

La revisión del estado del arte se lleva a cabo mediante la investigación de tipo exploratoria y mediante una revisión sistemática de la literatura (SLR) [65]. El proceso de SLR se realizó en tres etapas. En la primera etapa se seleccionan las fuentes y ecuaciones de búsqueda de la información. La estrategia utilizada en cuanto a la búsqueda de investigaciones y trabajos previos está asociada a fuentes de información, como bases de datos bibliográficas, descubridores académicos y repositorios institucionales de producción intelectual. Posteriormente, en la segunda etapa se definieron los criterios para seleccionar la información, asociados al problema de investigación. Los criterios de selección se definen en dos niveles. En el primer nivel los criterios son novedad, impacto y aplicabilidad. En el segundo nivel, los criterios son diversidad y aplicabilidad en de los servicios en el marco de un *CamIn*. Finalmente, en la tercera etapa, se realizó la evaluación de la calidad de los artículos seleccionados, aplicando los criterios de primer y segundo nivel. Ver detalle en tabla 3.

Tabla 3. Criterios de Evaluación Sistemática de la Bibliografía

Nivel	Criterio	Descripción	Indicadores	Modo de Evaluación
NIVEL 1. Se evalúan 16 servicios potenciales, resultado de la revisión de la literatura	Criterio 1-1 NOVEDAD	Nuevos aspectos que representan oportunidades de investigación, con base en la revisión del estado del arte	Publicaciones y bases de datos bibliográficas con alto impacto	ALTO. Si aplica de cuatro a cinco
			Publicaciones diversas y significativas	MEDIO. Si aplica de dos a tres
			Casos de éxito descritos en la revisión bibliográfica. El estado del arte relaciona soluciones similares a las del proyecto	BAJO. Si aplica de uno
	Criterio 1-2 IMPACTO	Asociado a la relevancia de servicios asociados a un campus inteligente, los cuales consideran el perfil de una persona	Publicaciones relacionadas con la investigación exploratoria	N/A. Si nada
			Nuevas formas de solucionar la problemática de la investigación.	
			Pocos trabajos con antecedentes	
			Considera el perfil enriquecido de la persona	ALTO. Si aplica de cuatro a cinco
			Servicio alineado con la visión del campus inteligente de la PUI	MEDIO. Si aplica de dos a tres
			La cobertura del proyecto beneficia a un número potencial de usuarios	BAJO. Si aplica de uno
Criterio 1-3 APLICABILIDAD	Considerada desde el punto de vista de la complejidad y retos en la implementación de los servicios adaptados, basados en el perfil de una persona	El proyecto tiene cobertura en diferentes zonas de la Universidad	N/A Si nada	
		El proyecto permite investigar en comunicación con diferentes actores o dependencias de la Universidad		
		Disponibilidad de los datos asociados al servicio	ALTO. Si aplica de cuatro a cinco	
		Factibilidad de gestión de los sensores para obtener los datos (simulación, bases de datos internas y externas, nuevos mecanismos, entre otros)	MEDIO. Si aplica de dos a tres	
		Existen tecnologías y paradigmas que permiten integrar las características de la idea	BAJO. Si aplica de uno	
		Es viable la gestión de recursos (humano, tecnológico, entre otros) para implementar el servicio	N/A. Si nada	
NIVEL 2. Se evalúan 9 servicios potenciales con los puntajes más altos, resultados de la evaluación del nivel 1	Criterio 2-1 DIVERSIDAD	Distinción entre los servicios potenciales en el contexto de un campus inteligente	Acompañamiento de Recurso Humano en la organización	ALTO. Si aplica a 4
			Caso diverso de servicios adaptados al contexto de un campus inteligente	MEDIO. Si aplica de dos a tres
			El caso de servicio adaptado utiliza características del perfil de una persona	BAJO. Si aplica de uno
	Criterio 2-2 APLICABILIDAD EN UN CAMPUS INTELIGENTE	Aplicabilidad en casos aterrizados y relación con los pilares de un campus inteligente.	Caso que utiliza tecnologías emergentes	N/A. Si nada
			Servicio potencial relacionado con el pilar Wellbeing	ALTO Si aplica de cuatro a cinco
			Servicio potencial relacionado con el pilar Knowledge	MEDIO. Si aplica a tres
			Servicio potencial relacionado con el pilar Environment	BAJO. Si aplica de uno o dos
			Servicio potencial relacionado con el pilar Governance	N/A. Si nada
			Servicio potencial relacionado con el pilar Society	

Fuente. Elaboración propia. El texto en color verde en la columna Modo de evaluación, indica el valor deseable del criterio.

2.2.1 Caracterización de Servicios Potenciales en un Campus Inteligente

Se encontraron dieciséis (16) grupos de servicios potenciales, de acuerdo con la revisión del estado del arte sobre los servicios adaptados al contexto en un *CamIn* y publicados como producto de la investigación *Muysca* [66]. Los servicios analizados se basan en la persona, en la

construcción de grandes centros de datos, considerando las características y cambios del entorno, con el fin de realizar la mejor adaptación posible asociada a cada usuario y contexto del campus universitario. Los grupos de servicios están asociados con virtualización de servicios, recomendaciones de contenido, detección y tratamiento de eventos anómalos, basados en Agentes e *IoT*, servicios que aportan a la toma de decisiones

Basados en Sistemas Inteligentes de Tutoría (ITS) e Inteligencia Artificial en Educación (AIED), basados en seguridad de la información, basados en interacción con el campus, basados en el uso del espacio, basados en calidad y satisfacción, basados en la investigación, basados en aspectos socioculturales, orientados a la ergonomía, basados en el tiempo, basados en el estado de la persona y basados en la percepción ética del servicio (se profundiza en ANEXO B)

2.2.2 Servicios Basados en Localización

Se espera que alrededor de 67 mil millones de dispositivos de *IoT* sean implementados en diferentes industrias para el año 2025, soportado en infraestructuras de sensores [67]. Otras tecnologías como RFID proporcionan sensores de bajo costo, los cuales permiten hacer el seguimiento de objetos de forma inalámbrica [68]. La localización inalámbrica permite identificar la ubicación y actividad de una persona, la cual es obtenida mediante señales, como las que facilita *Wifi* [4]. A partir de la detección de forma *non-intrusive*, de la condición, actividades y acciones mediante tecnologías emergentes, es posible aportar a la calidad de vida de una persona, mediante servicios personalizados, diseñados a partir de su interacción con el entorno.

En la tabla 4, se detalla la comparación de las características de perfiles enriquecido de diferentes autores. Algunos autores se centran en la investigación de aspectos como la localización, los gustos y preferencias, información de comportamiento a partir de datos transaccionales y no transaccionales (cuadro color amarillo); otros autores investigan aspectos del perfil enriquecido a partir de datos de sensores (recuadro azul); otros autores investigan características del perfil a partir de datos transaccionales (rectángulo color café); y finalmente el trabajo de grado Muysca se focaliza en el uso de datos transaccionales y no transaccionales (rectángulo color verde), para enriquecer el perfil a partir de la localización, observando de forma *non-intrusive* gustos, preferencias y hábitos que permitan obtener características del perfil en el marco de un *CamIn*.

Tabla 4. Comparación con el estado del arte

Trabajo	Perfil enriquecido del usuario	Servicios basados en localización	Servicios basados en Gustos y preferencias	Servicios basados en el comportamiento (hábitos)	Servicios basados en datos transaccionales
Díaz et al. (2008)	X	X			
Macías, Suarez y Lloret (2013)		X			
Khalid, Faisal y Rozaini (2014)		X			
Akpakwu et al. (2018)		X			
Mehdi et al. (2017)		X			
Huang y Gartner (2018)	X	X			
Perera et al. (2014)			X		
Adomavicius y Dietmar (2014)			X		
Ibarra, Augusto y Clark (2016)			X		
Preeja y Shivosubramani (2019)	X		X		
Li et al. (2015)				X	
Poslad et al. (2015)	X			X	
Takbiri et al. (2018)	X			X	
Villegas, Palacios y Luján(2019)				X	
Kang, Han y Kwon (2019)					X
Clark et al. (2016)					X
Roman, López y Mambo (2018)					X
Stoopa y Chiolerio (2014)					X
Muysca (Arias y González, 2020)	X	X	X	X	X

Fuente. Elaboración propia a partir del estado del arte. Convenciones: ✓ si aplica, X si no aplica.

2.2.3 Perfil de Usuario y Contexto

En la literatura, diferentes autores han tratado los aspectos asociados al perfil de una persona: datos básicos, gustos, preferencias, intereses, contexto de negocio, socio cultural, personalidad, estilo de aprendizaje, inteligencias múltiples, aspectos psicológicos, entre otros. Ver tabla 5.

Tabla 5. Análisis del Estado del Arte para Perfil de Usuario y Contexto

Perfil	Aspecto	[69]	[70]	[71]	[72]	[73]	[74]	[75]	[76]	[77]	[78]
Usua- rio	Datos Básicos	X	✓	X	X	X	X	X	X	✓	✓
	Preferencias	X	✓	X	X	✓	✓	X	✓	X	✓
	Gustos	X	X	X	X	X	✓	X	X	X	✓
	Intereses	X	✓	X	X	X	✓	X	✓	X	X
	Hábitos	X	X	X	✓	✓	✓	X	✓	X	X
	Pasatiempos	X	X	X	X	X	X	X	✓	X	✓
Con- texto	Espacio Temporal	✓	✓	X	X	✓	✓	✓	✓	X	✓
	Ambiental	✓	X	✓	✓	X	X	X	X	✓	X
	Social	X	X	X	X	X	X	X	X	X	X
	Reglamentación	X	X	X	X	X	X	✓	X	X	X
	Tecnológico	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Estructural	✓	X	✓	✓	X	X	✓	X	X	X

Fuente. Elaboración propia a partir del estado del arte. Convenciones: ✓ si aplica, X si no aplica.

Las diferentes características del perfil de una persona pueden ser recolectadas mediante tecnologías emergentes e información transaccional, permitiendo la integración de nuevas herramientas, para crear mejores servicios adaptados al contexto en el marco de un *CamIn*. Las características que pueden enriquecer el perfil de una persona a partir de la identificación de una persona en el campus.

2.2.4 Inteligencia Artificial para Detección de Comportamientos y Hábitos

Se han realizado diferentes aproximaciones de acuerdo con los grupos de servicios de localización, gustos y preferencias, comportamiento y aplicabilidad en un *CamIn*. Algunos autores se centran en la detección de patrones de localización, a partir de tecnologías como *Wifi* o *RFID*. Se observan oportunidades de investigación en cuanto a la detección del comportamiento, que permita detectar gustos y preferencias de forma *non-intrusive* mediante la localización, en el marco de un *CamIn*. A continuación, se analizan algunas investigaciones sobre detección de patrones o localización en espacios interiores o exteriores.

- Punto de conexión (p). Andrade, Cancela y Gama [79] indican que un punto es la representación de la conexión de un usuario a un lugar (AP asociado) en determinada fecha y hora, desde un dispositivo móvil. $p = (\text{latitud}, \text{longitud}, \text{tiempo})$. Para *Muysca*, se considera $p = (\text{id usuario encriptado}, \text{lugar de ubicación AP}, \text{fecha y hora})$.
- Trayectoria usuario (Tr). Una trayectoria [79] es una secuencia de puntos ordenados de la forma $Tr = (p1, p2, \dots, pn)$ donde $p1$ es el punto inicial de conexión a un AP y su fecha y hora de registro es menor a los siguientes puntos, de la forma p . Para *Muysca*, se considera que la secuencia de puntos Tr corresponde a las posiciones observadas de forma consecutiva (cada 5 minutos) para un mismo usuario.
- Ubicación de permanencia (Up). Representa un punto en la que un usuario permanece más tiempo que un determinado umbral de tiempo dentro de un umbral de distancia, de acuerdo con Wind, et. al. [80]. *Muysca* considera un Up cuando existen más de tres puntos consecutivos iguales en la trayectoria de un usuario. Una Up se denota por una tupla de la forma $Up = (\text{id usuario}, \text{etiqueta de ubicación}, \text{fecha y hora de inicio}, \text{fecha y hora fin})$. *Muysca* define como una Up a aquella cuya diferencia entre fecha y hora inicial y fecha y hora final sea mayor a 5 minutos.
- Diario usuario (Du). Un diario de usuario [79] es una secuencia de Up ordenadas de la forma $Du = (Up1, Up2, \dots, Upn)$, donde $Up1$ es la primera ubicación de permanencia de un usuario en un día determinado.
- Lugar Frecuente (Li). También llamado lugar de interés. Es un conjunto de Up . Un Li representa una región que es visitada frecuentemente por un determinado usuario, definido de la forma $Li = (\text{id usuario UP}, \text{etiqueta de ubicación}, \text{fecha y hora inicio}, \text{duración conexión})$. Un Li implica que el lugar posee algún significado para el usuario. Se usa un umbral de recuento para separar las ubicaciones que se visitan con más frecuencia que las otras que tienen un significado más fuerte para el usuario. *Muysca* establece que un Li debe asociar una duración mayor a 10 minutos. Un Li en el que un usuario se conecta con cierta frecuencia en un intervalo de tiempo determinado.

- Hábito de conexión a *Wifi* (H). Según Yoo, et al. [81] un hábito es una forma más o menos fija de pensar, querer o sentir que se adquiere a través de una repetición previa. Por ejemplo, ir de casa a la Universidad a estudiar de lunes a viernes a las 08:00 a.m. puede definirse como un hábito, o ir a la biblioteca tres veces por semana aproximadamente a las 4:00 p.m. también puede ser considerado un hábito. Por lo tanto, una tarea importante para identificar hábitos es filtrar los datos con el fin de mantener trayectorias justas que pueden representar algunos movimientos significativos como estos ejemplos. H se define de la forma $H = (id\ usuario\ Up, etiqueta\ de\ ubicación, número\ de\ conexiones, tiempo\ promedio\ de\ conexión, duración\ máxima, duración\ mínima, intervalo\ [día, semana, mes], fecha\ de\ inicio\ intervalo, fecha\ de\ fin\ intervalo)$. Para *Muysca*, un hábito se caracteriza por la repetición en la conexión a *Wifi*, realizada por parte de un usuario dentro del campus en un periodo de tiempo regularmente consecutivo.

2.2.5 Inteligencia Artificial para Servicios Inteligentes en un *CamIn*

Se han realizado diferentes aproximaciones en el uso de técnicas y algoritmos del contexto de inteligencia artificial, que contribuyen al desarrollo de servicios inteligentes en un *CamIn*. Áreas como el aprendizaje de máquina los sistemas expertos, el procesamiento de lenguaje natural y visión por computadora, permiten el desarrollo y evolución de servicios como recomendaciones a partir de la ubicación de una persona e interacción con el campus, detección de emociones o anomalías mediante cámaras, entre otros. En la tabla 6, se analiza el uso de tecnologías emergentes de acuerdo con el estado del arte, en relación con las dimensiones de un *CamIn*. Se observan oportunidades en el uso de tecnologías emergentes, para el desarrollo de servicios que contribuyen a dimensiones como bienestar, conocimiento, ambiente, gobierno y sociedad.

2.2.6 Deserción Estudiantil

El Ministerio de Educación Nacional de Colombia [82] define la deserción como “el abandono del sistema escolar por parte de los estudiantes, provocado por la combinación de factores que se generan tanto al interior del sistema como en contextos de tipo social, familiar, individual y del entorno”. Las universidades han abarcado esta problemática desde diferentes enfoques.

En la Pontificia Universidad Javeriana, diferentes autores han abordado el fenómeno a través de contribuciones, por ejemplo, análisis predictivos para identificar estudiantes que presentan un alto riesgo de deserción [83], estudio de determinantes del rendimiento académico y la deserción estudiantil [84] y desarrollo de sistemas de alertas tempranas para estudiantes de primer año en la Facultad de Ingeniería [85]. Este panorama, permite explorar las alertas actuales e identificar las oportunidades de enriquecimiento del perfil de un estudiante a partir de la localización, que aporte al estudio y gestión de medidas institucionales para prevenir la deserción estudiantil.

Tabla 6. Análisis del Estado del Arte sobre Inteligencia Artificial en un CamIn

Nro.	Servicio CamIn	Tecnología Emergente			Visión de CamIn					
		IoT	AmI	Analítica de datos	Técnica IA/análítica	Wellbeing	Knowledge	Enviroment	Governance	Society
1	Virtualización de servicios	✓	✓	✓		✓	X	✓	✓	X
2	Recomendaciones de contenido	✓	✓	X	Algoritmos evolutivos	✓	X	X	✓	X
3	Detección y tratamiento de eventos anómalos	X	✓	✓	Clasificador bayesiano. Lógica difusa. Machine Learning	✓	X	✓	✓	X
4	Servicios basados en Agentes e IoT	✓	X	X	Redes neuronales. Algoritmos evolutivos. Modelos de decisión de Markov. Sistema basado en reglas. Redes bayesianas. Modelos probabilísticos y estocásticos	✓	✓	✓	✓	X
5	Servicios que aportan a la toma de decisiones	X	X	✓	Lógica difusa. Mapas autoorganizados. Redes neuronales. Aprendizaje no supervisado	X	X	X	✓	X
6	Basados en Sistemas Inteligentes	X	✓	X		✓	✓	X	✓	X
7	Basados en seguridad de la información	X	X	X	Aprendizaje de máquina. Cadenas de Markov	X	X	X	✓	X
8	Basados en interacción con el campus	✓	X	✓	Aprendizaje de máquina. Árbol de decisión. Sistemas expertos	✓	✓	X	✓	X
9	Basados en calidad y satisfacción	X	✓	✓		✓	✓	X	X	X
10	Basados en la investigación	✓	✓	X		X	X	X	X	✓
11	Basados en el contexto	✓	✓	X	Sistemas basados en reglas. Redes bayesianas	X	✓	✓	X	✓
12	Orientados a la ergonomía	X	✓	✓		X	✓	X	X	✓
13	Basados en el tiempo	✓	✓	✓		✓	X	X	X	✓
14	Basados en el estado de la persona	X	X	X	Heurística	✓	X	X	X	X
15	Basados en la percepción del servicio	X	X	X	Redes neuronales artificiales. Lógica difusa. Modelos ocultos de Markov	✓	X	✓	X	X

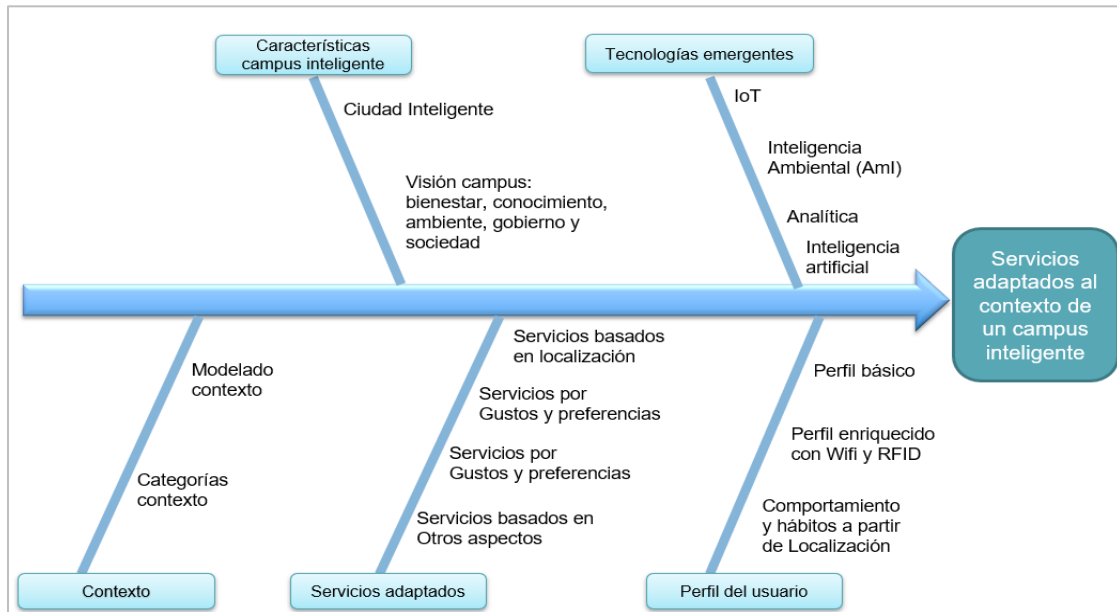
Fuente. [66].

2.2.7 Selección del Caso de Referencia

La selección del caso de referencia se realiza a partir de criterios de evaluación en dos niveles. En el primer nivel se evalúan los criterios en relación con los grupos de servicios identificados en el estado del arte para un *CamIn*, asociados a novedad, impacto, aplicabilidad, diversidad y enfoque *CamIn*. En el segundo nivel, se analizan criterios cómo diversidad y aplicabilidad. A partir de este análisis, se obtiene que el caso de referencia de *Muysca* es la detección y tratamiento de eventos anómalos. Por ejemplo, a partir de la detección de un aforo de un lugar (detectado mediante conexiones a *Wifi*), determinar si se excede la capacidad en la ocupación del lugar, notificando al personal administrativo en caso de superar el aforo. A partir del análisis del estado del arte, en la figura 2 se concluyen algunos aspectos que aportan al enriquecimiento de mecanismos del perfil de una persona en *Muysca*, a partir de la localización, en el marco de un *CamIn*.

El caso de referencia seleccionado corresponde a servicios basados en el uso del espacio, de acuerdo con características que enriquecen el perfil de una persona a partir de la localización en el contexto de un campus inteligente, aportando información al fenómeno de la deserción estudiantil.

Figura 2. Análisis causa-efecto a partir del análisis del estado del arte



Fuente. Elaboración propia

2.2.8 Consideraciones de Seguridad y Ética

En el contexto colombiano, existen diferentes leyes, decretos y lineamientos que permiten gestionar el tratamiento de datos personales, obtenidos de forma directa o mediante sensores. A continuación, se presentan algunas normativas sobre seguridad, privacidad y ética de la información. Ley estatutaria 1581 de 2012 [86]. Esta ley dicta disposiciones generales para la protección de datos personales. También existe la guía para la implementación del Principio de Responsabilidad Demostrada (*Accountability*) [87]. Hace parte del programa Integral de Gestión de Datos Personales y está dirigido a actores que deban dar cumplimiento del régimen general de Protección de Datos Personales y sean vigilados por la Superintendencia de Industria y Comercio. Finalmente, la Ley 1266 de 2008 [88] dicta disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales.

Muysca maneja los datos de forma anónima [86], validados por la Oficial de Protección de Datos PUJ. El enfoque de seguridad, privacidad y ética facilita que no se haga referencia explícita a una persona en particular. El análisis exploratorio se realiza en plataformas de hardware y software académico. Una vez que se finaliza la investigación *Muysca* en diciembre de 2020, se procede a eliminar los datos utilizados en la investigación.

3. TRABAJOS RELACIONADOS PUJ

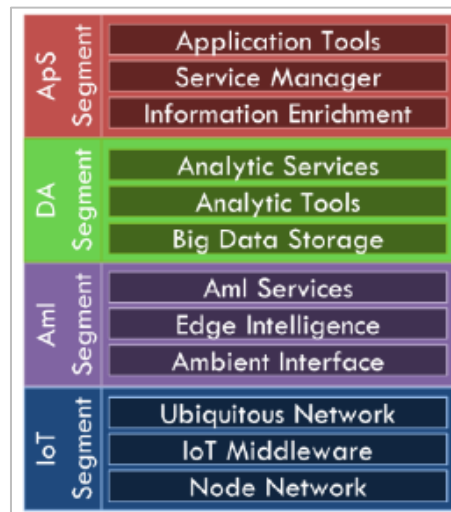
3.1 MODELO AMI-IOT *QUYSQUA*

Este trabajo de grado fue desarrollado por el Ingeniero Luis Barreto, bajo la dirección del Ingeniero Enrique González Guerrero, PhD. Este trabajo desarrolló un modelo extendido llamado *Quysqua* [27] en dos dimensiones; la primera dimensión se enfocó en el cuidado y bienestar de personas de la tercera edad y en el cuidado médico en casa de adultos mayores enfermos que necesitan atención. La segunda dimensión se enfocó en el ámbito informático, incorporando y profundizando el componente de herramientas genéricas para analítica.

Modelo *Quysqua*

El modelo *Quysqua* (palabra muisca que quiere decir cuidado y/o bienestar) se describe en la figura 3 [27]. Está definido por capas, cuatro grandes segmentos, cada uno compuesto por capas que definen las funcionalidades básicas, para implementar una solución destinada a brindar herramientas de *data analytics* (DA) de datos de asistencia y bienestar a un adulto mayor.

Figura 3. Trabajo Relacionado PUJ. Modelo *Quysqua*



Fuente. Trabajo de grado *Quysqua* [27]

El segmento *IoT* se centró en las funcionalidades de configuración de redes de nodos (sensores y actuadores), integración de tecnologías de comunicación heterogéneas, y gestión de datos (homogenizar, transformar y validar) de sensores y envío de órdenes a actuadores, como servicios ubicuos. En el segmento *AmI* consideró las capas asociadas al consumo de datos suministrados por el segmento *IoT*, aplicación de una inteligencia de borde para la toma de decisiones, y generación de servicio de envío y recepción de mensajes. El segmento *DA* se enfocó en el desarrollo de un modelo conceptual independiente de la tecnología, infraestructura o proveedor de Big Data. El componente de *Edge Intelligence*, recibe la notificación con los detalles de los datos procesados y dependiendo de la configuración de señales que reciba, presenta los

servicios AmI en modelos de comunicación *request/response* y/o *publish/subscribe* al Segmento *DA*. Con el seguimiento individual de pacientes, fue posible identificar alertas y sus calificaciones de riesgo, identificar hábitos de comportamiento, comparativos semanales de las últimas lecturas con respecto a las primeras, entre otros. Finalmente, el segmento *Aps* facilitó un conjunto de servicios adaptados de *IoT* y/o AmI y/o *DA*, visualización y presentación de la información y un conjunto de herramientas para desarrollar aplicaciones.

Validación del Modelo *Quysqua*

La validación del modelo *Quysqua* consideró cuatro pruebas. La primera, respecto al caso de referencia. Se aplicó el modelo en una dimensión física asistiendo al cuidado del ritmo cardiaco del adulto mayor; y una segunda dimensión emocional monitoreando el nivel de entrenamiento generado cuando el adulto mayor observa sus programas de televisión favoritos. La segunda, la comparación con trabajos ya reconocidos en el ámbito de *data analytics*. Se probó que el modelo *Quysqua*, consideró las funciones necesarias propuestas y validadas en esos trabajos. La tercera, la validación con expertos. Se validaron los aspectos no funcionales del modelo propuesto, que confirmó la coherencia y potencialidad de implementación. Finalmente, la cuarta validación estuvo asociado al desarrollo de un prototipo. Se validó la funcionalidad integrada de los trabajos previos PUJ. Se implementó un *sistema de agentes inteligentes (SMA)* necesarios para demostrar el componente de herramientas genéricas de análisis de datos. Se validó la aplicabilidad del modelo en los casos de referencia.

3.2 PROYECTO *SMART-UJ*

El proyecto *Smart-UJ* [28] es un proyecto liderado por la Facultad de Ingeniería de la PUJ. El proyecto considera las dimensiones de un campus inteligente. Se plantea el fortalecimiento del campus universitario, en las dimensiones de bienestar (*Wellbeing*) para vivir mejor, conocimiento (*Knowledge*) para aprender, crear e innovar, medioambiental (*Environment*) para cuidar la casa común, gobernabilidad (*Governance*) para administrar eficientemente los recursos y comunidad (*Community*) para transformar la comunidad. En el marco *Smart-UJ* se plantea la creación de equipos de trabajo con visión multidisciplinar que, frente a una necesidad planteada, ofrezcan una mirada desde el punto de vista de la necesidad puntual, como desde la tecnología que servirá como apoyo para cubrirla.

Servicios Inteligentes

Un servicio inteligente es un “servicio digital autónomo, proactivo, adaptable y personalizable, que a partir de la analítica y de información obtenida en tiempo real permiten tomar decisiones y realizar acciones en forma más oportuna” [28]. En el marco *Smart-UJ*, un servicio inteligente tiene cuatro niveles.

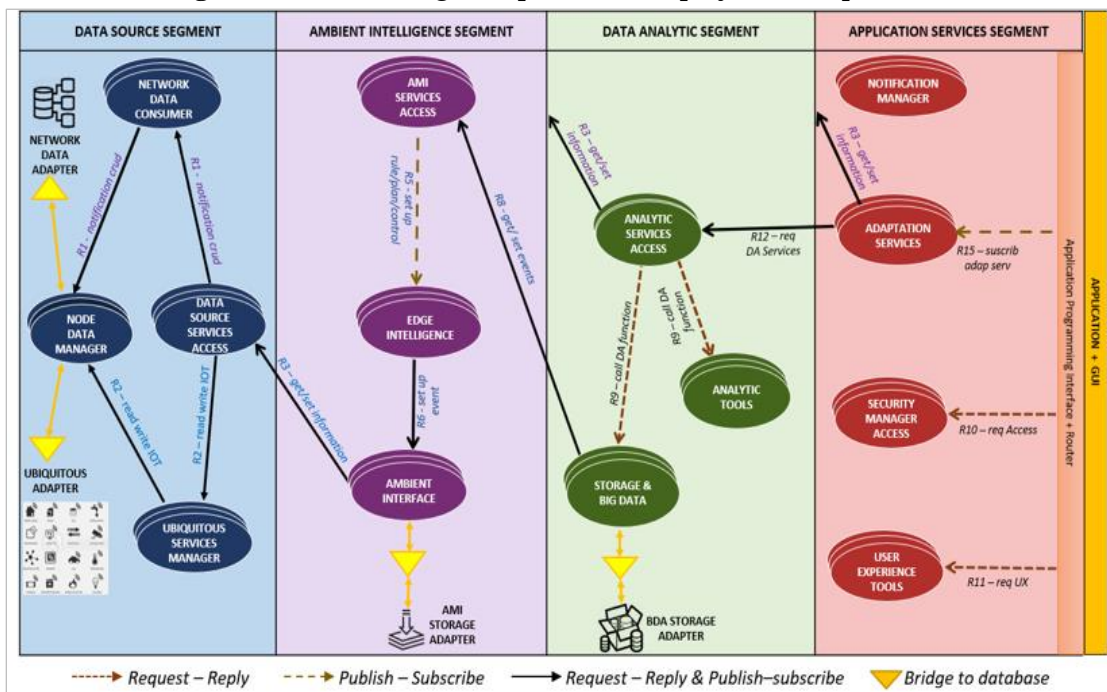
En el primer nivel, los servicios funcionan de forma reactiva en respuesta a un evento detectado en el ambiente. La respuesta a una solicitud realizada por un usuario mediante una aplicación o un sitio web interactivo. Por ejemplo, una aplicación que permite a un estudiante buscar un consejero y establecer una cita con él, se convierte en un servicio digital reactivo iniciado por la necesidad del estudiante. En el segundo nivel, se programa el sistema para que realice un monitoreo automático de condiciones programadas en el sistema, mediante la experiencia de

los usuarios. Por ejemplo, la experiencia del consejero le permite programar automáticamente reuniones de consejería con sus estudiantes, considerando épocas del semestre donde es común la solicitud de citas de consejería. En el tercer nivel, se encuentran los servicios predictivos, que incluyen las características del primer y segundo nivel también incluyen el uso de analítica predictiva. Por ejemplo, el uso de información académica de los estudiantes, que conduzca a la identificación de situaciones particulares, mediante el uso de analítica predictiva. Finalmente, en el cuarto nivel se encuentran los servicios inteligentes que adicionan a las características de los niveles inferiores, que apoyan la toma de decisiones en tiempo real, mediante el uso de diferentes fuentes de información. Por ejemplo, identificar comportamientos académicos como la asistencia a clase.

Modelo de Agentes para Construcción de Perfil de una Persona

En la figura 4, se describe el modelo de agentes construido por el equipo multidisciplinario del proyecto *Smart-UJ*. Este modelo, detalla el nivel de inteligencia que debe tener cada uno de los agentes, definidos y representados mediante óvalos. Posteriormente, se realiza el análisis de las metas asociadas a cada agente. Finalmente se definen roles e interacciones necesarias para que el servicio inteligente sea funcional: este ejemplo constituye la creación del perfil de una persona a partir de la localización captada mediante sensores como GPS y tecnologías de interconexión inalámbrica como *Wifi*.

Figura 4. Modelo de agentes para crear el perfil de una persona



Fuente. Proyecto Smart-UJ [28]

A continuación, se describen los roles e interacciones del modelo de agentes presentado en la figura 4.

1. Se activa el API *User Experience Tools (UET)*, asociado a la creación del perfil del usuario (R11).
2. Se activa el API asociado a la autenticación, para consultar o consumir la información del perfil de una persona (R10).
3. Se activa el API de suscripción a los cambios realizados en el servicio del perfil de una persona (R15).
4. Se activa el componente *Adaptation Services (AS)*, para obtener la información académica de la persona (R3, R1).
5. El componente *AS* activa el proceso de analítica, asociado a la consulta de lugares de conexión de una persona en el campus (R12).
6. El componente *Analytic Services Access (ASA)* realiza el llamado al componente *Storage & Big Data (SBD)*, solicitando la información de la consulta realizada en el paso 5 (R9).
7. El componente *SBD* obtiene los eventos relacionados con la ubicación del usuario (R8).
8. El componente *Ambient Intelligent Services Access (AMISA)* valida y activa las reglas y eventos locales del sistema, asociados al perfil de la persona solicitado (R5, R6).
9. El componente *Ambient Interface (AIN)* obtiene los datos (homogenizados) de los sensores de localización de la persona, solicitud paso 1 (R3, R2).
10. El componente *ASA* realiza el llamado al componente *Analytics tools (AT)*, solicitando la formación del perfil de la persona en cuanto a características y gustos de localización (R9).
11. El componente *AS* recibe y actualiza la información del perfil de la persona (R3, R1).
12. El API *UET* recibe y muestra la información del perfil actualizado de la persona (R11).

En la figura 5, se describe la arquitectura de alto nivel del modelo *Quysqua Smart-UJ*. El modelo consta de cuatro segmentos: *segmento Data Source (DS)*, *segmento Ambient Intelligence (AmI)*, *segmento Data Analytics (DA)* y *segmento Application Service (AS)* [23]. Si bien el modelo presenta segmentos y componentes genéricos, cada aplicación de este modelo puede hacer uso de los elementos del esquema general del modelo, de acuerdo con la naturaleza de los servicios a desarrollar, donde no es obligatorio el uso de todas las capas y componentes. En segmento *DS* se representa en color azul, segmento *AmI* se representa en color morado, segmento *DA* se representa en color verde y segmento *AS* se representa en color rojo.

Figura 5. Arquitectura de alto nivel Quysqua Smart-UJ

Application Service (AS) Segment	Notification Manager
	User Experience Tools
	Security Manager Access
	Adaptation Services
Data Analytic (DA) Segment	Analytic Services Access
	Analytic Tools
	Storage & Big Data
Ambient Intelligence (AmI) Segment	AmI Services Access
	Edge Intelligence
	Ambient Interface
Data Source (DS) Segment	Data Source Services Access
	Network Data Consumer
	Ubiquitous Services Manager
	Node Data Manager

Fuente. Proyecto Smart-UJ [28]

Segmento Data Source – DS. El segmento *DS* concentra las capas que gestionan el consumo de datos extraídos de sensores o fuentes transaccionales, conversión de datos físicos a equivalencias lógicas en el contexto de análisis e interfaz de comunicación con otros segmentos. Las capas que conforman el segmento *DS* son: *Node Data Manager*, *Ubiquitous Services Manager*, *Network Data Consumer* y *Data Source Services Access*. A continuación, se detalla el alcance de cada una de las capas del segmento DS.

- *Node Data Manager*. Capa encargada de leer, validar y almacenar los datos obtenidos de sensores y otras fuentes de información.
- *Ubiquitous Services Manager*. Capa encargada de gestionar y dar acceso a los datos provenientes de dispositivos *IoT*.
- *Network Data Consumer*. Capa encargada de realizar la equivalencia de nombres lógicos vs. ubicación Física para los datos gestionados en la capa *Node Data Manager*.
- *Data Source Services Access*. Capa encargada de filtrar la información requerida desde el segmento *Ambient Intelligence*, gestión de acceso a los datos solicitados desde la capa *AmI* y validar la seguridad para el consumo de los datos.

Segmento Ambiente Intelligence – AMI. El segmento *AmI* contiene las capas encargadas del consumo de datos provenientes del segmento *DS*, gestión de inteligencia de borde para la toma de decisiones reactivas o a largo plazo e interface de comunicación y gestión de mensajes con otros segmentos. Las capas que conforman el segmento *AmI* son: *Ambient Interface*, *Edge Intelligence* y *AmI Services Access*. A continuación se detalla el alcance de cada una de las capas del segmento *AmI*.

- *Ambient Interface*. Capa encargada de detectar y gestionar los eventos sobre los datos provenientes del segmento *Data Source Services Access*.
- *Edge Intelligence*. Capa encargada de gestionar las reglas reactivas asociadas a los eventos creados en la capa *AmI Services Access*.

- *AmI Services Access*. Capa encargada de gestionar la interacción con el usuario, gestionar mensajes y gestionar los servicios del segmento AmI mediante modelos de comunicación request/response y/o publish/subscribe (valida la seguridad para el consumo de los datos).

Segmento Data Analytic – DA. El segmento *DA* concentra las capas encargadas de la gestión de herramientas y algoritmos de analítica que permiten generar conocimiento y valor al contexto de análisis, mediante los datos provenientes de los segmentos AmI y DS. Las capas que conforman el segmento *DA* son: *Storage & Big Data*, *Analytic Tools* y *Analytic Services Access*. A continuación se detalla el alcance de cada una de las capas del segmento *DA*.

- *Storage & Big Data*. Capa encargada de consumir los servicios de los segmentos *AmI* y *DS*, y limpiar, preparar y analizar los datos iniciales requeridos por los modelos analíticos.
- *Analytic Tools*. Capa encargada de ejecutar los modelos de analítica (descriptiva, diagnóstica, predictiva, prescriptiva y cognitiva), que permitan generar conocimiento sobre el negocio; con base en los datos consumidos desde la capa *Storage & Big Data*.
- *Analytic Services Access*. Capa encargada de gestionar el acceso a los servicios *DA* por parte de la capa *adaptation services* y validar la seguridad para el consumo de la información.

Segmento Application Service – AS. El segmento *AS* contiene las capas encargadas de gestionar la adaptación de información, presentación, visualización y administración de seguridad para el acceso al consumo de información solicitada por parte de GUI y aplicaciones diversas. Las capas que conforman el segmento *DA* son: *Adaptation Services*, *Security Manager Access*, *User Experience Tools* y *Notification Manager*. A continuación se detalla el alcance de cada una de las capas del segmento *AS*.

- *Adaptation Services*. Capa encargada de realizar la construcción del perfil enriquecido de una persona y gestionar la adaptación de la información de acuerdo con el despliegue (dispositivo de acceso) y el contenido (contexto y usuario).
- *Security Manager Access*. Capa encargada de la gestión de la seguridad en el acceso al consumo de los servicios del modelo.
- *User Experience Tools*. Capa encargada de consumir y visualizar los servicios de analítica de acuerdo con la información consumida de la capa *Adaptation Services*.
- *Notification Manager*. Capa encargada de gestionar las notificaciones sobre eventos del modelo.

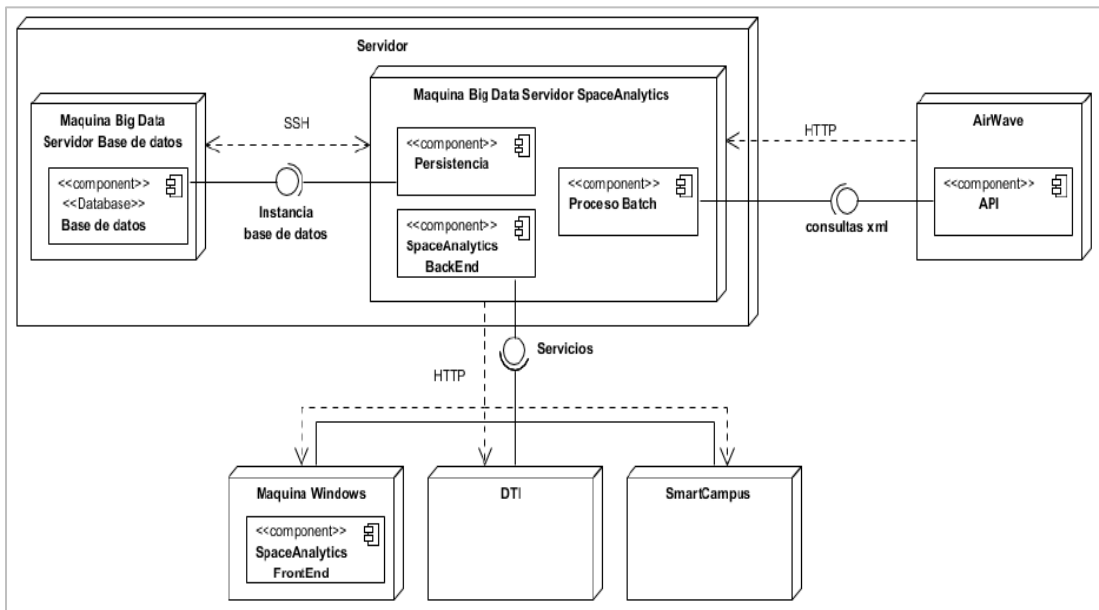
3.3 SPACE ANALYTICS

Space Analytics [89] es un trabajo de grado, desarrollado por Alejandro Torres, et al., bajo la dirección del Ingeniero Alejandro Sierra Múnera. El sistema provee información útil para la toma de decisiones en cuanto al uso de los espacios, mediante el uso de datos recolectados por la red *Wifi* de la PUJ. Se hace uso de técnicas descriptivas y predictivas de analítica, con

el fin de observar patrones y comportamientos de las personas a través del tiempo en lugares específicos.

En la figura 6, se presenta la arquitectura de *Space Analytics*. Consta de cuatro nodos, de los cuales tres pertenecen al sistema, el nodo restante, refiere a la plataforma de *Airwave* con el componente del API. En el nodo *Máquina Big Data Servidor Base de Datos* se almacenan los datos en una base de datos no relacional. En el nodo *Máquina Big Data Servidor SpaceAnalytics* se encuentra el proceso *batch* para la recolección de información y el proceso *backend* que permiten la serialización y exposición de los servicios. Finalmente, el nodo *Máquina Windows*, que gestiona los componentes *frontend* del proyecto.

Figura 6. Arquitectura trabajo relacionado Space Analytics



Fuente. Trabajo de grado *Space Analytics* [89]

Se realizó un prototipo de la interfaz de usuario para los servicios ocupación de tiempo, top de espacios y predicción de ocupación, con resultados satisfactorios. Se conoció el estado de la población que ocupa un espacio en la Biblioteca, de manera inmediata y precisa a las condiciones normales y anormales de los espacios.

3.4 IAM: MODELO INTEGRADO DE ADAPTACIÓN

El modelo Integrado de Adaptación (*IAM*) fue desarrollado por Ángela Carrillo, et al. [5]. Describe los componentes del modelo de adaptación, que considera las características del usuario, de su dispositivo de acceso y del contexto en el cual se realiza la interacción del usuario con el sistema. El modelo *IAM* considera cuatro componentes, detallados en la figura 7. El primer componente, es el módulo de presentación, que establece la información que el usuario desea ver desplegada en el dispositivo de acceso. El segundo componente, describe el módulo contextual, que detalla las características de interacción del usuario con el sistema. En el tercer componente, se describe el módulo de conexión inalámbrica, que detalla las características de

hardware y software, que define la tecnología de conexión más apropiada para que el usuario acceda mediante el dispositivo. Finalmente, el cuarto componente especifica el módulo de contenido, que facilita la definición de las características tanto de un individuo como de una comunidad.

Figura 7. Modelo de trabajo relacionado IAM



Fuente. IAM: Modelo Integrado de Adaptación [5]

Validación del modelo IAM

La validación del modelo *IAM*, se realizó a través de diferentes trabajos asociados a cada uno de los módulos propuestos. El módulo de presentación fue validado mediante el módulo de adaptación a la presentación de la información *NOMAD*, compuesto por siete capas: perfil físico, red, perfil actual, filtrado, sesión de dispositivo, sesión de usuario y servicios. El módulo de conexión inalámbrica fue validado, mediante el clasificador taxonómico de hardware, software y componente lógico mediante árboles de decisión. El módulo contextual se validó mediante el modelo *MOCA* (Módulo Contextual de Adaptación), a través de características espacio temporal, ambiental, actividad, usuario y estructural. El perfil individual de usuario se validó mediante el modelo *MAIPU* (Modelo de Adaptación de Información basado en Perfil de Usuario), que expuso las preferencias de usuario, de actividades relacionadas, de resultados, gustos, intereses e información básica. Finalmente, el módulo de perfil grupal se validó a través del modelo *MAICO* (Modelo de Adaptación de la Información para comunidades), mediante componentes como roles, hábitos, normas, políticas, preferencias sociales, entre otros.

3.5 ESCENARIOS DE DESERCIÓN ESTUDIANTIL EN LA PUJ

El trabajo de grado “Arquitectura basada en redes neuronales para identificación de escenarios de deserción estudiantil en la facultad de ingeniería de la Pontificia Universidad Javeriana”, fue desarrollado por Ceidy Álvarez [90], bajo la dirección Javier Francisco López. El trabajo permitió definir escenarios de deserción, para un programa de educación superior en una institución, que permitan apoyar el proceso de toma de decisiones en cuanto a la generación de estrategias de retención estudiantil.

Validación de la Arquitectura

La arquitectura propuesta consideró componentes de recolección de datos, procesamiento de datos y componente de aprendizaje y evaluación de algoritmo genético y red neuronal artificial para definir los determinantes de deserción. La arquitectura fue evaluada a través de seis casos específicos, para un programa académico presencial de educación superior. Se encontró un vector, dentro de una población definida a través de un algoritmo genético, con la mejor evaluación de fitness, calculada por medio de la estructura de una red neuronal definida. El análisis de la evaluación permitió reconocer los discriminantes definidos para algunos escenarios generados, dentro del conjunto de prueba.

3.6 FORTALEZAS Y APORTES DE TRABAJOS RELACIONADOS

En la tabla 7, se describe el análisis de las fortalezas de los trabajos relacionados, en cuanto a los aportes a al uso de herramientas de analítica y técnicas de inteligencia artificial.

Tabla 7. Fortaleza de trabajos relacionados

Trabajo Relacionado	FORTALEZAS	
	Conceptual	Uso herramientas de Analítica e Inteligencia Artificial
Integración de Herramienta De Analítica en el Modelo AMI-IOT QUYSQUA	Marco común, que de la posibilidad a tener un impacto mayor en la calidad de vida, bienestar y cuidado médico remoto de la persona de la tercera edad en su hogar.	Clustering
Proyecto de Planeación Universitaria SMART-UJ	Diseño de servicios inteligentes soportados en tecnologías avanzadas que transformen a la sociedad y que impacten positivamente los pilares misionales del campus PUJ.	Modelo de agentes
Space Analytics	Sistema que provee información asociada al uso de los espacios, apalancándose en los datos recolectados por la red Wifi de la PUJ, a través del tiempo en lugares específicos.	Modelo descriptivo a través de metodología CRISP-DM para análisis grupal. Modelo predictivo con red neuronal
IAM: Modelo Integrado de Adaptación	Modelo de adaptación, que considera las características del usuario, de su dispositivo de acceso y del contexto en el cual se realiza la interacción del usuario con el sistema.	Árboles de decisión
Identificación de escenarios de deserción estudiantil en la PUJ	La generación temprana de escenarios que permitan identificar las variables que influyen directamente en la deserción que son críticas para un grupo de estudiantes, y que son importantes en el desarrollo de estrategias de retención estudiantil de cualquier institución universitaria.	Uso de algoritmos genéticos y redes neuronales para identificar determinantes de deserción académica.

Fuente. Elaboración propia

A continuación, se detallan los aportes de cada trabajo, con respecto al trabajo de grado *Muysca*.

Trabajo relacionado Quysqua. (a) Modelo que integra características *Aml, IoT y data analytics*. Estos elementos, permiten observar de forma *non-intrusive*, el estado de un usuario

y facilitar servicios inteligentes en el marco de un *CamIn*. (b) Suscripción a cambios o actualizaciones en las características del perfil de una persona y heterogeneidad de los datos en el contexto de un *CamIn*.

Trabajo relacionado *Smart-UJ*. (a) Servicios Inteligentes. Diseño de servicios inteligentes en el contexto de deserción, que permita acompañar a un estudiante desde el rol de consejero académico, a partir de la identificación de un usuario en el campus. (b) Modelo de agentes. Como elemento importante en la construcción de prototipos. Este modelo, permite definir y representar las metas, roles e interacciones de los agentes asociados al servicio de enriquecimiento del perfil, a partir de la localización. En este contexto, se plantea la oportunidad de enriquecer el modelo de agentes, para las notificaciones en los cambios o actualizaciones del perfil de una persona, a partir de la localización.

Trabajo relacionado *Space Analytics*. (a) Representación de la información. Uso de Árboles para representar la estructura jerárquica de cada edificio, donde cada nodo inicialmente tiene la capacidad y los tags pertinentes a cada área que representa. Oportunidad de enriquecer los prototipos con otros lugares del campus universitario y de forma personalizada mediante el uso de nuevos datos del perfil de un estudiante. (b) Despliegue del sistema. Arquitectura y tecnologías de implementación orientada a procesos *batch* y *backend*, a partir de datos capturados mediante *Wifi*. Oportunidad de continuar el enriquecimiento del modelo, mediante otras fuentes que enriquezcan el perfil de una persona como *RFID*.

Trabajo relacionado *IAM*. (a) Construcción de perfil de usuario. Construcción del perfil de usuario, a partir de las preferencias, gustos e intereses de conexión en el campus. (b) Construcción de perfil de contexto. Características del perfil de contexto, que permitan enriquecer el perfil de un estudiante en el marco de un campus inteligente. Por ejemplo, la ubicación como parte del componente espacio temporal.

Trabajo relacionado *Arquitectura basada en redes neuronales para identificación de escenarios de deserción estudiantil en la PUJ*. (a) Contexto de deserción. Características del análisis de deserción mediante factores individuales, académicos, institucionales y socioeconómicos. (b) Enriquecimiento perfil de estudiante. A partir de información contextual, como lugares y tiempos de permanencia en el ámbito geográficos, que faciliten información para el tratamiento del fenómeno de la deserción en la PUJ. Oportunidad de enriquecer el perfil de una persona mediante el análisis de datos de programas con diferentes niveles de deserción.

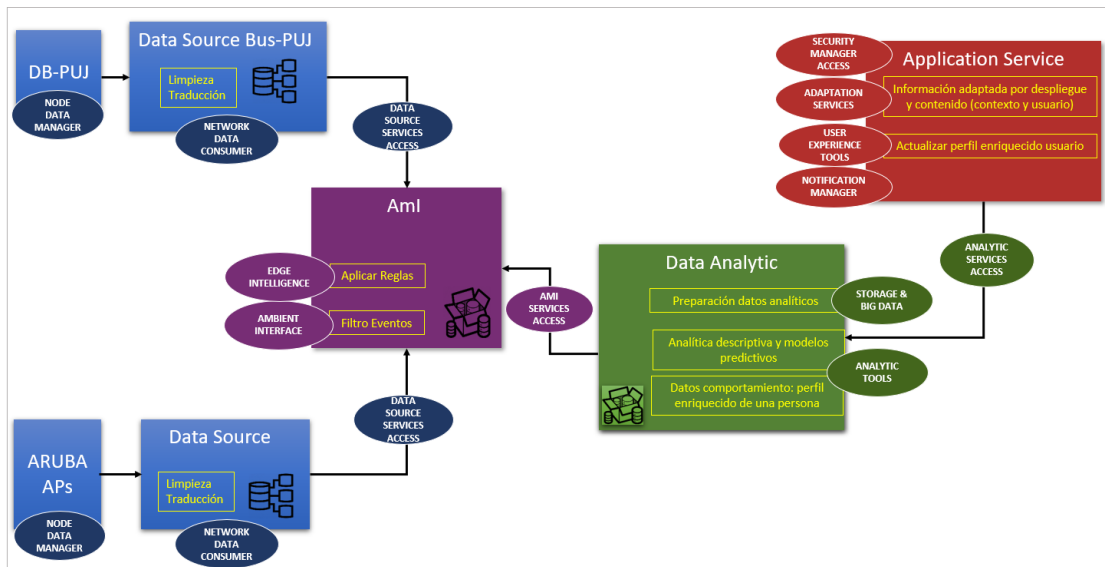
4. MODELO MUYSCA

Este capítulo implementa la segunda fase de desarrollo propuesta en el numeral 1.5.2 del presente documento, asociada al diseño de *Muysca*. Se detalla el diseño un módulo de la arquitectura, que permitirá obtener elementos que enriquecen el perfil de un estudiante, a partir de los patrones de comportamiento, como lugares frecuentes capturados con redes inalámbricas (*Wifi*) y hábitos bibliográficos tomados a partir de radiofrecuencia (*RFID*).

4.1 DESCRIPCIÓN GENERAL DEL MODELO MUYSCA

En la figura 8, se presenta la aplicación de la arquitectura de alto nivel del *Quysqua Smart-UJ* aplicada a los servicios propuestos por *Muysca* para enriquecer el perfil de una persona. Se realiza la equivalencia en código de colores. En segmento *DS* se representa en color azul, segmento *AmI* se representa en color morado, segmento *DA* se representa en color verde y segmento *AS* se representa en color rojo.

Figura 8. Aplicación del modelo *Quysqua Smart-UJ* en *Muysca*



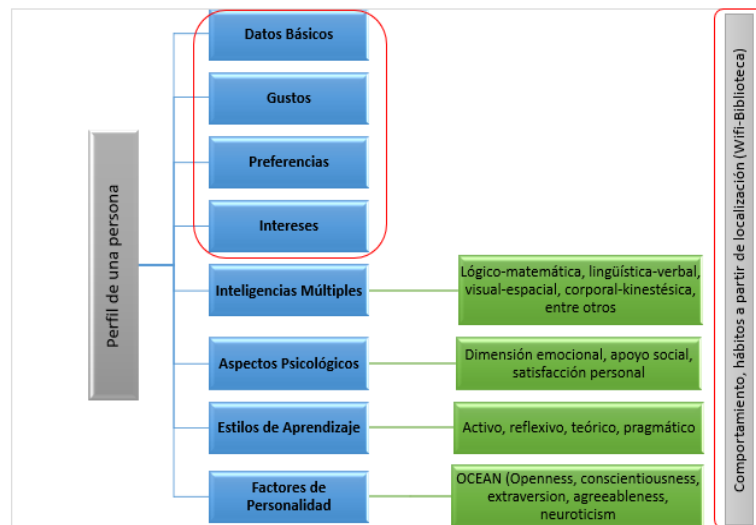
Fuente. Proyecto *Smart-UJ* [28] aplicado a *Muysca*

Se profundiza en el desarrollo y evolución de la capa *Adaptation Services*, perteneciente al segmento *AS*. En esta capa se incluyen componentes de información enriquecida de acuerdo con el modelo de adaptación *IAM*, en cuanto a los módulos de presentación, de contenido y de contexto. El modelo *Muysca*, propone la construcción del perfil enriquecido de una persona a partir de la localización en un Campus, mediante la conexión *Wifi* y mediante la observación del comportamiento bibliográfico captado mediante *RFID*.

4.2 PERFILES DE ENRIQUECIMIENTO MUYSCA

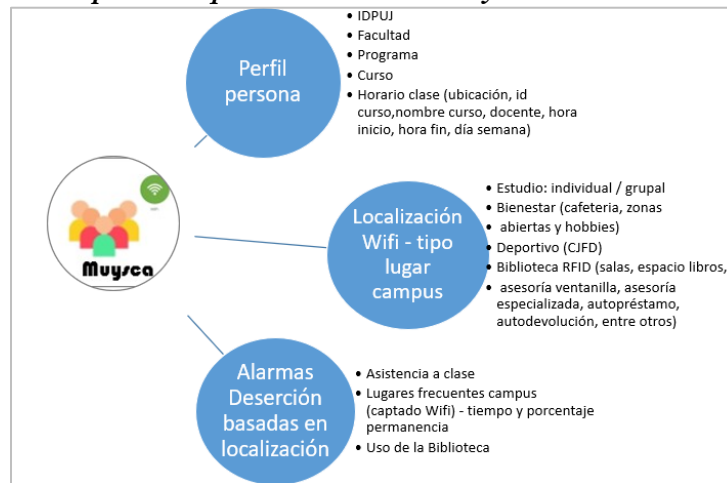
A partir del análisis del estado del arte y de los trabajos relacionados, se consideran elementos importantes, para el desarrollo del diseño y ampliación de la capa *Adaptation Services*, perteneciente al segmento *AS* del modelo de arquitectura genérico. El modelo de enriquecimiento de perfiles de usuario en *Muysca* se realiza a partir de IAM [5] y se muestra en la figura 9, considerando elementos del perfil de una persona (marcados con rectángulos rojos), como datos básicos, gustos, preferencias e intereses, localización en el campus mediante *Wifi* y *RFID* (préstamos recursos de la biblioteca) y datos académicos de PeopleSoft como el horario del estudiante.

Figura 9. Componentes del perfil enriquecido de una persona en Muysca



Fuente. Elaboración propia.

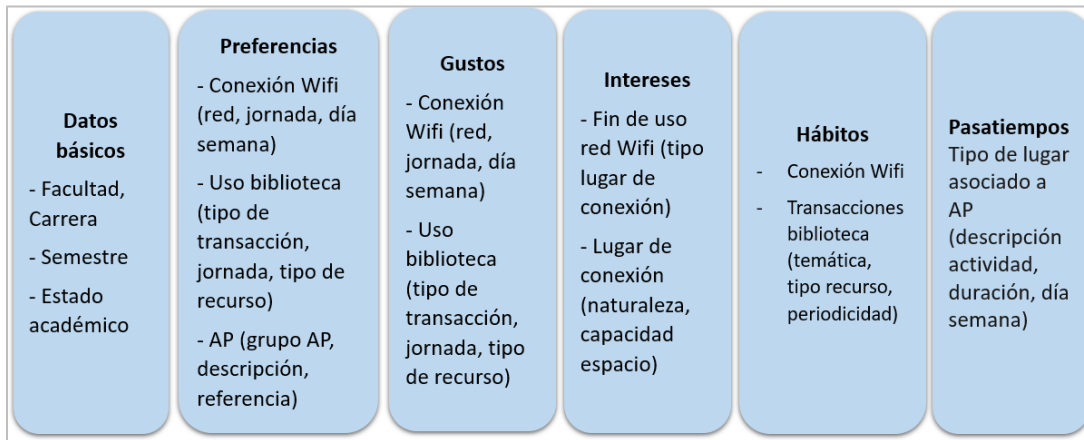
Luego de validar los datos de la investigación, se procedió a identificar los elementos que permitieran la construcción de los servicios de Muysca, como el perfil de una persona, aspectos del contexto en cuanto a localización y las alarmas de deserción importantes para complementar los servicios del caso de referencia. Los elementos que enriquecen los servicios del caso de referencia Muysca se detallan en la figura 10, en cuanto al perfil de una persona, aspectos de localización en el campus e identificación de servicios relacionados con las alarmas de deserción basadas en localización.

Figura 10. Elementos para enriquecer servicios de Muysca

Fuente. Elaboración propia

4.2.1 Perfil de Usuario

En la figura 11, se describe el perfil de usuario del modelo *Muysca*. Los aspectos considerados son datos básicos, preferencias, gustos, intereses, hábitos y pasatiempos en cuanto a las conexiones a *Wifi* y transacciones registradas mediante *RFID* en la biblioteca general, por parte de un estudiante.

Figura 11. Perfil de usuario del modelo Muysca

Fuente. Elaboración propia

4.2.2 Perfil de Contexto

En la figura 12, se describe el perfil de contexto del modelo *Muysca*. Se consideran aspectos como las características espacio temporales, datos ambientales de conexión, información social, reglamentación, tecnológico y aspectos estructurales.

Figura 12. Perfil de contexto del modelo Muysca

Fuente. Elaboración propia

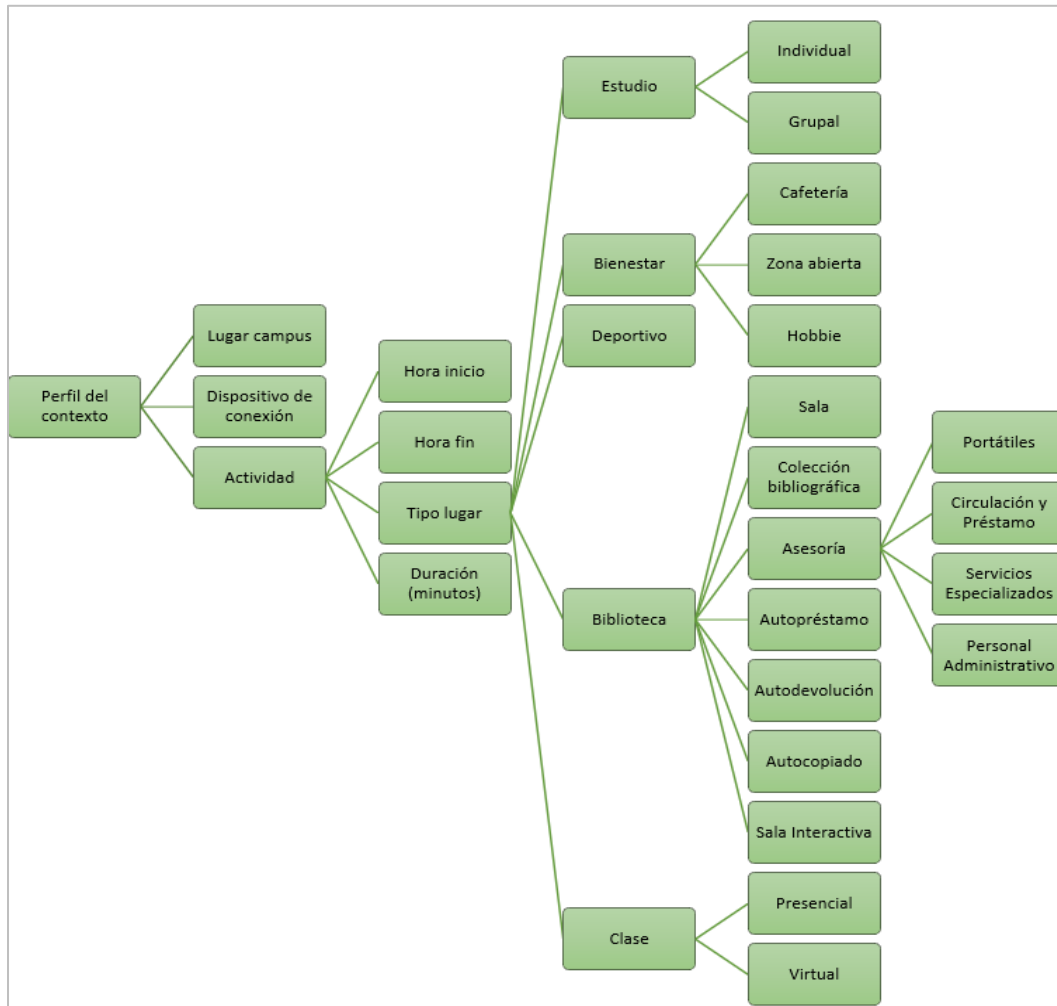
En cuanto a la representación de la información de un lugar de conexión de forma lógico, se construyó un árbol de representación de información, detallado en la figura 13. Esta taxonomía, permite identificar los lugares de permanencia de un usuario, así como el tipo de lugar, el ámbito y la dimensión de *SMART-UJ* asociada.

Los datos de la taxonomía de lugares de un usuario en el campus se obtienen a partir del análisis del estado del arte y de la validación de negocio realizadas con los equipos *SATIS* y Biblioteca General. Según Andrade, Cancela y Gama [79] existen características espacio-temporales para identificar lugares significativos que frecuentan las personas. A partir del nombre físico de un AP, se procede a realizar la equivalencia a las características relacionadas con la problemática de la investigación, por ejemplo, piso del edificio asociado, ámbito del proyecto Smart-UJ o dimensión de CamIn, como se muestra en la tabla 8.

Tabla 8. Mapeo de lugares de permanencia para el modelo Muysca

Grupo AP Airwave	Biblioteca
Nombre FÍSICO AP	55.5.5_55
ID AP	55555
LAN MAC Address AP	00:55:00:55:00:00
AP TYPE	Aruba AP 5555
Nombre LÓGICO Lugar Campus	Sala de Estudio Grupal Vásquez p4
Tipo Lugar Campus	Ambiente de Estudio Grupal
Piso Edificio	Piso 4
Nombre Edificio PUJ	Biblioteca - Ed. 3 Jesús María Fernández, S.J.
Ámbito Smart-UJ	Ambientes de Aprendizaje
Dimensión Campus Inteligente Smart-UJ	Aprendizaje

Fuente. Elaboración propia a partir de API Airwave

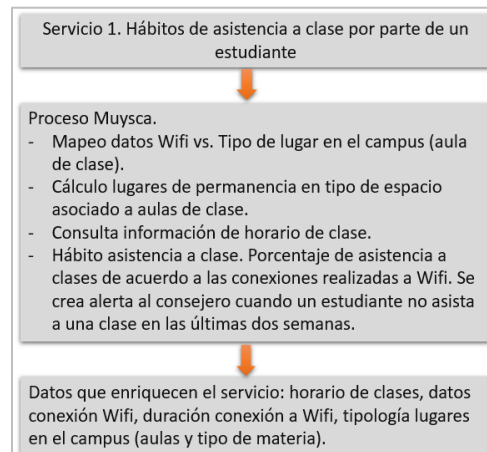
Figura 13. Representación de lugares en el campus PUJ

Fuente. Elaboración propia

4.3 SERVICIOS MUYSICA ASOCIADOS AL CASO DE REFERENCIA

De acuerdo con el caso de referencia definido en el numeral 2.2.7, se definen los siguientes servicios de enriquecimiento del perfil (SEP)

Asistencia a clase. En primer lugar, se realiza un mapeo de conexiones a AP en relación con los lugares del campus. Posteriormente, se detectan hábitos a partir de localización en lugares tipo aula clase y se compara la información con los datos transaccionales académicos para detectar el porcentaje y tiempo de conexión a AP en lugares relacionados con aulas o espacios de clase presencial, como se muestra en la figura 14.

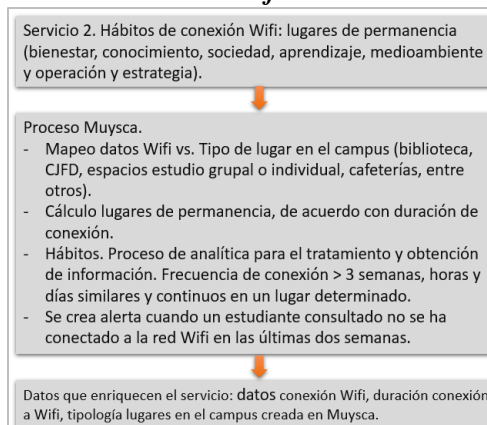
Figura 14. Descripción servicio de asistencia a clase

Fuente. Elaboración propia

A continuación, se describen los indicadores del servicio de asistencia a clase.

- Se calcula el porcentaje de tiempo de conexión (Wifi) en lugares cercanos a aulas de clase, filtrado por un periodo de tiempo.
- Minutos de conexión a Wifi según la naturaleza del lugar en Smart-UJ aula de clase.

Actividades frecuentes. Para lograr este servicio, se realiza un mapeo de conexiones a AP en relación con los lugares del campus. Posteriormente, se detectan hábitos de actividad de acuerdo con la localización (lugar y tipo), como se detalla en la figura 15.

Figura 15. Descripción servicio de actividades frecuentes

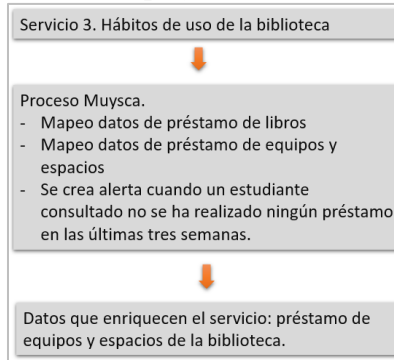
Fuente. Elaboración propia

A continuación, se describen los indicadores del servicio de actividades frecuentes.

- Se calcula el porcentaje de tiempo de conexión (Wifi) en lugares cercanos a aulas de clase, filtrado por un periodo de tiempo.
- Minutos de conexión a Wifi según la naturaleza del lugar en Smart-UJ.

Gustos y preferencias temáticos. Se identifican las materias más consultadas de acuerdo con los préstamos de material bibliográfico, equipos y espacios realizados en la Biblioteca General, como se muestra en la figura 16.

Figura 16. Descripción servicio gustos y preferencias temáticos



Fuente. Elaboración propia

El indicador del servicio de gustos y preferencias temáticos calcula el porcentaje de tiempo de conexión (Wifi) en lugares cercanos a aulas de clase, filtrado por un periodo de tiempo y número de ítems prestados en un periodo de tiempo. El sistema Muysca considera información transaccional y no transaccional para enriquecer el perfil de una persona. Los mecanismos que permiten enriquecer el perfil corresponden a la extracción de características de comportamiento y hábitos de conexiones *Wifi* en el campus, tomadas a partir de la localización en el campus mediante fuentes como *Wifi* y *RFID*.

4.4 PROCESO DE DETECCIÓN DE HÁBITOS EN MUYSKA

Un hábito se define en Muysca tomando como referencia los siguientes conceptos.

- Punto de conexión (p) [79]. $p = (id\ usuario\ encriptado, lugar\ de\ ubicación\ AP, fecha\ y\ hora)$.
- Trayectoria usuario (Tr) [79]. Para Muysca, se considera que la secuencia de puntos Tr corresponde a las posiciones observadas de forma consecutiva (cada 5 minutos) para un mismo usuario. $Tr = (p1, p2, \dots, pn)$ donde p1 es el punto inicial de conexión a Wifi.
- Ubicación de permanencia (Up) [80]. $Up = (id\ usuario, etiqueta\ de\ ubicación, fecha\ y\ hora\ de\ inicio, fecha\ y\ hora\ fin)$.
- Diario usuario (Du) [80]. $Du = (Up1, Up2, \dots, Upn)$, donde Up1 es la primera ubicación de permanencia.
- Hábito de conexión a Wifi (H) [81]. $H = (id\ usuario\ Up, etiqueta\ de\ ubicación, número\ de\ conexiones, tiempo\ promedio\ de\ conexión, duración\ conexión, intervalo\ [día, semana, mes], fecha\ de\ inicio\ intervalo, fecha\ de\ fin\ intervalo)$.

Posteriormente, se aplican los siguientes filtros para detectar un hábito a partir de la localización en el campus.

- Consulta de datos de conexiones a Wifi (incluyendo la duración de la conexión)
- Conformación de clústeres (mayor a tres repeticiones por día de la semana y jornada)
- Caracterización adicional para cada clúster. Duración promedio y desviación estándar.
- Se calcula la duración de conexión (hora inicio menos hora fin de conexión).

5. VALIDACIÓN DEL MODELO MUYSKA

Este capítulo implementa la tercera fase de desarrollo propuesta en el numeral 1.5.3 del presente documento, correspondiente a la validación de la arquitectura *Muysca*. Se valida la arquitectura propuesta mediante la implementación parcial siguiendo un análisis exploratorio mediante *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) [91].

5.1 METODOLOGÍA ANALÍTICA

La metodología *CRISP-DM* [91] establece un proyecto de minería de datos como una secuencia de fases: comprensión del negocio (fase número 1), compresión de los datos (fase número 2), preparación de los datos (fase número 3), modelado (fase número 4), evaluación (fase número 5) y despliegue (fase número 6).

5.1.1 Comprensión del Negocio

Esta tarea considera el contexto de la investigación, los objetivos de negocio y los criterios de éxito empresarial.

Objetivos de Negocio. En cuanto al TG *Muysca*, se plantea el siguiente objetivo de negocio: suministrar fuentes de datos adicionales para el enriquecimiento del perfil de un usuario, mediante datos de conexión de redes *Wifi* en el campus universitario. Criterios de éxito:

Identificar si un usuario asiste a clase, de acuerdo con la conexión realizada a *Wifi* mediante cualquier dispositivo, con el usuario encriptado del estudiante.

Identificar los lugares de interés para una muestra de usuarios de primer semestre de Ingeniería de Sistemas y de Comunicación Social, de acuerdo con la conexión realizada a *Wifi* mediante cualquier dispositivo en el campus.

Evaluación de la Situación. En el trabajo de grado *Muysca* se consideran los siguientes supuestos.

- El estudiante es quien se conecta a *Wifi*, con el usuario institucional. Posteriormente, el ID del estudiante pasa por un proceso de anonimización.
- Dado que algunos lugares del campus funcionan 24 horas, se analizará la información de lunes a domingo y de 1:00 a.m. a 11:59 p.m. La información de conexión *Wifi* es tomada del API de Aruba Airwave.
- La información de conexión de *Wifi* proviene de siete grupos de AP creados con fines de investigación, en la plataforma de gestión de *Wifi* de la Pontificia Universidad Javeriana. Estos son: primer piso Ed. 2 Fernando Barón, S.J.; sótano 1 a piso 4 Biblioteca - Ed. 3 Jesús María Fernández, S.J.; sótano y piso 1 Centro Javeriano de Formación Deportiva; cafetería piso 2 Ed. 3 Gabriel Giraldo, S.J.; piso 4 Ed. 67 José Rafael Arboleda, S.J.; piso 5 Ed. 20 Jorge Hoyos Vásquez, S.J. y Cafetería La Pecera.
- La información de conexión de *Wifi* para todos los grupos de AP se encuentra desde el 22 de noviembre de 2019 a 30 de junio de 2020. Dadas las condiciones de calendario académico, diversidad de fuentes de datos y la pandemia COVID-19, se realizará el

análisis exploratorio detallado en la presente metodología analítica desde el 1 de enero a 31 de marzo de 2020.

Objetivos de Minería de Datos. El objetivo de minería de datos corresponde a la búsqueda de clústeres de acuerdo con la conexión a *Wifi* dentro del campus universitario, detectando los lugares de interés por día de la semana y hora del día. Criterios de éxito:

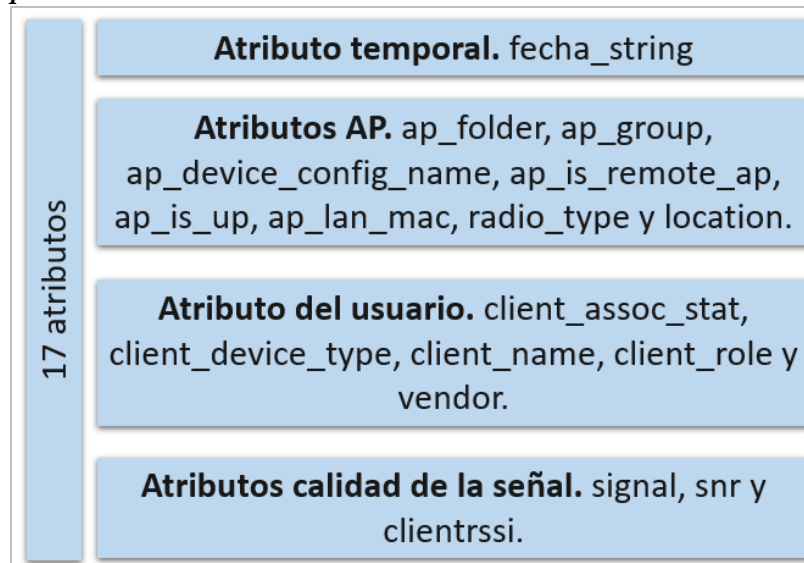
- Identificar los lugares de interés de un usuario durante la semana, horas del día y lugares del campus.
- Identificar variables en los datos que permitan enriquecer el perfil de un usuario, a partir de los datos cosechados de Wifi.

Plan de Proyecto. En el caso del TG *Muysca* el plan corresponde al calendario y alcance definido en el curso de trabajo de grado 2.

5.1.2 Comprensión de los Datos

Recopilación Inicial de Datos. En la figura 17 se detallan las características iniciales de los datos recopilados y construidos mediante el ETL de *Muysca*.

Figura 17. Recopilación inicial de los datos



Fuente. API Airwave y datos de ETL Muysca para atributo location

Descripción de los Datos. El conjunto de datos tiene 19 atributos y 6220 registros. En la tabla 9 se especifican algunos datos estadísticos como percentil, media y estándar de los valores numéricos del conjunto de datos.

Tabla 9. Descripción inicial de los datos

		Hora	Minuto	clientrsss	signal	snmp_uptime	snr
Id Estudiante	object						
Fecha Texto	object						
Hora	int64	count	6220.000000	6220.000000	6152.000000	6207.000000	6.220000e+03
Minuto	int64						6154.000000
ap_device_config_name	object	mean	13.041640	30.373473	37.823309	-57.636700	3.160822e+06
ap_group	object						37.816380
ap_lan_mac	object	std	4.052879	17.273643	10.731852	14.167929	3.423308e+06
client_device_type	object						10.737175
client_radio_mode	object	min	0.000000	0.000000	3.000000	-96.000000	0.000000e+00
client_role	object						3.000000
clientrsss	float64	25%	10.000000	15.000000	32.000000	-62.000000	3.936645e+05
place	object						32.000000
operational_mode	object	50%	13.000000	30.000000	38.000000	-55.000000	1.214851e+06
Programa	object						38.000000
radio_type	object	75%	16.000000	45.000000	46.000000	-48.000000	5.973988e+06
signal	float64						46.000000
snmp_uptime	int64	max	23.000000	59.000000	69.000000	-27.000000	1.300648e+07
snr	float64						69.000000
Hora completa texto	object						
dtype: object							

Fuente. Elaboración propia a partir de Python

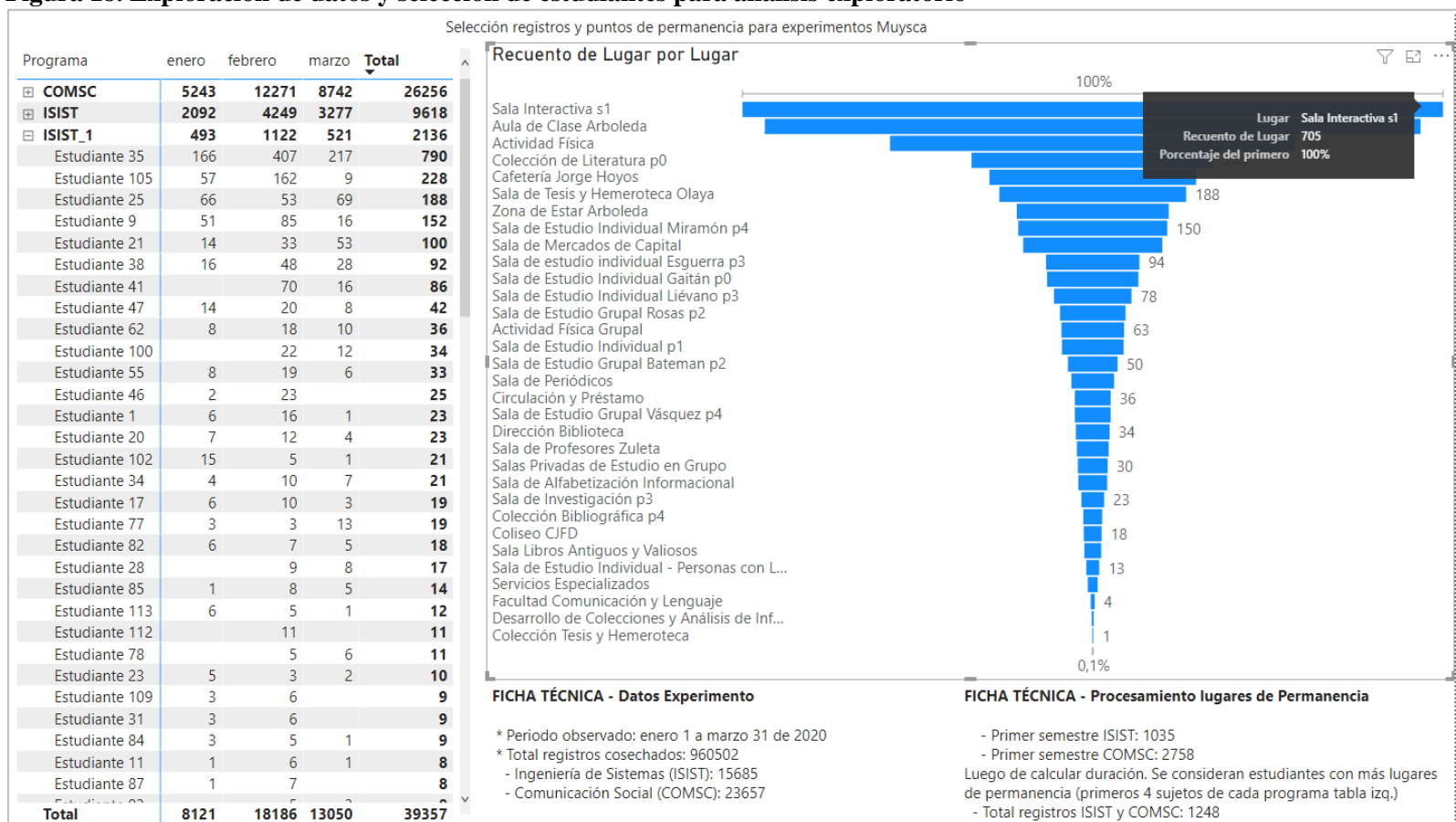
Exploración de los Datos. Se exploran los datos y se seleccionan doce estudiantes con el mayor número de conexiones a los AP (sin calcular duraciones), como se observa en la figura 18. En la tabla 10 se describen los estudiantes validados.

Tabla 10. Estudiantes seleccionados para análisis exploratorio

Usuario en Muysca-anonimizado
Estudiante 174
Estudiante 2797
Estudiante 342
Estudiante 635
Estudiante 36
Estudiante 37
Estudiante 56
Estudiante 88
Estudiante 2390
Estudiante 2820
Estudiante 3405
Estudiante 3758
Estudiante 105
Estudiante 25
Estudiante 35
Estudiante 9

Fuente. Elaboración propia a partir de datos transaccionales autorizados

Figura 18. Exploración de datos y selección de estudiantes para análisis exploratorio



Fuente. Elaboración propia a partir del servicio empresarial Power BI [92].

Verificación de la Calidad de los Datos. Al analizar el impacto de las variables en el análisis de lugares frecuentes y dado que ya se tiene la etiqueta de lugar asociado a un AP particular, las variables `clientrssi`, `signal`, `snmp_uptime` y `snr` no aportan a análisis exploratorio, puesto que no se busca ubicar un lugar a partir de estos valores.

5.1.3 Preparación de los Datos

Selección de los Datos. El análisis inicial consideró formar los clústeres por las variables `hora` y `día_semana` para, siendo `día_semana` una variable por construir a partir de la fecha. El dataset utilizado para la construcción de los modelos incluyó 6220 registros. El grupo de AP asociado corresponde a la biblioteca, para conexiones realizadas en los meses de enero, febrero y marzo de 2020.

Limpieza de los Datos. Las siguientes variables fueron retiradas del análisis, dado que son valores constantes y no representan variabilidad para el modelo: `ap_folder`, `ap_is_remote_ap`, `ap_is_up`, `radio_type`, `client_assoc_stat`, `client_radio_mode` y `operational_mode`. Las siguientes variables fueron retiradas del análisis, dado que el alcance de la investigación no considera la intensidad o ruido de la señal Wifi: `clientrssi`, `signal`, `snmp_uptime` y `snr`.

Construcción de Datos. Se tendrá en cuenta las siguientes fechas, correspondientes a días festivos de la muestra de estudio: 1 enero de 2020 – miércoles. Año Nuevo; 6 enero de 2020 - lunes. Día de los Reyes Magos y 23 marzo de 2020 – lunes. Día de San José. De acuerdo con el calendario académico e información institucional, otras fechas a considerar corresponden a: 13 a 17 de enero de 2020. Inducción de nuevos estudiantes; 20 de enero de 2020. Inicio primer periodo académico semestral y circular 006 de 2020. Suspensión de actividades académicas presenciales desde el 15 de marzo de 2020.

Se construyeron los siguientes atributos derivados, a partir de la variable `fecha_string`:

- `fecha_norm`. Fecha normalizada en formato día, mes, año. Ejemplo: 18/02/2020.
- `hora_norm`. Hora normalizada en formato hora, minutos, segundos. Ejemplo. 15:03:15.
- `mes_norm_num`. Mes normalizado en formato numérico. Ejemplo: 1 (equivale a enero).
- `mes_norm_text`. Mes normalizado en formato texto. Ejemplo: enero.
- `día_norm_num`. Día del mes normalizado en formato numérico. Ejemplo: 15.
- `día_semana`. Día de la semana normalizado en formato texto. Ejemplo: 1 (lunes).
- `num_sem_norm`. Número de la semana en formato numérico. Ejemplo: 11.

Integración de los Datos. Al validar los datos El valor mínimo varía: 0 para `hora_norm` y `minuto`; 1 para `mes_norm_num`, `día_norm_num` y `día_semana`; y 3 `num_sem_norm` (se encuentran datos a partir de la tercera semana). En los histogramas se observa que la hora con mayores conexiones a *Wifi* corresponde a las 3:00 p.m.; el mes con mayores conexiones a AP corresponde a febrero; los días del mes con mayores conexiones son los días 5, 11 y 14; el día de la semana con mayores conexiones a AP es el viernes y el número de semana en el que existen mayores conexiones es en la novena semana, observándose una asimetría en la curva de esta variable.

En cuanto a las variables categóricas se observa que el lugar con mayores conexiones a AP corresponde a la Sala Interactiva 1, seguido de la colección de literatura y la sala individual Esguerra del piso 3 de la Biblioteca General.

Formato de los Datos. En la tabla 11 se describen los formatos actualizados para los atributos.

Tabla 11. Formato de los datos para análisis exploratorio

id_estudiante	object
fecha_norm	datetime64[ns]
hora_norm	int64
minuto	int64
ap_device_config_name	object
ap_group	object
ap_lan_mac	object
client_device_type	object
client_role	object
place	object
programa	object
hora_completa	object
mes_norm_num	int64
mes_norm_text	object
dia_norm_num	int64
dia_semana	int64
num_sem_norm	int64
dtype:	object

Fuente. Elaboración propia en Python.

5.1.4 Modelado

Selección de las Técnicas de Modelado. El análisis exploratorio incluye los modelos detallados a continuación. De acuerdo con los métodos de clustering [51], se selecciona modelos que tengan escalabilidad muy grande, en donde se puedan comparar diferentes geometrías de los clústeres: KMEANS para geometría plana y DBSCAN para geometría no plana y tamaño de grupos desiguales. A continuación, se describen los modelos seleccionados.

- Modelo 1: ingenuo KMEANS.
- Modelo 2: KMEANS con k calculado con diagrama elbow.
- Modelo 3: Ingenuo DBSCAN
- Modelo 4: DBSCAN con datos escalados y un valor de épsilon calculado con el diagrama elbow.

Generación Diseño de Prueba. Inicialmente no se considera la división del dataset dado el tamaño de la muestra. La validación considera otras métricas de clustering.

Construcción de Modelos. Se ejecutan los modelos descritos en la tabla 12 para todos los datos y para los estudiantes seleccionados. Los datos se escalan con *MinMaxScaler*.

Evaluación de Modelos. Una vez seleccionado el número adecuado de clústeres y aplicado el algoritmo de clustering pertinente se tiene que evaluar la calidad de los de los mismos, de lo contrario, podrían derivarse conclusiones de agrupación que no se corresponden con la realidad. Se utiliza la validación interna *Silhouette width* [93], la cual cuantifica cómo de buena es la asignación que se ha hecho de una observación comparando su similitud con el

resto de las observaciones del mismo clúster frente a las de los otros clústeres. Para cada observación i , el *silhouette coefficient* (si) se obtiene del siguiente modo:

- Calcular la media de las distancias (ai) entre la observación i y el resto de las observaciones que pertenecen al mismo clúster. Cuanto menor sea ai mayor la similitud que tiene con el resto de las observaciones de su clúster.
- Calcular la distancia promedio entre la observación i y el resto de los clústeres. Entendiendo por distancia promedio entre i y un determinado clúster C como la media de las distancias entre i y las observaciones del clúster C .
- Identificar como bi a la menor de las distancias promedio entre i y el resto de los clústeres, es decir, la distancia al clúster más próximo (*neighbouring cluster*).

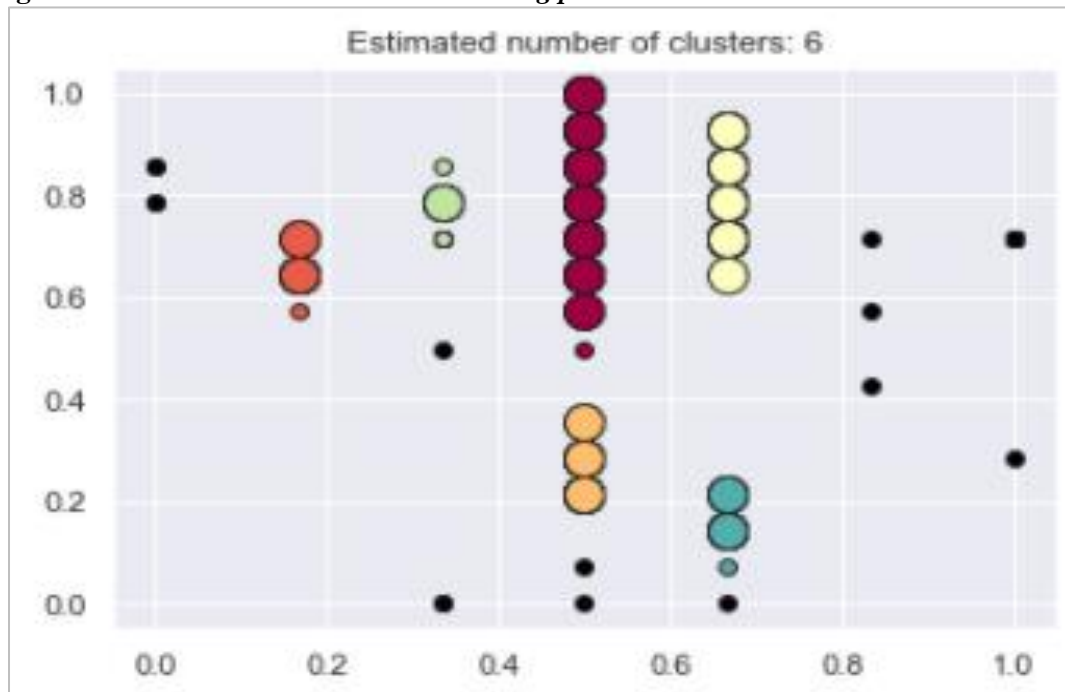
Su valor puede estar entre -1 y 1, siendo valores altos un indicativo de que la observación se ha asignado al clúster correcto. Cuando su valor es próximo a cero significa que la observación se encuentra en un punto intermedio entre dos clústeres. Valores negativos apuntan a una posible asignación incorrecta de la observación. Se trata por lo tanto de un método que permite evaluar el resultado del clustering a múltiples niveles: la calidad de asignación de cada observación por separado. Permitiendo identificar potenciales asignaciones erróneas (valores negativos de *silhouette*); la calidad de cada clúster a partir del promedio de los índices *silhouette* de todas las observaciones que lo forman. Si por ejemplo se han introducido demasiados clústeres, es muy probable que algunos de ellos tengan un valor promedio mucho menor que el resto y la calidad de la estructura de clústeres en su conjunto a partir del promedio de todos los índices *silhouette* [93].

5.1.5 Evaluación

Evaluación de Resultados. Los resultados detallados de los clústeres se evalúan y describen en el ANEXO G. El modelo que representa la diversidad y significado para el negocio es el modelo 4 DBSCAN con parámetros calculados mediante curva elbow. En la figura 19, se presentan el resultado del análisis de clustering para el Estudiante 105, de primer semestre de Ingeniería de Sistemas de la PUJ, en enero de 2020. Los clústeres formados se detallan a continuación.

- * Clúster naranja. Conexión a Wifi los martes en la tarde.
- * Clúster verde limón. Conexión a Wifi los miércoles en la tarde.
- * Clúster rojo oscuro. Conexión a Wifi los jueves en la tarde y noche.
- * Clúster naranja. Conexión a Wifi los jueves en la mañana.
- * Clúster amarillo. Conexión a Wifi los viernes en la tarde y noche.
- * Clúster verde. Conexión a Wifi los viernes en la mañana.

En el eje x se detallan los días de la semana, donde 0.0 equivale a lunes y 1.0 equivale al domingo. En el eje y se detalla las horas del día, en donde 0.0 equivale a las 00:01 a.m. y 1.0 equivale a 11:59 p.m.

Figura 19. Análisis de resultados de clustering para el estudiante 105

Fuente. Elaboración propia a partir de Python [51].

El servicio de asistencia a clase también se validó con contadores, dado que la información transaccional en relación con las conexiones realizadas a Wifi por parte de un estudiante, se pueden comparar fácilmente. Para el estudiante 105 de primer semestre de Ingeniería de Sistemas, se encontraron 7 observaciones asociadas al lugar aula de clase los días jueves después de las 4:00 p.m. En este caso, se debe validar con el horario de clase transaccional para conocer la clase, tipo de clase y naturaleza, en el caso de ser práctica o teórica, ampliando el valor para el negocio (*SATIS*).

Proceso de Revisión. En la tabla 12 se muestran los resultados de la métrica aplicado a cada modelo ejecutado. Se observan tiempos de ejecución aceptables para cada modelo. Por otro lado, se logra la caracterización de los clústeres a los que pertenece cada estudiante asociado a la investigación. El modelo 4 obtuvo en promedio 0,5729 en la métrica *Silhouette*, indicando que la mayoría de las observaciones se asignaron al clúster correcto. A diferencia del algoritmo KMEANS, DBSCAN no necesita de la especificación del número de clústeres deseado.

Tabla 12. Evaluación de modelos ejecutados

Usuario	Programa_5 emestre	Número registros (1 ene-21 mar 2020)	Modelo 1. KMEANS Ingenio (n_clusters)	Modelo 2. KMEANS (n_clusters elbow)	Modelo 3. DBSCAN Ingenio (eps cluster noise)	Modelo 4. DBSCAN (eps elbow cluster noise)	silhouett e_score Modelo 1	silhouette score Modelo 2	silhouette _score Modelo 3	silhouett e_score Modelo 4	Tiempo ejecución Modelo 1 (seg)	Tiempo ejecución Modelo 2 (seg)	Tiempo ejecución Modelo 3 (seg)	Tiempo ejecución Modelo 4 (seg)
Estudiante 105	ISIST_1	228	5	4	eps=0.11 6 21	eps=0.09 6 21	0.483	0.504	0.451	0.451	3	3	3	3
Estudiante 25	ISIST_1	188	5	4	eps=0.11 7 26	eps=0.10 7 30	0.545	0.537	0.522	0.518	3	3	3	3
Estudiante 35	ISIST_1	790	5	4	eps=0.11 10 2	eps=0.20 5 0	0.535	0.501	0.591	0.591	3	3	3	3
Estudiante 9	ISIST_1	152	5	4	eps=0.11 3 46	eps=0.08 3 28	0.606	0.564	0.390	0.462	3	3	3	3
Estudiante 2390	ISIST	477	5	4	eps=0.11 6 16	eps=0.08 6 16	0.602	0.553	0.484	0.484	3	3	3	3
Estudiante 2820	ISIST	334	5	3	eps=0.11 4 42	eps=0.08 4 42	0.618	0.774	0.697	0.697	3	3	3	3
Estudiante 3405	ISIST	782	5	4	eps=0.11 6 8	eps=0.18 6 8	0.536	0.524	0.266	0.266	3	3	3	3
Estudiante 3758	ISIST	321	5	3	eps=0.11 5 24	eps=0.08 5 24	0.808	0.730	0.751	0.751	3	3	3	3
Estudiante 36	COMSC_1	113	5	3	eps=0.11 4 26	eps=0.25 3 3	0.744	0.693	0.663	0.719	3	3	3	3
Estudiante 37	COMSC_1	213	5	3	eps=0.11 6 12	eps=0.23 6 8	0.623	0.625	0.634	0.640	3	3	3	3
Estudiante 56	COMSC_1	170	5	3	eps=0.11 4 14	eps=0.08 4 14	0.854	0.858	0.845	0.845	3	3	3	3
Estudiante 88	COMSC_1	129	5	4	eps=0.11 4 24	eps=0.25 3 24	0.728	0.663	0.660	0.581	3	3	3	3
Estudiante 174	COMSC	674	5	4	eps=0.11 9 8	eps=0.08 9 0	0.527	0.557	0.429	0.430	3	3	3	3
Estudiante 2797	COMSC	694	5	3	eps=0.11 5 17	eps=0.05 5 28	0.739	0.754	0.705	0.704	3	3	3	3
Estudiante 342	COMSC	596	5	3	eps=0.11 6 18	eps=0.10 6 18	0.710	0.680	0.716	0.716	3	3	3	3
Estudiante 635	COMSC	472	5	3	eps=0.11 5 12	eps=0.09 5 12	0.587	0.578	0.585	0.585	3	3	3	3
Todos los estudiantes		6220	5	6	eps=0.11	eps=0.048	0.422	0.427	0.295	0.299	3	3	3	3
<i>Promedio</i>			5	4			0,6275	0,6189	0,5696	0,5729	3	3	3	3

Fuente. Elaboración propia.

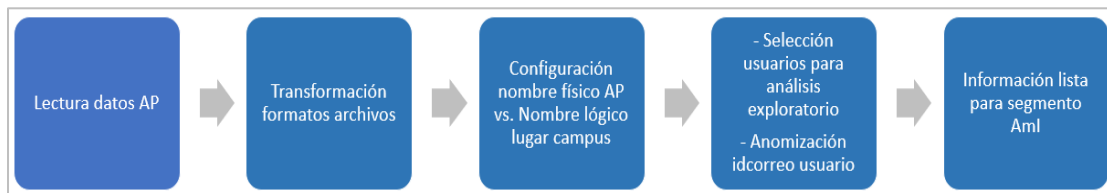
Determinación de los Sigüientes Pasos. Los siguientes pasos consideran el uso de la información de clústeres obtenida para enriquecer el perfil de una persona en el marco de un campus inteligente, a través de los servicios planteados por *Muysca*.

5.1.6 Implementación

A continuación, se detallan los componentes implementados tomando como referencia el marco conceptual de la arquitectura *Quysqua Smart-UJ*.

- En la capa *Adaptation Services*, perteneciente al segmento *AS* en cuando a la construcción del perfil de usuario y de contexto de una persona a partir de la localización en un Campus, mediante la conexión *Wifi* o *RFID*. Ver figura 20.
- Componentes *Node Data Manager* y *Network Data Consumer* de segmento *DS*. Esta implementación consideró el mapeo de lugares del campus en relación con el tipo de lugar y dimensión de *CamIn*, como se muestra en la figura 20.

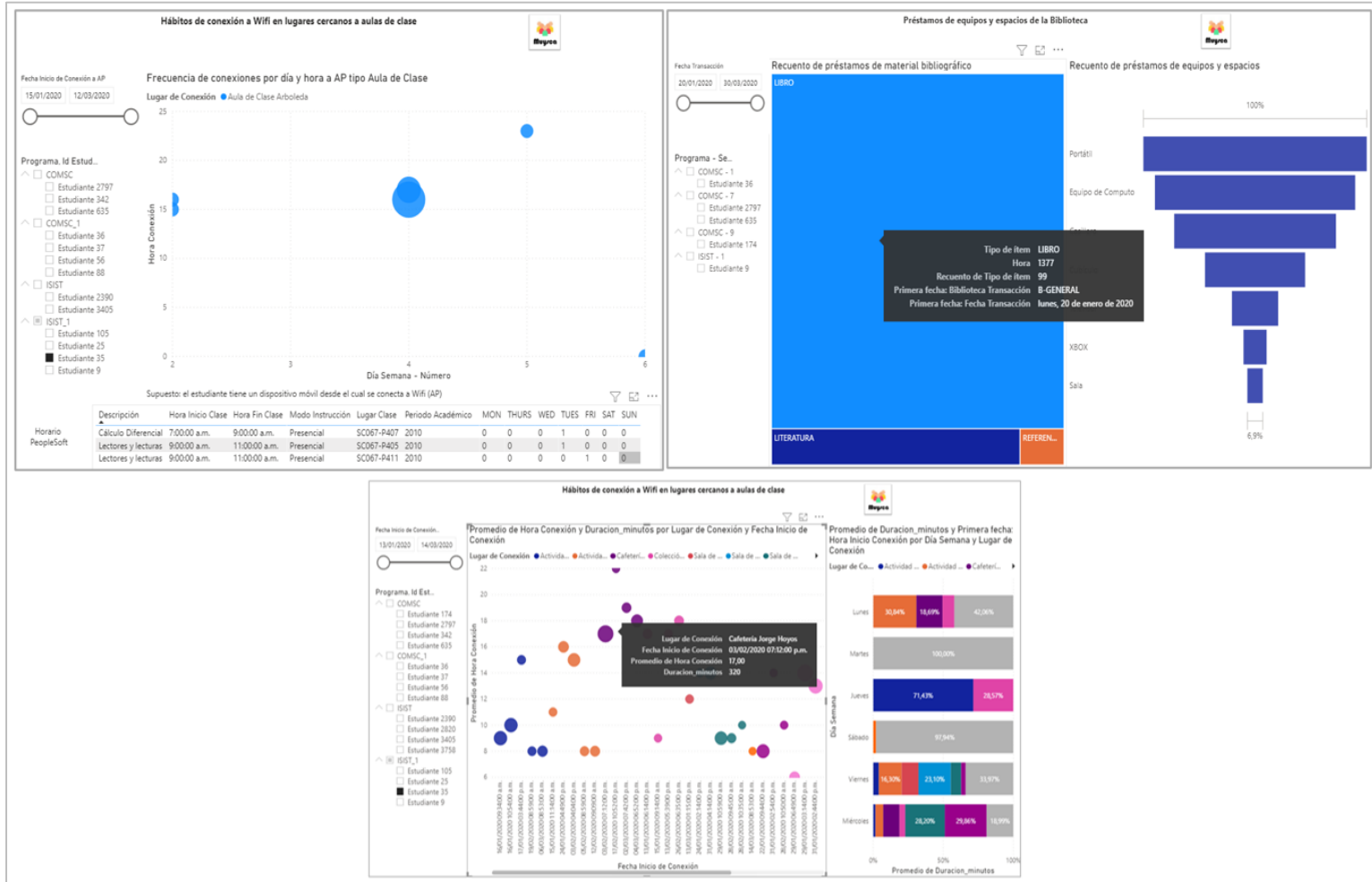
Figura 20. Descripción implementación componentes segmento DS



Fuente. Elaboración propia en herramienta Pentaho Data Integration.

- Consumo de los servicios de *Muysca* en la capa *User Experience Tools* del segmento *AS*, desde la aplicación *PowerBI*, gestionando el acceso a la información mediante niveles de seguridad y autenticación con *LDAP*, como se detalla en la figura 21.

Figura 21. Validación de resultado - caso de referencia Muysca



• Fuente. Elaboración propia en Power BI.

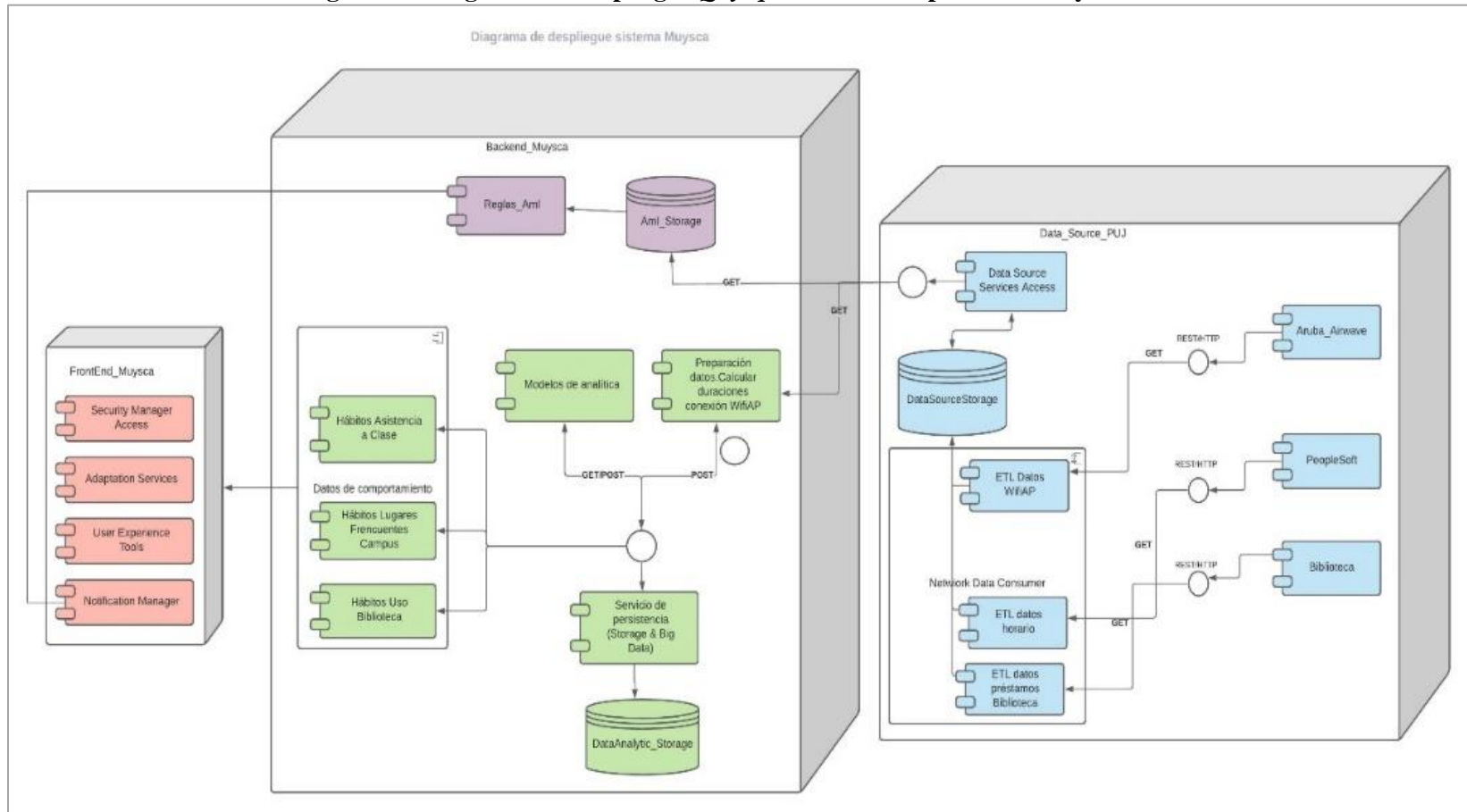
- Análisis exploratorio en la capa *Analytic Tools* del segmento *DA*. Análisis de clustering mediante DBSCAN por día de la semana y hora del día asociado a la conexión de un AP *Wifi* por parte de una persona. Ver título 5.1.4.

5.2 PROTOTIPO

En la figura 22 se representa el diagrama de despliegue considerado para aplicar la arquitectura *Quysqua Smart-UJ* a los servicios *Muysca*. La implementación parcial de *Muysca*, profundiza en el desarrollo de los siguientes componentes:

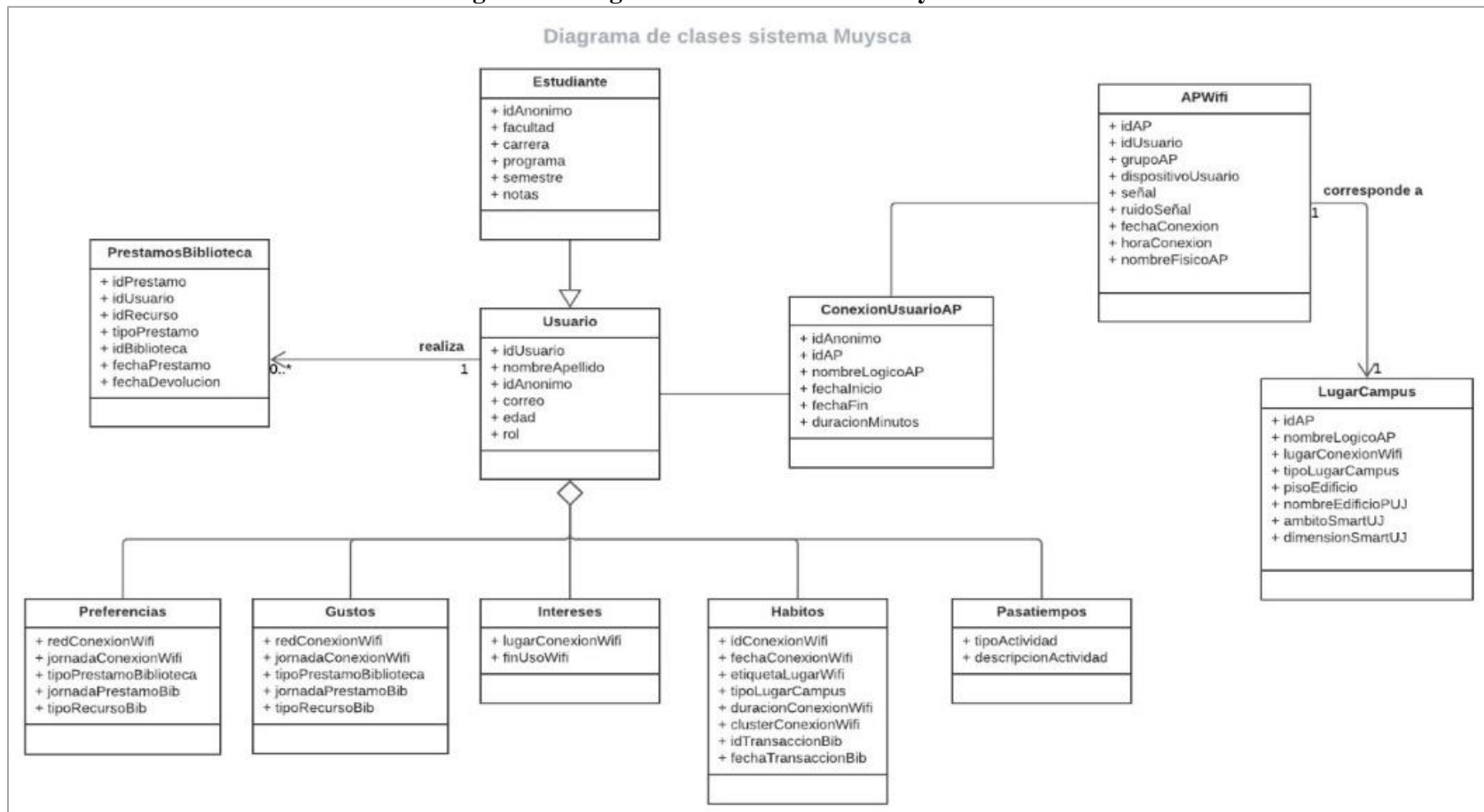
- En la capa *Adaptation Services*, perteneciente al segmento *AS* en cuando a la construcción del perfil de usuario y de contexto de una persona a partir de la localización en un Campus, mediante la conexión *Wifi* o *RFID*.
- Componentes *Node Data Manager* y *Network Data Consumer* de segmento *DS* contruidos en *Pentaho Data Integration*. Esta implementación consideró el mapeo de lugares del campus en relación con el tipo de lugar y dimensión de *CamIn*. Se utilizó el diagrama de clases para representar y modelar el perfil de usuario y de contexto de *Muysca* (ver figura 23).
- Consumo de los servicios de *Muysca* en la capa *User Experience Tools* del segmento *AS*, desde la aplicación *PowerBI*, gestionando el acceso a la información mediante niveles de seguridad y autenticación con *LDAP*.
- Análisis exploratorio en la capa *Analytic Tools* del segmento *DA*. Análisis de clustering mediante DBSCAN por día de la semana y hora del día asociado a la conexión de un AP *Wifi* por parte de una persona. El análisis se realizó mediante la librería *scikit-learn* del lenguaje Python.

Figura 22. Diagrama de despliegue *Quysqua Smart-UJ* aplicado a *Muysca*



Fuente. Modelo Quysqua Smart-UJ aplicado al modelo Muysca.

Figura 23. Diagrama de clases sistema Muysca



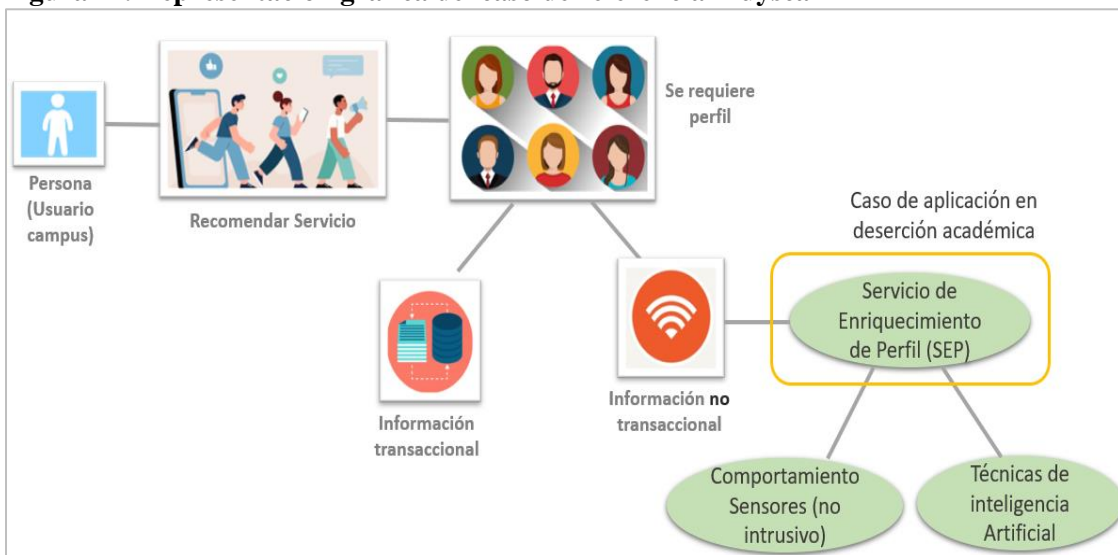
Fuente. Elaboración propia.

5.3 VALIDACIÓN DE RESULTADOS

5.3.1 Caso de Referencia

De acuerdo con el estado del arte y con la validación funcional en el caso de deserción académica con el Sistema de Alertas Tempranas de Intervención y Seguimiento (SATIS), se construyen tres épicas teniendo como inspiración la metodología *SCRUM* [94] (ANEXO D). El caso de referencia se observa en la figura 24, asociado al servicio de enriquecimiento del perfil (*SEP*) señalado con el recuadro amarillo, en donde a partir de fuentes transaccionales y no transaccionales asociadas a la localización de una persona en el campus se puede enriquecer el perfil para prestar mejores servicios adaptados al contexto.

Figura 24. Representación gráfica del caso de referencia Muysca



Fuente. Elaboración propia a partir del estado del arte y de la exploración del negocio

En la tabla 13 se valida conceptualmente el servicio de asistencia a clase, de acuerdo con la aplicación de la arquitectura *Quysqua Smart-UJ* aplicada a los servicios Muysca.

Tabla 13. Validación de arquitectura Quysqua Smart-UJ aplicada a los servicios Muysca

Application Service (AS) Segment	Notification Manager	El consejero académico o GUI se suscriben para recibir una notificación cuando se presenten actualizaciones en el perfil enriquecido asociado a hábitos de asistencia a clase.
	User Experience Tools	Consulta en herramientas de visualización del servicio de hábitos de asistencia a clase mediante API REST.
	Security Manager Access	Control de acceso y seguridad para consumo para los datos generados en las capas.
	Adaptation Services	Perfil de usuario y contexto actualizado con los datos enriquecidos de asistencia a clase de la capa <i>Analytic Tools</i> .
Data Analytic (DA) Segment	Analytic Services Access	Acceso y seguridad a los datos generado en el segmento <i>DA</i> .
	Analytic Tools	* Etapas <i>CRISP-DM</i> aplicar: selección, exploración, limpieza y transformación de datos del 1 de enero a 31 de marzo de 2020 * Modelos descriptivo y exploratorio sobre asistencia a clase, de acuerdo con conexión AP. Creación y ejecución de indicadores sobre asistencia a clase como: * Porcentaje de asistencia en el transcurso del semestre (<i>Wifi</i> vs. horario clase) * Tiempo de conexión de acuerdo con el tipo de lugar. Tiempo y porcentaje de conexión al AP asociado al área física donde se encuentra ubicado el salón asociado al curso inscrito por un estudiante.
	Storage & Big Data	Almacenamiento en base de datos analíticos de información retrospectiva en minutos sobre la duración de conexión a <i>Wifi</i> en lugar del campus tipo salón de clase asociada al horario de un estudiante.
Ambient Intelligence (AmI) Segment	AmI Services Access	Seguridad y acceso a información del segmento <i>AmI</i> .
	Edge Intelligence	* Edge Intelligent. Ejemplo. SI un estudiante no va a clase dos semanas ENTONCES notifique al consejero académico.
	Ambient Interface	Se almacenan datos limpios de la capa <i>Data Source Services Access</i> .
Data Source (DS) Segment	Data Source Services Access	Comunicación y envío de información solicitada desde la capa <i>Ambient Interface</i> .
	Network Data Consumer	Enrutador: Nombres lógicos vs. ubicación Física AP para el tipo de lugar aula de clase. Ejemplo 1: estudiantes conectados salón 205 Ed. 67 = AP 405 (granularidad gruesa) Ejemplo 2: clase martes SMA Estudiante 1 = AP202 6:00 p.m. (muy fino)
	Ubiquitous Services Manager	No aplica para Muysca
	Node Data Manager	Consumo datos: 1. Sensorial: <i>Wifi</i> , detalle dispositivo de acceso. Localización de un estudiante dentro del campus (siempre y cuando se conecte a red <i>Wifi</i> PUJ mediante dispositivo celular, tablet, PC, entre otros) 2. Información para perfil. ID anonimizado, facultad, programa, semestre, cursos inscritos, temas cursos inscritos, horario cursos inscritos, código cursos inscritos.

Fuente. Modelo Quysqua Smart-UJ aplicado al modelo Muysca.

5.3.2 Validación con Expertos

Se realiza la validación de los servicios propuestos en *Muysca* con expertos, inspirado en el *Modelo de Adaptación de Tecnología TAM* [95]. Se busca medir la utilidad percibida (*Perceived Usefulness*). Los servicios evaluados corresponden al caso de referencia descrito en el numeral 5.3.1: asistencia a clase, lugares frecuentes y gustos y preferencias temáticos en el campus (ANEXO H). La metodología de validación seguida considera los siguientes puntos.

1. Adaptación del método *TAM* para *Muysca*. El objetivo de la validación es evaluar con expertos la utilidad percibida con respecto a los servicios de *Muysca*: hábitos de asistencia a clase mediante conexión a *Wifi*, hábitos de conexión a *Wifi* en el campus (lugares de permanencia) y hábitos en el uso de la Biblioteca General.
2. Hipótesis del proceso de validación.
 - H1. El perfil enriquecido de *Muysca* sobre hábitos de asistencia a clase y lugares frecuentes de conexión a *Wifi* en el campus, puede contribuir a la prevención de deserción académica en la Pontificia Universidad Javeriana.
 - H2. El perfil enriquecido de *Muysca* sobre hábitos de asistencia a clase y lugares frecuentes de conexión a *Wifi*, puede aportar al desarrollo de servicios adaptados al contexto de un campus inteligente.
3. Diseño de investigación: muestra, instrumento.
 - Investigación de tipo descriptivo para conocer la percepción de los expertos y usuario potenciales del sistema *Muysca*.
 - Universo. Expertos asociados a la evaluación de los servicios del sistema *Muysca*. Características de expertos.
 - Algunos expertos conocen sobre el proyecto Sistema de Alertas Tempranas de Intervención y Seguimiento (SATIS).
 - Algunos expertos tienen roles representativos y de impacto funcional en la Universidad.
 - Algunos expertos tienen roles representativos y de impacto técnico en la Universidad.
 - Algunos expertos tienen roles representativos en cuanto a la investigación sobre adaptación y uso de datos de *Wifi*.
 - Muestra de estudio. Se escogió un muestreo no probabilístico, por conveniencia, dado que la opinión de los expertos y usuarios potenciales se hace conveniente. La muestra es escogida por el investigador responsable de realizar el muestreo (no al azar). La diversidad de expertos se considera más importante que la cantidad.
 - Selección instrumento. Se utilizará una encuesta construida en *Forms de Office 365* [96].
4. Construcción instrumento: secciones, escala de evaluación y criterios de aceptación. La encuesta se construyó en dos secciones. En la primera sección, se abordaron validaciones sobre la utilidad percibida de los servicios asistencia a clase, lugares frecuentes y gustos y preferencias temáticos en el campus (mayor detalle en numeral 4.3), en cuanto a contribuciones a la investigación, estudio de deserción estudiantil en la PUJ

y en el desarrollo de servicios adaptados al contexto. En la segunda sección, se abordaron aspectos relacionados con la intención de uso de los servicios Muysca en el futuro, de acuerdo con investigaciones o proyectos en los que participa el evaluador.

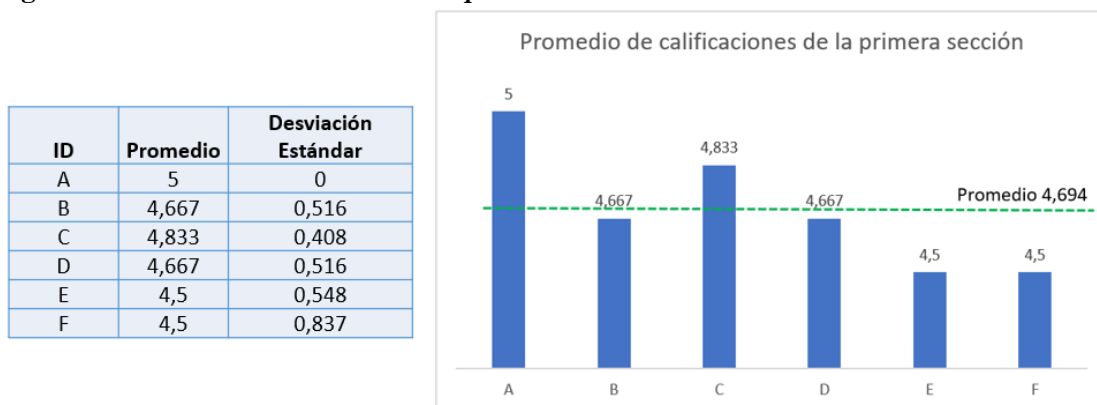
A continuación, se detalla la escala de evaluación.

1	2	3	4	5
Totalmente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Totalmente de acuerdo

A continuación, se detallan los criterios de aceptación.

- Por lo menos una respuesta de cada tipo de experto.
 - Cada una de las hipótesis se acepta con el promedio de calificaciones por sección de la encuesta, mayores a 4,00.
 - Plazo de respuestas. Hasta el sábado 5 de diciembre de 2020, a las 10:00 p.m.
5. Preparación presentación con evaluadores (piloto). Se construyó un video que detalla los servicios de *Muysca* en la URL <https://youtu.be/Koz20daRt54>
 6. Aplicación de instrumento. Encuesta de validación con expertos construida en *Forms*.
 7. Análisis de resultados. La encuesta fue respondida por ocho (8) expertos. Se cumplen los criterios de aceptación. Para mayor detalle ver ANEXO H. En la primera sección, los expertos estuvieron totalmente de acuerdo (5) en cuando a que "hábitos de asistencia a clase", a partir de información de localización basada en el Wifi, tiene potencial para contribuir positivamente al análisis y prevención de la deserción académica en la Universidad. Otra evaluación significativa (4,833) corresponde a que los "hábitos de lugares frecuentes en el campus", a partir de información de localización basada en el Wifi, tiene potencial para contribuir positivamente al estudio de la deserción académica en la Universidad (figura 25).

Figura 25. Análisis de resultados de la primera sección de la encuesta

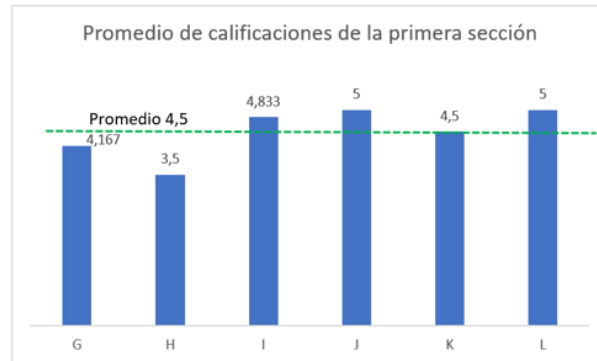


Fuente. Elaboración propia en Excel.

En la figura 26, se detalla el análisis de la evaluación de expertos de la segunda sección. Todos los expertos están totalmente de acuerdo (5) con el uso de 'alguno' de los servicios de *Muysca* en los sistemas de información o investigaciones en los que participan. También, todos los expertos están totalmente de acuerdo (5) en afirmar que servicios del sistema *Muysca* pueden enriquecer otros sistemas de información o investigaciones en curso

Figura 26. Análisis evaluación de expertos primera sección

ID	Promedio	Desviación Estándar
G	4,167	1,329
H	3,5	1,049
I	4,833	0,408
J	5	0
K	4,5	0,548
L	5	0



Fuente. Elaboración propia en Excel.

Dado el análisis de resultados se aceptan las hipótesis H1 y H2. En cuanto a las observaciones de la encuesta los expertos consideran que los servicios ofrecidos por el Sistema *Muysca* aportarán en los siguientes puntos.

- Servicios complementarios. En la medida que se puedan realizar cruces de información entre ellos. Por ejemplo, el tipo de libros tomados en préstamo en relación con las asignaturas inscritas. Igualmente permite complementar el perfil para poder ofrecer un mejor apoyo en consultoría al considerar los gustos de la persona.
- Información importante para la toma de decisiones relacionadas con la vida estudiantil en relación con la deserción académica.
- Diseño de nuevos servicios enfocados al bienestar, la investigación y el desarrollo integral en el Campus Javeriano.
- Fácil experiencia de uso tanto en PC de escritorio como en smartphones.
- Otras aplicaciones. Por ejemplo, teniendo presente esta situación de pandemia se puede analizar y generar esos espacios de estudio de acuerdo con los protocolos de bioseguridad. Otra aplicación, que permitiría realizar una priorización de servicios tecnológicos con esos datos capturados. Desde el equipo tecnológico se considera integre a más sistemas de información de la Universidad que pueden enriquecer aún más a *Muysca*, mostrando nuevas oportunidades de informes que permitan tomar decisiones de cara a la Planeación Universitaria.

6. CONCLUSIONES

Se realizó la selección del servicio de hábitos de lugares frecuentes de conexión que enriquecen el perfil de una persona en el marco de un campus inteligente. Se identificaron los servicios de asistencia a clase, lugares frecuentes y gustos y preferencias temáticos, mediante *Wifi*, *RFID* y algunos datos transaccionales. Lo anterior, permite dar cumplimiento al objetivo número uno del trabajo de grado de investigación *Muysca*.

El diseño del módulo de la arquitectura de *Muysca*, consideró la aplicación del modelo *Quysqua Smart-UJ*. A nivel tecnológico, se implementaron algunos componentes del modelo *Quysqua Smart-UJ* como la construcción del perfil de usuario y de contexto en la capa *Adaptation Services*, el mapeo de lugares del campus en relación con el tipo de lugar y dimensión de *CamIn* en el componente *Node Data Manager* y *Network Data Consumer* de segmento *DS*, el consumo de los servicios de *Muysca* en la capa *User Experience Tools* del segmento *AS*, gestionando el acceso a la información mediante niveles de seguridad y autenticación institucional.

A nivel analítico, se realizó el análisis exploratorio en la capa *Analytic Tools* del segmento *DA*, encontrando clústeres mediante *DBSCAN* por día de la semana y hora del día asociado a la conexión de un AP *Wifi* por parte de una persona. Las contribuciones a nivel tecnológico y a nivel analítico permiten dar cumplimiento al objetivo número dos del trabajo de grado de investigación *Muysca*.

A nivel conceptual, se realizó la caracterización y selección mediante criterios de los servicios potenciales en un *CamIn*, lo cual conllevó a la selección del caso de referencia, así como al diseño de servicios propuestos por *Muysca*. Como producto de la investigación, se elaboró un artículo científico aceptado en la conferencia *ICITS'21 - The 2021 International Conference on Information Technology & Systems (2021)*, titulado ‘Services based on the enriched profile of a person within a smart university’ [66].

A nivel institucional, se obtuvo una propuesta con nuevas características que enriquecen el perfil de una persona mediante la observación del comportamiento desde fuentes no transaccionales y transaccionales en el marco de consejería académica y *SATIS*, permitiendo la integración de diferentes fuentes de información y mostrando nuevas oportunidades para apoyar la toma de decisiones de la Universidad con datos de *Wifi* y *RFID*. Lo anterior, permite dar cumplimiento al objetivo número tres del trabajo de grado de investigación *Muysca*.

Como trabajo futuro se propone la implementación de los servicios *Muysca* en el *core* tecnológico de la Universidad, el diseño de nuevos servicios inteligentes en tiempo real que tomen como base los servicios de la presente investigación, considerando el perfil grupal del usuario y contexto *CamIn* mediante la localización. También se propone la continuación en la investigación de servicios adaptados al contexto de un *CamIn*, que permita realizar recomendación considerando modelos de adaptación. A nivel analítico, se propone, el consumo de más información, que permita realizar otros análisis con enfoque predictivo, asociado al perfil enriquecido de una persona.

7. ANEXOS

A continuación, se detallan los anexos del trabajo de grado de investigación *Muysca*.

Anexo	Descripción
ANEXO A	Análisis del estado del arte
ANEXO B	Caracterización de oportunidades y servicios identificados
ANEXO C	Indicadores y servicios de Muysca
ANEXO D	Diseño del módulo de la arquitectura ampliado
ANEXO E	Análisis exploratorio CRISP-DM
ANEXO F	Módulo de la arquitectura
ANEXO G	Validaciones del sistema Muysca
ANEXO H	Validación de los servicios Muysca con expertos
ANEXO I	Análisis de alertas y servicios Muysca a partir de la localización
ANEXO J	Autorización del negocio para el uso de los datos
ANEXO K	Sustentación TG MISYC MAINN

REFERENCIAS

- [1] M. Alvarez, G. López, E. Vázquez, V. Villagrà y J. Berrocal, «Smart CEI Moncloa: An ICo-based platform for people flow and environmental monitoring on a Smart University Campus,» *Sensors*, vol. 17, n° 12, pp. 1-24, 2017.
- [2] A. Atayero, S. Popoola, O. Adeyemi, D. Afolayan, M. Akanle, V. Adetola y E. Adetiba, «Trends and patterns of broadband Internet access speed in a Nigerian university campus: A robust data exploration,» *Data in Brief*, vol. 23, 2019.
- [3] A. Pacheco, P. Cano, E. Flores, E. Trujillo y P. Marquez, «A Smart Classroom Based on Deep Learning and Osmotic IoT Computing,» de *Congreso Internacional de Innovacion y Tendencias En Ingenieria, CONIITI 2018 - Proceedings*, 2018.
- [4] W. Wang, A. Liu, M. Shahzad, K. Ling y S. Lu, «Understanding and modeling of WiFi signal based human activity recognition,» de *MobiCom '15: Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, New York, USA, 2015.
- [5] A. Carrillo, F. Aragón, J. Cárdenas, J. Cristancho, M. Higuera, D. Marín, L. Niño, J. Nova, A. Orozco, A. Rico y A. Romero, «IAM: Modelo Integrado de Adaptación,» *Avances En Sistemas E Informática*, vol. 6, n° 3, pp. 145-162, 2009.
- [6] M. Henríquez y P. Palma, «Control Automático de Condiciones Ambientales en Domótica usando Redes Neuronales Artificiales,» *Información Tecnológica*, vol. 22, n° 3, pp. 125-139, 2011.
- [7] J. Agreda y E. González, «Ambient intelligence based multi-agent system for attend elderly people,» *9th Computing Colombian Conference*, pp. 115-120, 2014.
- [8] S. Cuomo, P. Michele, F. Piccialli, A. Galletti y J. Jung, «IoT-based collaborative reputation system for associating visitors and artworks in a cultural scenario,» *Expert Systems with Applications*, vol. 79, pp. 101-111, 2017.
- [9] L. Patiño y A. Cardona, «Review of some studies on university student dropout in Colombia and Latin America,» *Acta Universitaria*, vol. 23, n° 4, pp. 37-46, 2013.
- [10] M. Picasso, M. Villanelo y L. Bedoya, «Habits of reading and study and its relationship to academic achievement in dental students of a Peruvian University,» *KIRU*, vol. 12, n° 1, pp. 19-27, 2015.
- [11] F. Oswald, N. Schmitt, B. Kim, L. Ramsay y M. Gillespie, «Developing a biodata measure and situational judgment inventory as predictors of college student performance,» *Journal of Applied Psychology*, vol. 89, n° 2, pp. 187-207, 2004.
- [12] M. d. E. nacional, «Cifras de deserción académica durante el año 2015,» Bogotá, Colombia, 2015.
- [13] P. U. Javeriana, «Informe del rector al consejo de regentes año 2018,» Bogotá, Colombia, 2018.
- [14] C. Jaramillo, C. Ordóñez, A. González, M. León, M. Curiel, I. Palacios y O. Barrera, «Formulación del programa de prevención de la deserción: articulación de acciones en un programa transversal de acompañamiento para la Pontificia Universidad Javeriana,» de *Octava Conferencia Latinoamericana sobre el abandono en la educación superior*, Panamá, 2018.

- [15] E. Ismagilova, L. Hughes, Y. Dwivedi y K. Raman, «Smart cities: Advances in research—An information systems perspective,» *International Journal of Information Management*, vol. 47, pp. 88-100, 2019.
- [16] C. Galeano, D. Bellón, S. Zabala, E. Romero y V. Novoa, «Identificación de los pilares que direccionan a una institución universitaria hacia un Smart-Campus,» *Revista de Investigación, Desarrollo e Innovación*, vol. 9, n° 1, pp. 127-145, 2018.
- [17] M. Reguera, N. Sotelo, A. Barraza y M. Cortéz, «Estilo de vida y factores de riesgo en estudiantes de la Universidad de Sonora: un estudio exploratorio,» *Boletín Clínico Hospital Infantil del Estado de Sonora*, vol. 32, n° 2, pp. 63-68, 2015.
- [18] M. Eisenhauer, P. Rosengren y P. Antolin, «A development platform for integrating wireless devices and sensors into Ambient Intelligence systems,» *6th IEEE Annual Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops*, pp. 1-3, 2009.
- [19] G. Acampora, V. Loia y A. Vitiello, «A cognitive multi-agent system for emotion-aware ambient intelligence,» de *Symposium Series on Computational Intelligence - IA*, 2011.
- [20] A. Ilapakurti, J. Vuppalapati, S. Kedari, R. Vuppalapati y C. Vuppalapati, «Adaptive edge analytics for creating memorable customer experience and venue brand engagement, a scented case for Smart Cities,» de *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*, 2017.
- [21] I. Argote, R. Jimenez y J. Gómez, «Detección de patrones de deserción en los programas de pregrado de la Universidad Mariana de San Juan de Pasto, aplicando el proceso de descubrimiento de conocimiento sobre base de datos (KDD) y su implementación modelos matemáticos de predicción,» de *Cuarta Conferencia Latinoamericana sobre el Abandono en la Educación Superior*, Medellín, Colombia, 2014.
- [22] M. Enríquez, M. Fajardo y F. Garzón, «College students habits and study techniques. A general review,» *Psicogente*, vol. 18, n° 33, p. 166–187, 2015.
- [23] D. Kofoed, P. Sapiezynski, M. Furman y S. Lehmann, «Inferring Stop-Locations from WiFi,» *Plos One*, vol. 11, n° 2, 2016.
- [24] K. Seo, Y. Jang, Y. Kwon y R. Lee, «An ambient service model for providing structured web information based on user-contexts,» *Proceedings*, pp. 421-428, 2010.
- [25] K. Park, J. Park y S. Kim, «An effective model and management scheme of personal space for ubiquitous computing applications,» *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, vol. 38, n° 6, pp. 1295-1311, 2008.
- [26] S. Restrepo, J. Pezoa y D. Ovalle, «An Adaptive Architecture for Ambient Intelligence Based on Meta-modeling, Smart Agents, and Wireless Sensor Networks,» *IEEE Latin America Transactions*, vol. 12, n° 8, pp. 1508-1514, 2014.
- [27] L. Barreto, «Integración de Herramienta de Analítica en el Modelo AMI-IOT QUYSQUA,» Pontificia Universidad Javeriana, 2017.
- [28] E. González, M. Serrano, A. Montero y E. Ruiz, «PPU-Smart University,» Pontificia Universidad Javeriana, Bogotá, Colombia, 2019.

- [29] Pontificia Universidad Javeriana, «Informe del Rector al Consejo de Regentes,» Bogotá, Colombia, 2018.
- [30] R. Hernández, C. Fernández y P. Baptista, *Metodología de la Investigación*, México: McGraw-Hill, 2001.
- [31] S. Oyedepo, T. Adekeye, O. Leramo, O. Kilanko, O. Babalola, A. Balogun y M. Akhibi, «Assessment of energy saving potentials in covenant university, Nigeria,» *Energy Engineering*, vol. 113, n° 3, pp. 7-26, 2016.
- [32] E. De Angelis, A. Ciribini, L. Tagliabue y M. Paneroni, «The Brescia Smart Campus Demonstrator. Renovation toward a zero Energy Classroom Building,» *Procedia Engineering*, vol. 118, pp. 735-743, 2015.
- [33] T. Takpor y A. Atayero, «Integrating internet of things and ehealth solutions for students' healthcare,» de *Proceedings of the World Congress on Engineering*, London, 2015.
- [34] J. Tan, «Growth of industry clusters and innovation: Lessons from Beijing Zhongguancun Science Park,» *Journal of Business Venturing*, vol. 21, n° 6, p. 827-850, 2006.
- [35] E. Sánchez, «Modelo Informático de integración AmI-IoT para el Cuidado de Adultos Mayores,» Facultad de Ingeniería, Pontificia Universidad Javeriana, Bogotá, Colombia, 2017.
- [36] S. Kantareddy, I. Mathews, R. Bhattacharyya, I. Peters, T. Buonassisi y S. Sarma, «Long Range Battery-Less PV-Powered RFID Tag Sensors,» *IEEE Internet of Things Journal*, vol. 6, n° 4, pp. 6989-6996, 2019.
- [37] L. Chuling, Z. Xie y P. Peng, «A discussion on the framework of smarter campus,» de *2009 third international symposium on intelligent information technology application*, Shanghai, 2009.
- [38] J. Koch, J. Wettach, E. Bloch y k. Berns, «Indoor Localisation of Humans, Objects, and mobile Robots with RFID Infrastructure,» de *7th International Conference on Hybrid Intelligent Systems (HIS 2007)*, Kaiserlautern, Alemania, 2007.
- [39] J. Raper, G. Gartner, H. Karimi y C. Rizos, «A critical evaluation of location-based services and their potential,» *Journal of Location Based Services*, p. 5-45, 2007.
- [40] P. Enge y P. Misra, «Special issue on global positioning system,» *Proceedings of the IEEE*, vol. 87, n° 1, pp. 3-15, 1999.
- [41] G. Huang, Z. Hu, J. Wu y H. Xiao, «WiFi and Vision-Integrated Fingerprint for Smartphone-Based Self-Localization in Public Indoor Scenes,» *IEEE Internet of Things Journal*, vol. 7, n° 8, pp. 6748 - 6761, 2020.
- [42] T. García, A. García, A. Gómez, J. Botia, H. Hagraas, J. Dooley y V. Callaghan, «An adaptive learning fuzzy logic system for indoor localisation using Wi-Fi in Ambient Intelligent Environments,» de *2012 IEEE International Conference on Fuzzy Systems*, Brisbane, Australia, 2012.
- [43] T. Erickson, «Some problems with the notion of context-aware computing,» *ACM*, vol. 45, n° 2, pp. 102-104, 2002.
- [44] A. Aztiria, J. Augusto y R. Basagoiti, «Learning frequent behaviors of the users in intelligent environments,» *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, n° 6, pp. 1265-1278, 2013.

- [45] H. Kang, J. Han y G. Kwon, «An Ecological Approach to Smart Homes for Health Care Services: Conceptual Framework of a Smart Servicescape Wheel,» *JMIR Mhealth Uhealth*, vol. 7, n° 2, 2019.
- [46] C. Nebeker, T. Lagare, M. Takemoto, B. Lewars, K. Crist, C. Bloss y J. Kerr, «Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research,» *Translational Behavioral Medicine*, vol. 6, n° 4, pp. 577-586, 2016.
- [47] Real Academia Española, «Diccionario de la Lengua Española,» 2020. [En línea]. Available: <https://dle.rae.es/comportamiento>. [Último acceso: julio 2020].
- [48] L. Hadjas, C. Lüscher y L. Simmler, «Aberrant habit formation in the Sapap3-knockout mouse model of obsessive-compulsive disorder,» *Scientific Reports*, 2019.
- [49] Real Academia Española, «Diccionario de la Lengua Española,» [En línea]. Available: <https://dle.rae.es/h%C3%A1bito>. [Último acceso: 15 junio 2020].
- [50] P. Oliveira, P. Matos y P. Novais, «Behaviour Analysis in Smart Spaces,» de *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCoM/IoP/SmartWorld)*, Toulouse, France, 2016.
- [51] D. Cournapeau, «Scikit-learn,» de *Google Summer of Code*, Rocquencourt, Francia, 2010.
- [52] E. Schubert, J. Sander, M. Ester, H. Kriegel y X. Xu, «DBSCAN revisited, revisited: why and how you should (still) use DBSCAN,» *ACM Transactions on Database Systems (TODS)*, vol. 42, n° 3, p. 19, 2017.
- [53] R. Wood y T. Payne, *Competency based recruitment and selection. A practical guide*, Chichester: John Wiley & Sons, 1998.
- [54] G. González, M. Héctor, M. Duque, D. Nestor, C. Ovalle y A. Demetrio, «Modelo del Estudiante para Sistemas Adaptativos de Educación Virtual,» *Avances en Sistemas e Informática*, vol. 5, n° 1, pp. 199-206, 2008.
- [55] D. Duque, D. Ovalle y J. Jiménez, «Adaptive Model for Virtual Courses based on Intelligent Planning Techniques,» *Revista Avances en Sistemas e Informática*, vol. 4, n° 1, pp. 39-46, 2007.
- [56] J. González, A. Carrillo, R. Fabregat, L. Camargo, M. García y M. Pavlich, «RunayaySoft,» *International Journal of Web Information Systems*, vol. 15, n° 1, pp. 103-131, 2019.
- [57] L. Aballay, S. Aciar, C. Collazos y C. González, «Adaptation Model Content Based in Cultural Profile into Learning Environment,» *IEEE Latin America Transactions*, vol. 13, n° 2, pp. 490-495, 2015.
- [58] L. Goldberg, «The Structure of Phenotypic Personality Traits,» *American Psychologist*, vol. 48, n° 1, pp. 26-34, 1993.
- [59] F. Suárez, F. Buey y F. Herrero, «Styles and Learning Strategies in university students,» *Psicothema*, vol. 12, pp. 615-622, 2000.
- [60] L. Morales, «KONSEY 2.0 Sistema de Consejería,» de *Pontificia Universidad Javeriana*, Bogotá, Colombia, 2014.

- [61] M. Novoa, C. Nieto, C. Forero, C. Caycedo, M. Palma, M. Montealegre, M. Bayona y C. Sánchez, «Relación Entre Perfil Psicológico, Calidad De Vida Y Estrés Asistencial En Personal De Enfermería,» *Universitas Psychologica*, vol. 4, n° 1, pp. 63-76, 2005.
- [62] D. Garlan, S. Cheng y B. Schmerl, Increasing system dependability through architecture-based self-repair, *Lecture Notes in Computer Science*, Springer-Verlag, Inc., New York, NY, Berlin: Springer, 2007.
- [63] D. Abowd, P. Brown, M. Smith y P. Steggles, «Towards a better understanding of context and context-awareness,» *International Symposium on Handheld and Ubiquitous Computing*, pp. 304-307, 1999.
- [64] A. Dey, «Understanding and using context,» *Personal and Ubiquitous Computing*, vol. 5, pp. 4-7, 2001.
- [65] D. Gough, S. Oliver y J. Thomas, *An Introduction to Systematic Reviews*, London: Sage, 2017.
- [66] V. Arias y E. González, «Services based on the enriched profile of a person within a smart university,» de *ICITS'21 - The 2021 International Conference on Information Technology & Systems*, Santa Elena, Ecuador, 2021.
- [67] IHS Markit, «The Internet of Things: A Movement Not a Market Critical IoT Insights,» London, U.K., 2017.
- [68] S. Miles, S. Sarma y J. Williams, *RFID Technology and Applications*, New York, USA: Cambridge, 2008.
- [69] L. Soyoungh, «Intercultural transitions in higher education: international student adjustment and adaptation,» de *Higher Education*, 2020.
- [70] S. Mohamed, A. Enjy, W. Nourhan, A. Adel y W. Ray , «Medical Education Adaptations Post COVID-19: An Egyptian Reflection,» de *Journal of Medical Education and Curricular Development*, 2020.
- [71] J. Hanson, «Effectual thinking and music education: One view of creative adaptation in an underserved urban middle school,» de *International Journal of Music Education*, 2020.
- [72] Y. Shi y J. Xie, «A study of personalized training strategies for vocational education based on big data systems,» de *The Journal of Physics: Conference Series*, 2020.
- [73] Y. Cheng y X. Bu, «Research on key technologies of personalized education resource recommendation system based on big data environment,» *The Journal of Physics: Conference Series*, vol. 1437, n° 1, 2020.
- [74] Z. Zhang y H. Min, «Analysis on the construction of personalized physical education teaching system based on a cloud computing platform,» de *Wireless Communications and Mobile Computing*, 2020.
- [75] C. Porcel, J. Herce, J. Bernabé, A. Tejeda y E. Herrera, «Trust based fuzzy linguistic recommender systems as reinforcement for personalized education in the field of oral surgery and implantology,» *International Journal of Computers, Communications and Control*, vol. 15, n° 3, 2020.
- [76] H. Zitouni, S. Meshoul y C. Mezioud, «New contextual collaborative filtering system with application to personalized healthy nutrition education,» *Journal of King Saud University - Computer and Information Sciences*, 2020.

- [77] H. Gasimov, «Modelling support systems for selecting professions for applicants in the content of personalization of education,» *EUREKA, Physics and Engineering*, vol. 2020, n° 2, pp. 83-97, 2020.
- [78] S. Filipiak y B. Łubianka, «Big five personality traits among polish students facing the transformations of the education system,» *Current Issues in Personality Psychology*, vol. 7, n° 2, pp. 98-108, 2019.
- [79] T. Andrade, B. Cancela y J. Gama, «Discovering locations and habits from human mobility data,» *Annals of Telecommunications*, 2020.
- [80] D. Wind, P. Sapiezynsk y M. Furman, «Inferring Stop-Locations from WiFi,» *PLoS ONE*, vol. 11, n° 2, 2016.
- [81] W. Yoo, H. Kim y M. Shin, «Stations-oriented indoor localization (SOIL): A BIM-Based occupancy schedule modeling system.,» *Building and Enviroment*, 2020.
- [82] M. d. E. Nacional, «Definición de deserción,» Bogotá, Colombia, 2020.
- [83] K. Fuentes y C. Sánchez, «Análisis predictivo para la identificación de estudiantes que presentan un alto riesgo de deserción en la Pontificia Universidad Javeriana. Evaluación de factores asociados,» Pontificia Universidad Javeriana, Bogotá, Colombia, 2017.
- [84] L. Girón y D. González, «Determinantes del rendimiento académico y la deserción estudiantil, en el programa de Economía de la Pontificia Universidad Javeriana de Cali,» Pontificia Universidad Javeriana, Cali, Colombia, 2005.
- [85] L. Burgos, M. Rojas y C. Saldarriaga, «SAPIENS : Sistema de Alertas Tempranas para estudiantes de primer año de la Facultad de Ingeniería de la Pontificia Universidad Javeriana,» Pontificia Universidad Javeriana, Bogotá, Colombia, 2018.
- [86] C. d. l. República, «Ley estatutaria 1581 de 2012». Colombia Patente Diario Oficial No. 48.587, 18 octubre 2012.
- [87] S. d. I. y. Comercio, «Guía para la implementación del Principio de Responsabilidad Demostrada (Accountability)». Colombia 2020.
- [88] C. d. l. República, «Ley estatutaria 1266 de 2008». Colombia Patente Diario Oficial No. 47.219, 31 diciembre 2008.
- [89] A. Torres, A. Contreras, A. Rodríguez, J. Rodríguez y S. Salamanca, «Space Analytics,» de *Pontificia Universidad Javeriana*, Bogotá, Colombia, 2019.
- [90] C. Álvarez, «Arquitectura basada en redes neuronales para identificación de escenarios de deserción estudiantil en la facultad de ingeniería de la Pontificia Universidad Javeriana,» de *Pontificia Universidad Javeriana*, Bogotá, Colombia, 2014.
- [91] P. Chapman, J. Clinton, Keber, T. Khabaza, T. Reinart, C. Shearer, R. Wirth, Y. Cheng y X. Bu, « Step-by-step data mining guide Research on key technologies of personalized education resource recommendation system based on big data environment,» *The Journal of Physics: Conference Series*, vol. 1437, n° 1, 2000.
- [92] O. 365, «Servicio Empresarial Power BI,» 2014.
- [93] A. Lengyel y Z. Botta-Dukát, «Silhouette width using generalized mean—A flexible method for assessing clustering efficiency,» *Ecology and Evolution*, vol. 9, n° 23, pp. 13231-13243, 2019.
- [94] K. Schwaber y J. Sutherland, «Scrum Guide™,» 2020. [En línea].

- [95] D. Adams, R. Nelson y P. Todd,, «Perceived usefulness, ease of use, and usage of information technology: a replication,» *MIS Quarterly*, vol. 16, n° 2, pp. 227-247, 1992.
- [96] O. 365, «Tecnología de Microsoft Forms,» 2016.