

**Universidad ORT Uruguay**  
**Facultad de Ingeniería**

# **Fenotipado Digital**

Entregado como requisito para la obtención del título de Ingeniero en Sistemas

**Santiago González – 145583**

**Jorge Kinosian - 183470**

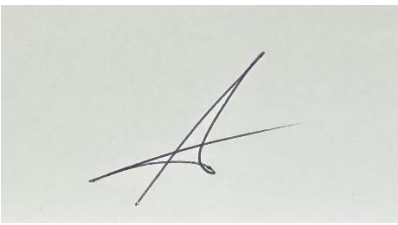

**Tutor: Gerardo Maturro**

**2025**

## Declaración de autoría

Nosotros, Santiago González y Jorge Kinosian, declaramos que el trabajo que se presenta en esa obra es de nuestra propia mano. Podemos asegurar que:

- La obra fue producida en su totalidad mientras realizábamos la tesis de grado;
- Cuando hemos consultado el trabajo publicado por otros, lo hemos atribuido con claridad;
- Cuando hemos citado obras de otros, hemos indicado las fuentes. Con excepción de estas citas, la obra es enteramente nuestra;
- En la obra, hemos acusado recibo de las ayudas recibidas;
- Cuando la obra se basa en trabajo realizado conjuntamente con otros, hemos explicado claramente qué fue contribuido por otros, y qué fue contribuido por nosotros;
- Ninguna parte de este trabajo ha sido publicada previamente a su entrega, excepto donde se han realizado las aclaraciones correspondientes.

	
<b>Santiago Gonzalez</b>	<b>Jorge Kinosian</b>
<b>Montevideo, 20 de Marzo de 2025</b>	

## **Agradecimientos**

Queremos agradecer, en primer lugar, a nuestras familias, que nos apoyaron en cada etapa de este proceso. También a nuestros docentes, por su orientación y enseñanzas, y a nuestros compañeros, con quienes compartimos este camino. Un agradecimiento especial a Gerardo Maturro, por su valiosa guía y apoyo a lo largo del trabajo.

## Abstract

**Contexto:** La presente tesis explora el concepto de fenotipado digital y su aplicación en la salud mental, fundamentado en la idea del “fenotipo extendido” y en la convergencia entre psicoinformática y análisis de datos. A través del uso de dispositivos móviles (*smartphones* y *wearables*) y técnicas de monitoreo pasivo, se aborda cómo la recolección continua de información puede brindar un panorama más preciso de la conducta y el estado de salud de las personas.

**Objetivo:** El objetivo principal fue analizar las bases teóricas del fenotipado digital, su evolución histórica y el potencial de esta metodología en la detección y seguimiento de trastornos mentales. Además, se planteó diseñar e implementar un prototipo funcional para poner en práctica los conceptos investigados.

**Método:** Se realizó una revisión bibliográfica sobre fenotipado digital, incluyendo aspectos de privacidad, consentimiento informado y modelos de arquitectura de captura de datos. Paralelamente, se llevó a cabo el desarrollo de una aplicación móvil y un *backend* para recolectar información de sensores (GPS, sensor de luz, bloqueo y desbloqueo).

**Resultados:** El prototipo demostró la factibilidad de integrar sensores múltiples en una sola plataforma, permitiendo el registro continuo de variables de comportamiento. Asimismo, la implementación de buenas prácticas de privacidad (anonimización y consentimiento informado) facilitó la aceptación del usuario. Los resultados evidenciaron patrones útiles para caracterizar la movilidad, el descanso y la interacción social, aspectos clave para el monitoreo de síntomas.

**Conclusiones:** El fenotipado digital ofrece un enfoque prometedor para mejorar la evaluación, prevención y manejo de condiciones mentales. La incorporación de datos en tiempo real puede anticipar recaídas y posibilitar intervenciones más oportunas. No obstante, es fundamental reforzar la ética, la seguridad y la privacidad de la información para garantizar un uso responsable y confiable de estas tecnologías.

## Palabras clave

Fenotipado digital, Salud mental, Monitoreo pasivo, Monitoreo activo, Biomarcadores digitales, Psiquiatría, *Mobile sensing*, Aprendizaje automático (*Machine Learning*), Big Data, Privacidad, Consentimiento informado, Psico informática, Tecnologías móviles

# Índice

1.	Introducción .....	12
1.1.	Definición de “Fenotipado Digital” .....	12
1.2.	Alcance y objetivos de la tesis .....	13
1.3.	Áreas de aplicación .....	13
1.4.	Estructura de la tesis y desarrollo de un prototipo .....	14
1.5.	Conclusión.....	15
2.	Conceptos fundamentales sobre fenotipado digital.....	17
2.1.	Fenotipado digital.....	17
2.2.	Contexto .....	18
2.3.	Context Aware.....	18
2.4.	Mobile sensing .....	19
2.5.	Personalidad digital .....	20
2.6.	Sensing architecture .....	20
2.7.	Tipos de sensores .....	22
2.8.	Registro semántico .....	24
2.9.	Clínica digital y Monitoreo Continuo .....	24
2.10.	Datos activos y pasivos .....	26
2.11.	Monitoreo Pasivo .....	27
2.12.	Biomarcadores y Biomarcadores Digitales .....	27
2.13.	Biomarcador de Control Cognitivo .....	28
2.14.	Biomarcador de Memoria de Trabajo .....	28
2.15.	Consentimiento Informado.....	28
2.16.	Modelo de Privacidad para la Recolección de Datos Móviles.....	30
3.	Origen y evolución histórica .....	33

3.1.	Fundamentos conceptuales del “Fenotipado extendido” al “Fenotipado digital” ....	33
3.2.	Aparición formal del término y definiciones iniciales .....	33
3.3.	Ventajas tempranas y ejemplos iniciales.....	34
3.4.	Expansión en psiquiatría y salud mental .....	34
3.5.	Evolución del campo y madurez tecnológica.....	35
3.6.	Contribuciones al conocimiento y retos futuros.....	36
3.7.	Perspectiva integradora: el futuro del fenotipado digital .....	36
4.	Áreas de aplicación .....	38
4.1.	Fenotipado Digital y su Evolución Conceptual.....	38
4.2.	Aplicaciones en Salud Mental.....	39
4.3.	Investigación en Big Data y Aprendizaje Automático.....	40
4.4.	Psicoinformática y Nuevos Campos de Estudio .....	40
4.5.	Atención Médica Digitalmente Soportada .....	41
4.6.	Áreas de Expansión: Salud Pública y Más Allá.....	41
4.7.	Desafíos y Consideraciones Éticas.....	42
4.8.	Conclusiones .....	42
5.	Consideraciones éticas en el fenotipado digital .....	44
5.1.	Privacidad de los datos y capitalismo de vigilancia.....	44
5.2.	El consentimiento informado como piedra angular .....	45
5.3.	Responsabilidad de los comités de ética e instituciones .....	46
5.4.	Riesgos de despersonalización, patologización y exclusión .....	47
5.5.	Diseño centrado en el usuario y consideraciones de aceptabilidad.....	47
5.6.	Retos técnicos y validez de los datos en escenarios reales .....	48
5.7.	Perspectivas futuras y líneas de acción .....	48
5.8.	Cumplimiento Normativo: GDPR y Otras Regulaciones .....	49
5.9.	Transparencia, Consentimiento Informado y Potenciales Riesgos .....	50
5.10.	Conclusiones .....	50

6.	Perspectivas interdisciplinarias .....	52
6.1.	Colaboración entre Ciencia, Tecnología y Sector Comercial .....	52
6.2.	Avances en Psiquiatría y Ciencias del Comportamiento .....	53
6.3.	Ampliación de Campos de Aplicación.....	54
6.4.	Perspectiva Crítica e Interdisciplinaria .....	56
6.5.	Plataformas, Herramientas y Equipos Multidisciplinarios.....	56
6.6.	Aplicaciones en la Prevención de Conductas de Riesgo.....	57
6.7.	Desarrollo Metodológico, Machine Learning y Poblaciones Diversas.....	58
6.8.	Hacia la Psiquiatría de Precisión y Nuevas Categorías Diagnósticas .....	59
6.9.	Perspectiva de la Psicoinformática y el Estudio de la Personalidad .....	60
6.10.	Ética y Empoderamiento del Paciente.....	61
6.11.	Convergencia Metodológica y Coexistencia con Evaluaciones Tradicionales ....	62
6.12.	Movilización General del Conocimiento y Sigüientes Pasos.....	62
6.13.	Conclusiones .....	63
7.	Tecnologías para el fenotipado digital .....	65
7.1.	Herramientas de desarrollo de software .....	65
7.2.	Técnicas de análisis y procesamientos de datos.....	68
	Etapas en el análisis y procesamiento de datos .....	69
	Técnicas de preprocesamiento, procesamiento y análisis de los datos .....	71
7.3.	Seguridad, privacidad y protección de datos.....	75
	Introducción a la Privacidad y Seguridad en actividades en línea .....	75
	Privacidad como concepto legal.....	76
	Privacidad como concepto técnico.....	77
	Medidas de protección y prevención.....	78
8.	Prueba de conceptos, prototipo funcional .....	81
8.1.	Introducción .....	81
8.2.	Objetivo del proyecto .....	81

8.3.	Alcance.....	82
8.4.	Requerimientos.....	83
	Funcionales.....	83
	No funcionales.....	84
8.5.	Diagrama de despliegue .....	86
8.6.	Modelo de gestión y ciclo de vida.....	87
8.7.	Tecnologías utilizadas .....	89
	Aplicación Móvil: Kotlin vs otras tecnologías.....	89
	Backend: Node.js vs otras tecnologías .....	90
	Base de datos: MySQL vs otras tecnologías .....	90
	Plataforma de visualización: Power BI vs Otras herramientas .....	91
8.8.	Decisiones clave.....	92
8.9.	Análisis de riesgos.....	93
	Riesgos técnicos .....	93
	Riesgos éticos.....	93
	Riesgos operativos.....	94
8.10.	Ilustraciones del prototipo y reporte.....	94
8.11.	Conclusiones .....	98
9.	Conclusiones generales .....	100
10.	Trabajos futuros.....	102
11.	Lecciones aprendidas .....	104
12.	Referencias bibliográficas .....	106

## Índice de tablas

Tabla 1 Comparación entre Kotlin y otras tecnologías .....	90
Tabla 2 Comparación entre Node.js y otras tecnologías .....	90
Tabla 3 Comparación entre MySQL y otras tecnologías .....	91
Tabla 4 Comparación entre Power BI y otras tecnologías .....	91

## Índice de ilustraciones

Ilustración 1 Clínica digital .....	24
Ilustración 2 Icono de SensoTrack .....	94
Ilustración 3 Pantalla principal de SensoTrack .....	95
Ilustración 4 Hoja principal del reporte.....	95
Ilustración 5 Hoja del sensor de ubicación del reporte .....	96
Ilustración 6 Hoja del sensor de luz del reporte .....	96
Ilustración 7 Hoja del sensor de bloqueo del reporte .....	97
Ilustración 8 Hoja de datos del reporte.....	97

# 1. Introducción

La presente tesis se inscribe en el campo emergente del fenotipado digital, un enfoque que consiste en recolectar y analizar datos generados a través de dispositivos electrónicos (especialmente teléfonos inteligentes (*smartphones*) y sensores portátiles) para describir y predecir conductas, estados de salud y rasgos psicológicos de las personas. Este concepto, como se profundiza en el documento de la tesis, surge de la convergencia entre la evolución tecnológica de las últimas décadas y las necesidades de áreas tan diversas como la salud mental, la ciencia de datos y la psico informática. A lo largo de este capítulo introductorio, se explican la definición del término, el alcance y los objetivos de la tesis, y se ofrecen pinceladas acerca de las principales áreas de aplicación en las que el fenotipado digital muestra un potencial significativo. Finalmente, se adelanta la existencia de un prototipo desarrollado durante la investigación, con miras a ilustrar y validar los conceptos presentados.

## 1.1. Definición de “Fenotipado Digital”

El fenotipado digital se puede definir como la cuantificación momento a momento del fenotipo humano (es decir, del conjunto de características observables de un individuo) mediante datos derivados de dispositivos digitales personales. En lugar de basarse exclusivamente en métodos clínicos tradicionales (encuestas, cuestionarios, entrevistas puntuales), el fenotipado digital aprovecha la tecnología para realizar una captura sistemática y continua de información. Ejemplos de estas fuentes de datos incluyen:

- **Sensores integrados en el teléfono móvil:** GPS para ubicación, acelerómetro para actividad física, micrófono para análisis de voz, entre otros.
- **Uso de aplicaciones y redes sociales:** frecuencia de conexión, tipos de aplicaciones empleadas, patrones de interacción social.
- **Dispositivos portátiles (*wearables*):** relojes inteligentes o pulseras de actividad que monitorizan variables fisiológicas (frecuencia cardíaca, patrones de sueño, pasos diarios).

A partir de estos indicadores, el fenotipado digital busca extraer lo que en el documento se denomina “biomarcadores digitales”: métricas objetivas que dan cuenta de estados psicológicos, rasgos de personalidad o condiciones de salud. El potencial clínico y científico de este enfoque está fundamentado en su capacidad de recolección continua y no invasiva de

información, lo cual abre la puerta a diagnósticos más tempranos, intervenciones más ágiles e incluso tratamientos personalizados.

## 1.2. Alcance y objetivos de la tesis

En esta tesis se analiza la forma en que el fenotipado digital puede aplicarse para:

- **Mejorar la comprensión de la conducta y la salud mental:** Se propone estudiar cómo la recopilación de datos pasivos (captados de manera casi imperceptible para el usuario) y datos activos (aportados conscientemente) posibilita obtener una visión más profunda y realista del estado diario de las personas.
- **Desarrollar sistemas de monitoreo continuo:** Se abordan las arquitecturas tecnológicas (llamadas *Sensing Architecture*) para capturar, procesar y analizar la información. Estas arquitecturas suelen integrar múltiples sensores, algoritmos de *machine learning* y sistemas de retroalimentación, ofreciendo así una visión integrada de la rutina y el contexto del usuario.
- **Implementar aplicaciones en escenarios reales:** Más allá de la teoría, la tesis describe el diseño y construcción de un prototipo que pone en práctica los conceptos de fenotipado digital. Esta implementación pretende demostrar la factibilidad del monitoreo digital, la usabilidad de la herramienta y la capacidad de generar datos confiables que puedan ayudar a profesionales de la salud y otros actores involucrados.

Al concluir este trabajo, se pretende que el lector cuente con un panorama amplio de los fundamentos teóricos del fenotipado digital, las áreas donde su aplicación es más prometedora y las consideraciones técnicas y éticas relevantes para su despliegue.

## 1.3. Áreas de aplicación

Una de las secciones más relevantes de la tesis se centra en las áreas de aplicación del fenotipado digital. El documento expone cómo este método se ha incorporado especialmente en la salud mental y la psiquiatría, pero también describe usos potenciales en la salud pública y en otros campos emergentes.

Salud Mental y Psiquiatría

- **Diagnóstico y Evaluación:** El fenotipado digital permite complementar los métodos clínicos convencionales (como cuestionarios de autoinforme o escalas diagnósticas) con

datos de la vida cotidiana del paciente. Por ejemplo, los cambios en la movilidad captados por el GPS o en la rutina de sueño detectados a través del acelerómetro pueden dar indicios tempranos de un cuadro depresivo o un episodio maníaco.

- **Monitoreo Continuo y Prevención de Recaídas:** La medición pasiva y permanente de señales de comportamiento podría servir para alertar a clínicos o cuidadores sobre riesgos inminentes de descompensación.
- **Personalización de Tratamientos:** A partir de los datos recogidos, es posible ofrecer intervenciones a medida, ya sea a través de notificaciones oportunas en el *smartphone* o consejos automatizados de autocuidado.

### Salud Pública y Epidemiología

- **Seguimiento de Crisis Sanitarias:** El fenotipado digital se ha utilizado para comprender cómo se desplazan y se comunican las personas durante epidemias o situaciones de emergencia.
- **Monitorización de Hábitos:** Los datos agregados de grandes grupos de usuarios pueden ayudar a identificar patrones de comportamiento que influyan en la salud de la población.

### Ámbitos Sociales y de Investigación en Comportamiento

- **Psico informática:** Convergencia entre psicología e informática para estudiar variables de personalidad, conductas de interacción virtual y efectos de las redes sociales.
- **Nuevas Herramientas de Big Data:** La gran cantidad de información captada puede analizarse usando técnicas de *machine learning*, descubriendo correlaciones antes inaccesibles y posibilitando aplicaciones de predicción en campos como la educación, el marketing responsable y la detección de comportamientos de riesgo.

En todos los casos, se recalca la importancia de salvaguardar la privacidad y la ética en la gestión de datos. El documento detalla marcos como el GDPR europeo, el HIPAA estadounidense y modelos específicos de protección y anonimización de la información.

## 1.4. Estructura de la tesis y desarrollo de un prototipo

Para abordar la complejidad del fenotipado digital, la tesis se organiza en varios capítulos. Además de una revisión teórica extensa sobre los conceptos, las tecnologías y la historia del

término, se profundiza en los lineamientos éticos y en aspectos metodológicos como la arquitectura de recolección de datos, los sensores empleados y las técnicas de análisis.

Un elemento clave del trabajo es la implementación de un prototipo funcional, cuyo objetivo es validar los principios teóricos y mostrar cómo la captura de datos puede realizarse de manera segura y respetuosa con la privacidad de los usuarios. Este prototipo, concebido de forma modular, reúne los siguientes componentes principales:

- **Aplicación móvil:** Encargada de recolectar variables de comportamiento (uso de aplicaciones, eventos de pantalla, llamadas, ubicación, etc.) y datos fisiológicos o contextuales (de sensores como la cámara, el micrófono y el GPS).
- **Servidor y base de datos central:** responsable del almacenamiento seguro, la anonimización y la administración de los datos.
- **Reporte:** Responsable de la generación y presentación de informes basados en los datos recolectados. Incluye mecanismos de visualización para facilitar el análisis.

La construcción y puesta en marcha de este prototipo son fundamentales para demostrar la viabilidad de las ideas expuestas, así como para abrir nuevas líneas de investigación y desarrollo en el futuro.

## 1.5. Conclusión

En síntesis, el fenotipado digital ofrece un abanico amplio de posibilidades para estudiar y mejorar la salud de las personas y los grupos. Su aporte central radica en combinar:

- Monitoreo continuo y pasivo de información (datos de comportamiento y contexto),
- Integración con métodos de análisis avanzados (*machine learning*, *big data*),
- Intervenciones y seguimientos clínicos más ágiles y personalizados.

No obstante, la tesis también pone de relieve que este enfoque acarrea desafíos trascendentes en materia de privacidad, consentimiento informado, calidad de datos y responsabilidad ética. Se requiere, por tanto, una aproximación interdisciplinaria que involucre tanto a ingenieros y científicos de datos como a profesionales de la salud, expertos legales y autoridades regulatorias.

En los siguientes capítulos, se profundiza en los fundamentos teóricos, las tecnologías disponibles y las consideraciones éticas que acompañan al fenotipado digital. Con ello, se

establecen las bases para el desarrollo y la explicación del prototipo que se ha construido como parte esencial de esta investigación, lo que permite exhibir de manera tangible el potencial de esta aproximación para el bienestar y la comprensión del ser humano en entornos digitales.

## 2. Conceptos fundamentales sobre fenotipado digital

El fenotipado digital se ha posicionado como una herramienta clave en el análisis de datos derivados de interacciones humanas, especialmente en campos como la salud, la tecnología y la inteligencia artificial. Para entender su importancia, es esencial revisar las definiciones más relevantes que lo describen.

### 2.1. Fenotipado digital

El término “fenotipado digital” (*digital phenotyping*) fue acuñado en 2015 por Jain et al., inspirándose en la noción de “fenotipo extendido” de Richard Dawkins (1982). Según se describe en el trabajo de [1], el fenotipo digital se conceptualiza como un fenotipo extendido que integra datos de actividad y comportamiento en línea con información clínica, con el propósito de optimizar el diagnóstico y el pronóstico médico.

La definición más citada en la literatura describe el fenotipado digital como la “cuantificación momento a momento del fenotipo humano a nivel individual mediante datos provenientes de dispositivos digitales personales”. Según se observa en [7], el enfoque busca combinar la información generada a través de la actividad en línea y otros comportamientos capturables con datos clínicos, con miras a mejorar tanto el diagnóstico como el pronóstico en medicina y salud mental.

En [1] también señala que este concepto ha evolucionado para incluir elementos como informes digitales autogenerados y mediciones fisiológicas obtenidas mediante sensores digitales, lo que amplía su alcance hacia una caracterización más rica e integrada del individuo.

Además [1] menciona que el fenotipado digital se utiliza cada vez más como una medida clínica para la toma de decisiones en atención sanitaria. [1] pone el ejemplo de *Mindstrong Health Services*, que utiliza una aplicación en *smartphones* para crear biomarcadores digitales y proporcionar servicios de telemedicina a individuos con enfermedades mentales graves. Los biomarcadores digitales pueden detectar el deterioro clínico temprano no reportado por el paciente, permitiendo intervenciones de baja intensidad a través de la telemedicina antes de que ocurra una crisis.

## 2.2.Contexto

Desde la perspectiva de Dourish “el contexto es una noción resbaladiza...”. El contexto puede considerarse como algo que rodea la actividad y que puede convertirse en contenido dependiendo del enfoque de la actividad [2].

Lieberman y Selker lo ven desde la perspectiva de la programación informática. "Se puede considerar que el contexto es todo lo que afecta el cálculo, excepto la entrada y salida explícitas".

Según Dey et al, “El contexto es cualquier información que pueda utilizarse para caracterizar la situación de una entidad. Una entidad puede ser una persona, lugar u objeto que se considere relevante para la interacción entre un usuario y una aplicación, incluidos el usuario y las aplicaciones mismas” [2].

Según Gray y Salber, lo definen “El contexto percibido son propiedades que caracterizan un fenómeno. Se detectan y son potencialmente relevantes para las tareas respaldadas por una aplicación o los medios por los cuales se realizan esas tareas.” [2].

Dadas estas definiciones podemos armar nuestra propia definición. Cualquier información que rodea y afecta a una entidad, ya sea una persona, un lugar o un objeto, y que resulta relevante para la interacción entre un usuario y un sistema se conoce como contexto. Este contexto abarca tanto el entorno objetivo externo como el estado subjetivo interno del usuario, y requiere que el usuario proporcione datos durante la interacción. Exceptuando las entradas y salidas explícitas, el contexto incluye todo lo que afecta el cálculo y las tareas, y se caracteriza por características detectables que pueden ser relevantes para las aplicaciones y los medios mediante los cuales se realizan las tareas.

## 2.3.Context Aware

Un sistema es *Context-Aware* si utiliza el contexto para proporcionar información y servicios relevantes al usuario, donde la relevancia depende de la tarea de los usuarios [2].

En cuanto a aplicaciones *context-aware*, se refiere a aplicaciones diseñadas para adaptarse y responder a las condiciones y situaciones del entorno en el que se utilizan. Estas aplicaciones toman en cuenta información como la identidad del usuario, el tiempo y la ubicación para

presentar información relevante, ejecutar servicios automáticamente y etiquetar el contexto para su posterior recuperación.

Una aplicación *context-aware* es aquella que puede adaptarse y responder dinámicamente a las circunstancias y condiciones del entorno en el que se utiliza, utilizando información contextual relevante para proporcionar servicios personalizados y mejorar la experiencia del usuario [2].

## **2.4.Mobile sensing**

El concepto de *Mobile Sensing* hace referencia a un enfoque metodológico que aprovecha dispositivos y plataformas digitales para la recopilación de datos sobre el comportamiento humano [3]

Por su parte, [4] amplía esta definición al destacar que *Mobile Sensing* implica el uso de dispositivos móviles, en particular *smartphones*, para la captura y almacenamiento de datos. Estos dispositivos, equipados con una variedad de sensores como GPS, acelerómetros y giroscopios, permiten una recolección continua de información tanto del entorno como del comportamiento del usuario, sin requerir una intervención activa.

Esta metodología resulta especialmente relevante en el campo de la psico informática, ya que permite obtener mediciones objetivas del comportamiento y del contexto del usuario, proporcionando un marco robusto para estudios y aplicaciones en áreas como la salud, la inteligencia artificial y el análisis de datos personales [4].

Además, [4] menciona que las redes sociales ubicuas (USN) combinan tecnologías de computación ubicua con redes sociales tradicionales, permitiendo interacciones sociales continuas y personalizadas en diversos contextos. Estas redes aprovechan datos contextuales recogidos por sensores en *smartphones*, como la ubicación y la actividad del usuario, para adaptar y enriquecer la experiencia social.

En el desarrollo de las USN, el *Mobile Sensing* desempeña un papel crucial, ya que permite la recolección continua y pasiva de datos del usuario sin requerir acciones explícitas por parte de este. Los sensores en dispositivos móviles capturan información detallada sobre la ubicación, los movimientos y las actividades cotidianas del usuario. Estos datos se utilizan para personalizar la experiencia social, ofreciendo recomendaciones y sugerencias basadas en el contexto actual del usuario [4].

## **2.5. Personalidad digital**

Según [5], la personalidad digital se refiere a la identidad y personalidad única que los individuos desarrollan y presentan en el ámbito digital. Este concepto incluye las características, comportamientos y formas de autoexpresión que las personas exhiben a través de su presencia en línea, interacciones y actividades en diversas plataformas digitales y entornos virtuales. Los individuos tienen la capacidad de construir y curar sus personalidades digitales mediante la selección cuidadosa de la información que comparten, la representación de rasgos específicos y su participación en actuaciones virtuales.

El papel de la personalidad digital es particularmente relevante en la era digital, ya que influye directamente en cómo las personas se presentan, interactúan con los demás y navegan por el entorno en línea. Además de ser un medio de autoexpresión y construcción de identidad, permite a los individuos moldear y gestionar sus personajes en línea. Esta dimensión es crucial para la autopresentación, las interacciones sociales y la formación de comunidades en plataformas digitales. Asimismo, la personalidad digital tiene implicaciones importantes en la marca personal y la identidad profesional, ya que permite a los individuos destacar sus habilidades y establecer una presencia significativa en línea [5].

Por otra parte, [5] destaca los efectos psicológicos y sociales asociados a la personalidad digital, los cuales son de gran alcance. A nivel individual, la presión por mantener una imagen curada en línea puede generar un aumento del estrés, ansiedad y una disminución de la autoestima. La exposición constante a versiones idealizadas de la vida de otras personas puede fomentar sentimientos de insuficiencia, contribuyendo a una cultura de comparación y dudas personales. Además, la difuminación de los límites entre el mundo digital y el real puede afectar negativamente las interacciones sociales y las relaciones, generando un sentido de desconexión y soledad.

En un nivel social, la personalidad digital influye en las normas y valores de la sociedad, moldeando la manera en que las personas se perciben a sí mismas y a los demás. Esto subraya su impacto tanto a nivel individual como colectivo en el mundo interconectado de hoy [5].

## **2.6. Sensing architecture**

Según [5], el término *Sensing Architecture* se refiere al marco e infraestructura que permite la recolección, interpretación y análisis de datos relacionados con el comportamiento del usuario,

sus preferencias y las señales ambientales en entornos digitales. Esta arquitectura combina sensores, métodos de recolección de datos y algoritmos para capturar y procesar información.

La *Sensing Architecture* implica la integración de múltiples sensores, como cámaras, micrófonos, acelerómetros y GPS, en dispositivos o plataformas digitales. Estos sensores permiten supervisar la recolección de datos provenientes de las interacciones del usuario y su entorno. Por ejemplo, un *smartphone* puede utilizar sensores para detectar gestos táctiles, capturar fotografías o videos y medir movimientos corporales. Los datos capturados por los sensores son procesados y analizados mediante algoritmos y técnicas de aprendizaje automático, los cuales convierten los datos sensoriales en bruto en información útil, como patrones y correlaciones. Por ejemplo, los algoritmos pueden analizar los movimientos del usuario para identificar estados emocionales o interpretar datos ambientales para ofrecer sugerencias personalizadas.

La *Sensing Architecture* consta de diversos componentes y funcionalidades que trabajan en conjunto para recopilar y analizar datos en entornos digitales. Entre estos componentes se encuentran los sensores, las técnicas de recolección de datos, los algoritmos de procesamiento de datos y los sistemas de retroalimentación.

- **Sensores:** Los sensores son esenciales en la *Sensing Architecture*, ya que recopilan distintos tipos de datos relacionados con el entorno y las interacciones humanas. Estos sensores están presentes en dispositivos como *smartphones*, dispositivos portátiles y plataformas digitales, e incluyen cámaras, micrófonos, acelerómetros, giroscopios y sensores de proximidad, entre otros. Cada tipo de sensor recoge diferentes tipos de datos, como visuales, de audio, de movimiento o ambientales.
- **Métodos de recolección de datos:** Las estrategias utilizadas para recopilar datos de los sensores se denominan métodos de recolección de datos. Estos determinan la frecuencia, duración y contexto de la recolección de datos, y pueden incluir transmisión en tiempo real, muestreo periódico, captura activada por eventos y recolección iniciada por el usuario.
- **Métodos de procesamiento de datos:** Una vez adquiridos los datos, se aplican métodos de procesamiento que incluyen algoritmos de aprendizaje automático, modelos estadísticos y técnicas de reconocimiento de patrones. Estos algoritmos interpretan los datos sensoriales crudos para identificar patrones, correlaciones y anomalías,

extrayendo información relevante sobre el comportamiento del usuario, sus preferencias, emociones o elementos del entorno.

- **Mecanismos de retroalimentación:** Los sistemas de retroalimentación son fundamentales en el diseño de la *Sensing Architecture*, ya que permiten mejorar los procesos de recolección y análisis de datos. Los usuarios pueden proporcionar retroalimentación directamente a través de entradas explícitas o sistemas automatizados, que ajustan la arquitectura de sensores en función de las interacciones y resultados observados. Este ciclo continuo de retroalimentación contribuye a mejorar la precisión y eficacia del sistema con el tiempo.

Finalmente, [5] describe cómo la *Sensing Architecture* lleva a cabo funciones clave como la adquisición de datos, el preprocesamiento, la extracción de características, la fusión de datos y la toma de decisiones. La captura de datos consiste en la recopilación de información sensorial en bruto, mientras que el preprocesamiento incluye tareas como el filtrado, limpieza y normalización de los datos para prepararlos para el análisis. La extracción de características identifica los aspectos más relevantes de los datos preprocesados, los cuales son utilizados como insumos para análisis posteriores o procesos de toma de decisiones.

## 2.7. Tipos de sensores

Según [5], en el análisis de la personalidad digital se utilizan diversos sensores para capturar y recopilar datos relacionados con el comportamiento, las preferencias y el entorno circundante de los individuos. Estos sensores desempeñan un papel crucial en la comprensión y construcción de personalidades digitales. Entre los más comunes se encuentran las cámaras, los micrófonos y los sensores GPS, aunque también se emplean otros sensores específicos en función de las necesidades del análisis.

### Cámaras

Las cámaras se utilizan ampliamente para capturar datos visuales en entornos digitales. Estas permiten grabar imágenes y videos, lo que facilita a los investigadores el análisis de expresiones faciales, gestos y otras señales visuales. A través de estas observaciones, las cámaras ofrecen valiosos conocimientos sobre emociones, comunicación no verbal e interacciones físicas dentro del ámbito digital.

### Micrófonos

Los micrófonos son sensores esenciales para capturar datos de audio. Estos permiten analizar patrones de habla, tono y contenido. Al examinar los datos de audio, los investigadores pueden obtener información sobre características vocales, patrones lingüísticos e incluso estados emocionales. Los micrófonos, además, contribuyen a la interpretación de interacciones habladas y señales auditivas en contextos digitales.

### Sensores GPS

Los sensores GPS, combinados con servicios basados en la ubicación, proporcionan información sobre el contexto geográfico de los individuos. Capturan datos relacionados con ubicaciones, movimientos y preferencias de lugares específicos. Gracias a estos sensores, los investigadores pueden comprender los patrones de movilidad, preferencias espaciales y la influencia de la ubicación en los comportamientos digitales.

### Otros sensores

Además de los sensores mencionados, existen otros tipos que desempeñan un papel importante en el análisis de la personalidad digital:

- **Acelerómetros:** Presentes en *smartphones* y dispositivos portátiles, detectan movimientos y miden cambios en la orientación. Proporcionan información sobre actividades físicas, gestos y movimientos de los individuos dentro de entornos digitales.
- **Biosensores:** Incluyen monitores de ritmo cardíaco y sensores electro dermales, entre otros. Capturan datos fisiológicos, como ritmo cardíaco, conductancia de la piel y patrones de sueño. Estos sensores permiten a los investigadores analizar estados emocionales, niveles de estrés y el bienestar general de los individuos.
- **Sensores ambientales:** Estos sensores, como los de temperatura y humedad, recopilan datos sobre el entorno físico. Proporcionan información contextual que influye en los comportamientos y preferencias de los individuos dentro de espacios digitales, ayudando a comprender el impacto del entorno físico en las personalidades digitales.

En conjunto, estos sensores y sus capacidades de recolección de datos permiten una visión integral del comportamiento humano en entornos digitales, haciendo posible un análisis más profundo y preciso de la personalidad digital [5].

## 2.8.Registro semántico

El registro semántico se refiere al proceso de captura y análisis de las interacciones de los usuarios con el objetivo de inferir contexto, sentimientos e intenciones [5]. A través del procesamiento de datos provenientes de diversas fuentes digitales, esta técnica permite extraer información significativa sobre el comportamiento y las emociones de los individuos [5].

La importancia del registro semántico radica en su capacidad para mejorar la interacción entre los usuarios y las plataformas digitales. Al incorporar este tipo de análisis, las personalidades digitales pueden generar respuestas más personalizadas y contextualmente relevantes, optimizando la experiencia del usuario en entornos digitales [5].

## 2.9.Clínica digital y Monitoreo Continuo

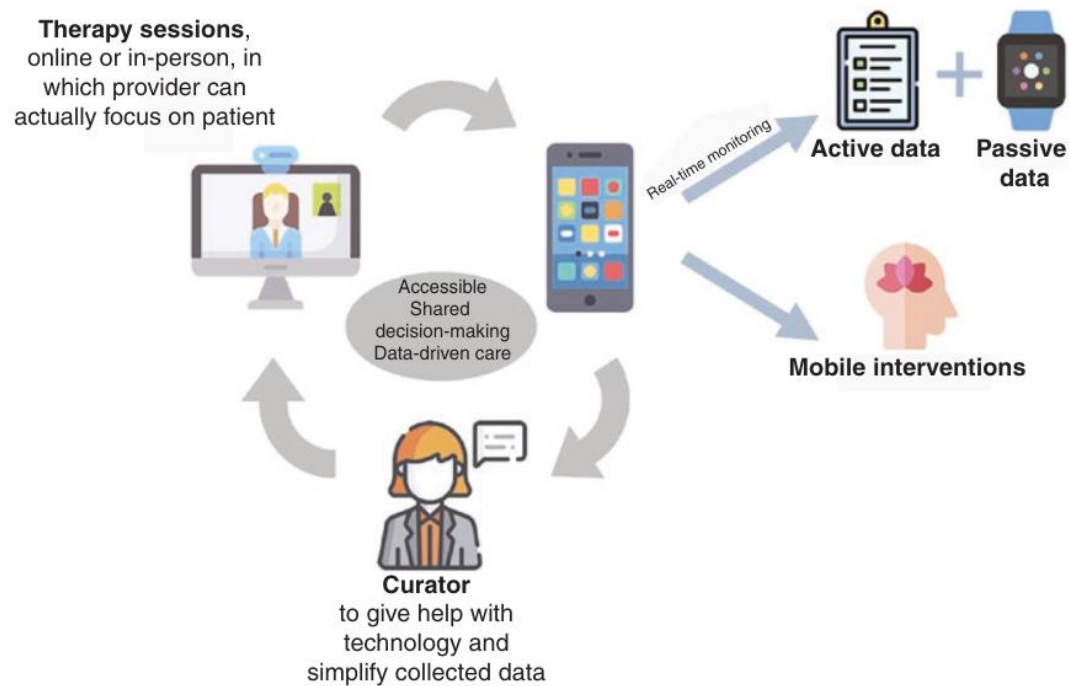


Ilustración 1 Clínica digital

Según [6], hasta ahora los profesionales de la salud han dependido principalmente de los síntomas que los pacientes recuerdan mencionar durante sus consultas. Sin embargo, los síntomas experimentados entre dos citas médicas suelen estar sujetos a sesgos, ya que muchos de ellos son olvidados o reportados de manera imprecisa. Además, los diagnósticos y tratamientos se basan en gran medida en la interpretación subjetiva de los clínicos sobre esta información, lo que puede comprometer la precisión del abordaje médico. En este contexto, el

monitoreo continuo del paciente se presenta como una solución que permite a los profesionales de la salud acceder a datos más precisos sobre la condición real de sus pacientes. Asimismo, tanto el automonitoreo como las intervenciones móviles pueden empoderar a los pacientes en la gestión de su propio tratamiento [6].

Uno de los cambios esperados en la evolución hacia la Clínica Digital es la inclusión del monitoreo en tiempo real de los síntomas mediante teléfonos inteligentes y otros sensores. Este proceso puede implementarse de dos formas:

**Monitoreo activo:** El paciente participa activamente completando cuestionarios o reportando cambios en su sintomatología a través de una aplicación móvil.

**Monitoreo pasivo:** Se utilizan datos capturados automáticamente por sensores como la medición del ritmo cardíaco, el análisis de patrones de voz mediante micrófonos o el conteo de pasos registrados en dispositivos móviles.

A través de estos mecanismos, no solo los clínicos obtienen información más precisa y confiable sobre sus pacientes, sino que también pueden utilizar calculadoras clínicas y herramientas analíticas avanzadas para tomar decisiones médicas basadas en datos [6].

Para garantizar una interpretación más intuitiva de los datos recopilados, se espera la incorporación de curadores de datos que actúen como intermediarios entre los pacientes y los clínicos. Según [6], estos curadores son principalmente científicos e ingenieros con experiencia en *big data*, ciencias de la salud, modelado predictivo y análisis de datos. Su función es acceder a los datos recopilados por los dispositivos de los pacientes fuera de las consultas médicas y generar representaciones visuales a través de gráficos y modelos analíticos. Esto facilita una interpretación más clara y estructurada de la información, proporcionando nuevas perspectivas clínicas.

La disponibilidad de datos curados permitirá a los proveedores de salud optimizar las modificaciones en los tratamientos de manera más eficiente y mejorar la comunicación con los pacientes. Además, estos datos están diseñados para ser comprensibles tanto para los profesionales médicos como para los pacientes sin formación especializada, fomentando una mayor participación en la toma de decisiones sobre su salud [6].

## 2.10. Datos activos y pasivos

Los datos activos son aquellos proporcionados directamente por el usuario, como al realizar una búsqueda en internet o responder cuestionarios [6]. En contraste, los datos pasivos se refieren a la captura de información sin requerir la participación activa del usuario, limitándose únicamente a la autorización para acceder a datos contextuales almacenados en sus dispositivos [6].

Según [1], existen dos técnicas principales de captura de datos: activa y pasiva. La captura activa implica la participación directa del usuario, como completar encuestas o diarios electrónicos. Este método permite obtener datos específicos y detallados, aunque puede resultar invasivo y depender del cumplimiento del usuario. En contraste, la captura pasiva recolecta datos automáticamente a través de sensores integrados en dispositivos móviles, como el seguimiento de la ubicación GPS y los patrones de actividad física. La combinación de ambas técnicas proporciona una imagen más completa y precisa del comportamiento y estado mental de los individuos [1].

Los datos pasivos constituyen la fuente primaria de información para el fenotipado digital, ya que permiten obtener datos detallados sobre la vida diaria y el estado de ánimo del paciente [6]. Por ejemplo, los datos de GPS pueden revelar cómo varía el número de ubicaciones visitadas y el tiempo pasado en diferentes lugares a lo largo del tiempo, en relación con episodios de cambios en el estado de ánimo [6]. De manera complementaria, los datos registrados en el acelerómetro pueden aportar información sobre los patrones de movilidad física y el gasto metabólico [6].

Asimismo, los registros de voz obtenidos a través de micrófonos incorporados pueden utilizarse para detectar marcadores vocales de estado de ánimo, identificando síntomas de episodios emocionales, como tristeza o excitación [6]. A su vez, el uso del teléfono inteligente en sí mismo constituye una fuente valiosa de información: los eventos de encendido y apagado de la pantalla podrían indicar ansiedad, mientras que los patrones de sociabilidad pueden inferirse a partir de la interacción con aplicaciones de redes sociales, mensajes de texto SMS, registros de llamadas y el uso de aplicaciones de chat [6].

El consumo de multimedia y contenido digital también puede ofrecer pistas sobre el estado interno de una persona. Factores como los estilos de música reproducidos en plataformas de transmisión, el tiempo dedicado a ver series de televisión en exceso o el tipo de productos

adquiridos mediante aplicaciones de entrega pueden proporcionar información relevante sobre su bienestar emocional [6]. Además, los datos fisiológicos, como la frecuencia cardíaca, pueden ser capturados por dispositivos portátiles e integrarse al conjunto de información clínica [6].

En contraste, los datos activos requieren la participación consciente de los usuarios para su recopilación. Aunque estos métodos no presentan la misma innovación que los datos pasivos, constituyen una fuente confiable de información complementaria para el fenotipado digital, proporcionando validez convergente a los datos obtenidos por otros medios [6].

### **2.11. Monitoreo Pasivo**

El monitoreo pasivo ha emergido como una herramienta revolucionaria en la evaluación de los trastornos del estado de ánimo, especialmente en el diagnóstico y tratamiento del trastorno bipolar [1]. Tradicionalmente, los métodos de evaluación dependen de entrevistas clínicas y cuestionarios, los cuales son subjetivos y capturan información en puntos específicos del tiempo [1]. Esta falta de continuidad y objetividad en la recolección de datos puede derivar en diagnósticos erróneos y tratamientos inadecuados [1].

En contraste, el monitoreo pasivo permite la recolección continua de datos a través de dispositivos que los usuarios ya utilizan en su vida diaria, como teléfonos móviles y dispositivos portátiles, proporcionando una evaluación más completa y precisa de los estados afectivos y cognitivos de los pacientes [1].

El diagnóstico del trastorno bipolar se ha basado tradicionalmente en métodos esporádicos y subjetivos, como entrevistas clínicas y cuestionarios [1]. Sin embargo, la dependencia de la autoevaluación del paciente y la ausencia de un monitoreo continuo limitan la capacidad de capturar la naturaleza fluctuante de los síntomas de este trastorno [1].

### **2.12. Biomarcadores y Biomarcadores Digitales**

Los biomarcadores se describen como características individuales cuantificables que permiten identificar procesos patológicos [6]. Estos incluyen medidas anatómicas, bioquímicas y fisiológicas, utilizadas para la detección y prevención de enfermedades, así como para evaluar la respuesta a intervenciones terapéuticas [6].

En el contexto digital, surge el término biomarcadores digitales, el cual hace referencia a marcadores de procesos individuales obtenidos mediante sensores y herramientas

computacionales [6]. Estos representan una extensión de los biomarcadores tradicionales, adaptados a la era digital, facilitando el monitoreo continuo de variables relevantes para la salud [6].

Asimismo, los biomarcadores digitales han demostrado ser especialmente prometedores como indicadores de episodios depresivos, permitiendo una evaluación más objetiva de los cambios en el estado emocional y el comportamiento de los individuos [6].

### **2.13. Biomarcador de Control Cognitivo**

El control cognitivo se refiere a la capacidad de un individuo para regular y dirigir sus procesos de pensamiento de manera intencional y consciente [1]. Este concepto abarca habilidades clave como la atención sostenida, la inhibición de respuestas impulsivas, la flexibilidad cognitiva (capacidad de alternar entre tareas o pensamientos) y la planificación [1].

En el contexto del fenotipado digital, el biomarcador de control cognitivo se evalúa a través del tiempo de reacción en interacciones con dispositivos móviles, como toques y deslizamientos en la pantalla del *smartphone*. Además, la variabilidad en estos tiempos de reacción proporciona información sobre la capacidad del individuo para mantener y ajustar su atención y respuestas a estímulos de manera consistente a lo largo del tiempo [1].

### **2.14. Biomarcador de Memoria de Trabajo**

La memoria de trabajo es un sistema cognitivo que permite a las personas mantener y manipular temporalmente información para la realización de tareas cognitivas complejas, como la comprensión, el aprendizaje y el razonamiento [1]. Este tipo de memoria es fundamental para actividades que requieren la retención y uso de información en tiempo real, como recordar un número de teléfono mientras se marca o seguir instrucciones complejas [1].

En el contexto del fenotipado digital, el biomarcador de memoria de trabajo se evalúa mediante tareas que requieren recordar y manipular información presentada en el *smartphone*, como patrones de toques específicos que deben ser repetidos en una secuencia determinada [1].

### **2.15. Consentimiento Informado**

En el ámbito sanitario, el fenotipado digital se considera un indicador de riesgo que puede aportar información relevante para la toma de decisiones clínicas [1]. En los Estados Unidos,

el consentimiento informado, que debe explicar adecuadamente los riesgos y beneficios involucrados, es un requisito suficiente para recibir servicios de atención sanitaria [1]. Sin embargo, en Europa, bajo el Reglamento General de Protección de Datos (GDPR), el consentimiento informado por sí solo puede no ser suficiente para el procesamiento de datos personales en el ámbito sanitario [1].

El GDPR o "Reglamento (UE) 2016/679" regula el procesamiento de datos personales de personas físicas dentro de la Unión Europea. Esto incluye actividades como la recopilación, almacenamiento y análisis de datos personales [1]. Además, el GDPR impone numerosos requisitos a quienes procesan datos (procesadores de datos) o a quienes tienen responsabilidad sobre ellos (controladores de datos) [1]. Los sujetos de datos, es decir, las personas a las que se refieren los datos, cuentan con un amplio conjunto de derechos, como el derecho a ser informados y el derecho a solicitar la eliminación de sus datos procesados [1].

Por defecto, cualquier procesamiento de datos personales requiere el consentimiento informado del sujeto de datos, salvo en ciertas excepciones, como cuando los datos son necesarios por razones legales o para cumplir un contrato con la persona [1]. En estudios psicológicos, el consentimiento informado suele ser la base para la recopilación y procesamiento de datos personales. No obstante, este consentimiento no exime al controlador o procesador de datos del cumplimiento de las obligaciones impuestas por el GDPR, las cuales incluyen derechos como el acceso a los datos en un formato portátil y la oposición al procesamiento de datos, incluso después de haber sido otorgado el consentimiento [1].

En los Estados Unidos, el fenotipado digital en el ámbito sanitario se considera información de salud protegida (PHI). Su acceso y uso están regulados por normativas estatales y federales [1]. La Ley de Portabilidad y Responsabilidad de Seguros de Salud (HIPAA), promulgada en 1996, establece disposiciones de privacidad y seguridad de datos con el fin de proteger la información médica [1]. La Regla de Privacidad de HIPAA permite que las entidades cubiertas, como proveedores de atención médica y planes de salud, utilicen y divulguen información PHI identificable sin necesidad del consentimiento del individuo cuando sea para fines de tratamiento [1].

En [4] se menciona que la recolección de datos mediante dispositivos móviles proporciona información detallada y sensible sobre los usuarios, lo que plantea preocupaciones significativas en materia de privacidad. Estos datos pueden incluir ubicación del usuario,

patrones de comportamiento e interacciones sociales, lo que requiere la implementación de medidas estrictas de protección para garantizar un uso seguro y ético de la información personal [4].

El cumplimiento de regulaciones de privacidad, como el Reglamento General de Protección de Datos (GDPR) de la Unión Europea, es fundamental en este contexto [4]. Estas normativas estipulan que se debe obtener el consentimiento explícito de los usuarios antes de la recopilación de datos y que se debe proporcionar información clara y comprensible sobre qué datos serán recopilados, cómo se utilizarán y quién tendrá acceso a ellos [4].

El consentimiento informado es un principio clave en este proceso. Los usuarios deben comprender de manera clara qué datos se están recolectando y con qué propósito, así como tener la capacidad de aceptar o rechazar su participación en cualquier momento [4]. Proporcionar información accesible y transparente es esencial para permitir que los usuarios tomen decisiones informadas sobre su participación en la recolección de datos.

Para garantizar la protección de la información recolectada, es fundamental la implementación de medidas de seguridad robustas, tales como la anonimización y el cifrado de datos [4]. Estas técnicas ayudan a minimizar los riesgos, asegurando que los datos no puedan ser fácilmente asociados con individuos específicos, incluso en caso de accesos no autorizados [4].

## **2.16. Modelo de Privacidad para la Recolección de Datos Móviles**

El Modelo de Privacidad para la Recolección de Datos Móviles (PM-MoDaC, por sus siglas en inglés) consta de nueve medidas diseñadas para minimizar los riesgos de privacidad y proporcionar a los usuarios mayor control sobre sus datos [4]. Estas medidas incluyen:

- **Consentimiento del Usuario (PM A):** Antes de instalar la aplicación, se debe informar al usuario sobre qué datos se recopilan y con qué propósito. Este aspecto se cubre en una política de privacidad, que el usuario debe aceptar antes de usar la aplicación [4].
- **Permitir a los Usuarios Ver sus Propios Datos (PM B):** Los usuarios deben poder visualizar los datos que se están recopilando sobre ellos, lo que les permite tomar decisiones informadas sobre su compartición. También deben contar con la opción de exportar todos los datos almacenados sobre ellos [4].

- **Opción de Optar por No Participar (PM C):** Los usuarios deben tener la posibilidad de optar por no participar y solicitar la eliminación de sus datos en cualquier momento. Esta medida es fundamental para respetar la autonomía del usuario [4].
- **Aprobación por la Comisión de Ética (PM D):** Los estudios psicológicos o médicos que involucren la recolección de datos generalmente requieren aprobación previa por parte de una comisión de ética o un comité de revisión, el cual supervisa y valida las prácticas de privacidad aplicadas en la investigación [4].
- **Identificadores Aleatorios (PM E):** En lugar de utilizar un inicio de sesión tradicional, que podría vincular datos sensibles con detalles personales, se implementan identificadores aleatorios para proteger la privacidad del usuario [4].
- **Anonimización de Datos (PM F):** Se emplean técnicas de anonimización de datos, como funciones hash unidireccionales, para ofuscar información identificable, como números de teléfono y direcciones Bluetooth, minimizando el riesgo de exposición de datos personales [4].
- **Sistema de Permisos (PM G):** El sistema de permisos de Android permite a los usuarios controlar qué datos o sensores son accesibles por una aplicación, facilitando una solicitud de permisos clara y comprensible [4].
- **Transferencia Segura (PM H):** La protección de la información durante la transmisión de datos es esencial. Para ello, se utilizan conexiones seguras que impiden accesos no autorizados al transferir datos del dispositivo del usuario al servidor [4].
- **Identificación sin Vínculo a Datos Recopilados (PM I):** Los sistemas deben ser capaces de identificar a los usuarios sin vincularlos directamente a los datos recopilados, asegurando la separación entre identidad y datos sensibles [4].

El PM-MoDaC fue implementado y evaluado en la aplicación TYDR (*Track Your Daily Routine*), una herramienta diseñada para la investigación en psico informática, que recopila datos de uso y sensores de *smartphones* [4]. Durante su implementación, se integraron las nueve medidas del modelo para garantizar la privacidad de los usuarios.

La evaluación del PM-MoDaC se llevó a cabo mediante un estudio que analizó la disposición de los usuarios a compartir sus datos y cómo esta variabilidad estaba influenciada por factores demográficos y rasgos de personalidad. Los resultados indicaron que los usuarios más jóvenes eran menos propensos a compartir sus datos, lo que destaca la importancia de adaptar las estrategias de privacidad según el grupo demográfico [4].

El PM-MoDaC no se limita únicamente a la psico informática, sino que está diseñado para ser aplicable en otros ámbitos donde la privacidad en la recolección de datos móviles es una preocupación fundamental [4].

### **3. Origen y evolución histórica**

A lo largo de las últimas décadas, el concepto de “fenotipado digital” ha tomado forma a partir de la conjunción de ideas previas en biología evolutiva y de los rápidos avances tecnológicos en dispositivos móviles, análisis de datos y aprendizaje automático. Para comprender plenamente la historia y evolución de este término, es fundamental revisar sus raíces conceptuales en la biología, su formalización en la práctica clínica y de investigación, así como su creciente impacto en la psicología y la psiquiatría.

#### **3.1. Fundamentos conceptuales del “Fenotipado extendido” al “Fenotipado digital”**

La historia del fenotipado digital se remonta, al menos conceptualmente, a la obra del biólogo evolutivo Richard Dawkins en 1982, Dawkins acuñó el término “fenotipo extendido” (“*extended phenotype*”), destacando que el fenotipo no debía limitarse únicamente a los rasgos biológicos directamente observables, sino que incluía también aquellas manifestaciones conductuales y las interacciones del organismo con su entorno. De este modo, la conducta y las repercusiones ambientales forman parte de la expresión fenotípica.

Aunque esta idea se planteó en el ámbito de la biología evolutiva, sentó las bases para entender que el “rastreo” o las manifestaciones de un individuo pueden trascender lo meramente físico. Esta concepción se convertiría años después en el sustento teórico para el “fenotipado digital”, al extender el concepto de fenotipo a los datos y comportamientos que las personas generan a través de sus interacciones digitales.

#### **3.2. Aparición formal del término y definiciones iniciales**

Tal como se describe en la Sección 1.1.1, el “fenotipado digital” surge formalmente en 2015 con la publicación de Jain et al. Sin embargo, el germen teórico proviene de la idea del “fenotipo extendido” de Dawkins.

A raíz de esta base conceptual, en [7], se refuerza la idea de que la mayor parte de la recopilación de datos para el fenotipado digital ocurre de manera pasiva (*passive sensing*). Es decir, los individuos no tienen que ingresar manualmente información, sino que sus dispositivos (teléfonos inteligentes, relojes o pulseras inteligentes, etc.) capturan datos en segundo plano.

De esta forma, se logra una medición más objetiva y continua del comportamiento humano cotidiano.

### **3.3. Ventajas tempranas y ejemplos iniciales**

Desde sus primeras aplicaciones, el fenotipado digital mostró ventajas importantes sobre los métodos tradicionales de medición del comportamiento y la salud mental. En [7], se señalan algunos ejemplos tempranos que evidenciaron el potencial de esta aproximación:

Detección involuntaria de temblores a través del movimiento del cursor del ratón (mouse), lo que ofrecía pistas sobre la presencia o severidad de ciertas alteraciones motoras.

Uso de consultas en motores de búsqueda (ej. Microsoft Bing) para la detección de condiciones neurodegenerativas, basándose en los patrones de interacción y búsqueda de información.

Estos casos representan el potencial del fenotipado digital para identificar biomarcadores conductuales o cognitivos a gran escala, de manera no intrusiva y con un grado de inmediatez y alcance sin precedentes.

### **3.4. Expansión en psiquiatría y salud mental**

Objetivo principal en psiquiatría

La adopción del fenotipado digital en psiquiatría surgió como respuesta a las limitaciones de las técnicas tradicionales de evaluación, como cuestionarios, entrevistas clínicas o escalas estandarizadas que suelen depender de la memoria retrospectiva, están sujetas a sesgos de deseabilidad social o no permiten un seguimiento continuo.

*“Digital phenotyping is free from these limitations and has the potential to provide a new and much richer perspective on human behavior.”* [8].

La posibilidad de recopilar datos a través de dispositivos personales (incluidos *smartphones* y *wearables*) permite monitorear a los pacientes en tiempo real y en su entorno cotidiano, resultando en flujos de datos objetivos y de alta resolución.

Integración con métodos tradicionales y superación de sesgos

De manera coherente, [9] pone de relieve la racionalidad de este nuevo enfoque al comparar sus bondades con los métodos clásicos de autoinforme o con la evaluación de un tercero (*proxy-rated measures*), mencionando que la objetividad de los datos pasivos es un componente fundamental para mejorar la precisión diagnóstica.

Asimismo, estudios en ámbitos como la bipolaridad han demostrado las limitaciones de la evaluación clínica esporádica (por ejemplo, con entrevistas mensuales) al no captar las fluctuaciones diarias o semanales de los síntomas. El fenotipado digital, mediante la recolección continua de datos de movilidad, patrones de sueño, interacciones sociales y uso de dispositivos, ofrece una mirada longitudinal más detallada.

### 3.5. Evolución del campo y madurez tecnológica

Fases históricas y adopción creciente

La literatura describe diversos momentos históricos en la evolución del fenotipado digital. Una categorización común identifica:

1. **Primera etapa:** caracterizada por el uso de dispositivos anteriores a los *smartphones* (como el Nokia N95) y aplicaciones ad hoc para la recolección de datos.
2. **Segunda etapa:** correlacionada con la popularización de sistemas operativos como Android e iOS, que facilitaron el desarrollo y la distribución masiva de aplicaciones de seguimiento.
3. **Tercera etapa:** marcada por la existencia de plataformas maduras, ya sean de código abierto o comerciales, capaces de recopilar y analizar grandes volúmenes de datos en múltiples dominios (depresión, ansiedad, trastornos del sueño, Parkinson, esquizofrenia, trastorno bipolar, detección temprana de riesgo suicida, entre otros).

En sintonía con esto, [8] señala que el campo está pasando de prototipos a plataformas de investigación con múltiples funcionalidades y destaca el surgimiento de un conjunto común de características conductuales relevantes para los principales objetivos de investigación en salud mental y neurociencia.

Nuevos datos sobre la adopción de dispositivos

El crecimiento de los teléfonos inteligentes y dispositivos *wearables* (relojes, pulseras) ha sido determinante para la expansión del fenotipado digital. En [10], se informa que el 84% de la

población mundial ya posee un *smartphone* y cerca del 10% cuenta con algún dispositivo vestible (“*wearable device*”). Estos aparatos generan, de manera continua, datos de múltiples sensores (acelerómetro, GPS, micrófono, pulsómetro, etc.) y aplicaciones, conformando la base del “fenotipo digital” de cada persona.

### 3.6. Contribuciones al conocimiento y retos futuros

#### Aportes al diagnóstico y la investigación

El fenotipado digital ofrece ventajas únicas: la posibilidad de capturar “huellas digitales” de comportamiento y fisiología de forma pasiva, continua y en contextos de vida real. Esto puede:

- **Reforzar diagnósticos** al complementar las evaluaciones clínicas con datos objetivos.
- **Ayudar a la detección temprana** de descompensaciones o síntomas incipientes (p.ej., cambios en la frecuencia de uso del teléfono, alteraciones en patrones de sueño o desplazamiento).
- **Personalizar intervenciones** en salud mental (*just-in-time interventions*), diseñadas específicamente en función de la conducta y el contexto de cada individuo.

#### Debates éticos y desafíos metodológicos

No obstante, su rápida adopción también plantea desafíos importantes, especialmente en lo referente a privacidad, seguridad de datos, consentimiento informado y gobernanza de la información. La captura constante de datos personales y conductuales obliga a investigadores, clínicos y legisladores a definir políticas y marcos éticos que garanticen la protección de la intimidad y los derechos de los usuarios.

En el plano metodológico, existe el riesgo de un “exceso de entusiasmo” sobre las capacidades predictivas y diagnósticas de las tecnologías digitales. Para que el fenotipado digital cumpla su promesa de revolucionar la psiquiatría y la psicología clínica, se requiere una validación rigurosa, algoritmos transparentes y reproducibles, y la integración de disciplinas (psicología, informática, estadística, ética, salud pública).

### 3.7. Perspectiva integradora: el futuro del fenotipado digital

El fenotipado digital ha avanzado desde una idea inspirada en la biología evolutiva de Richard Dawkins (“fenotipo extendido”) hasta consolidarse como un campo multidisciplinar que

combina salud mental, psico informática, computación y estadística avanzada. Sus primeras aplicaciones demostraron su potencial para mejorar la objetividad en la medición de conductas y estados clínicos, así como para abarcar grandes poblaciones de forma pasiva y continua.

A medida que la tecnología madura, se están estableciendo enfoques estandarizados y plataformas escalables para la captura de datos, su almacenamiento seguro y su posterior análisis. Este proceso de “maduración del campo” implica no solo avances en la ingeniería de software y hardware, sino también el desarrollo de marcos conceptuales que permitan identificar y consensuar qué comportamientos digitales son más relevantes para la investigación y la práctica clínica.

En definitiva, el fenotipado digital se perfila como una herramienta cada vez más influyente para:

- **Mejorar la precisión diagnóstica y optimizar intervenciones** personalizadas.
- **Producir conocimiento** sobre la dinámica real de enfermedades psiquiátricas y otros trastornos.
- **Transformar** la forma en que se abordan los problemas de salud mental, al pasar de evaluaciones puntuales a una monitorización longitudinal y contextualizada.

No obstante, el éxito final de esta disciplina dependerá en gran medida de la forma en que se gestionen los retos asociados a la privacidad, la ética de la investigación y la fiabilidad de los análisis de datos a escala masiva. Con la proliferación de los dispositivos inteligentes y la creciente sofisticación de métodos de machine learning, todo indica que el fenotipado digital continuará expandiéndose y forjando nuevos paradigmas de conocimiento y práctica clínica en los años venideros.

## 4. Áreas de aplicación

El fenotipado digital se comprende como la recopilación y análisis de datos generados mediante el uso de dispositivos electrónicos, especialmente *smartphones* y *wearables*, con el fin de describir y predecir conductas y estados de salud. Esta definición se ha ido perfeccionando con el tiempo, integrando variables como la actividad física (acelerometría), patrones de movilidad (GPS), interacciones de teclado o voz, e incluso la llamada "*exhaust*" digital, que incluye el historial de redes sociales y los registros de uso de aplicaciones [11].

De acuerdo con la publicación [6], el fenotipado digital se concibe como una oportunidad para ampliar el alcance del diagnóstico, el monitoreo y la intervención de los trastornos mentales, al extraer información contextual y de comportamiento de la cotidianidad de los pacientes. Con la incorporación de estas medidas y la facilidad de recolección de datos continuos, se espera mejorar la precisión y la rapidez con que se detectan cambios en la salud mental.

### 4.1. Fenotipado Digital y su Evolución Conceptual

De la Biología a la Era Digital

Tradicionalmente, el término "fenotipo" se refería a los rasgos observables de un organismo en el contexto de la genética y la biología. Sin embargo, en la actualidad, el término se ha extendido para abarcar los rastros y datos que dejan los individuos en su interacción cotidiana con dispositivos electrónicos. Esta extensión del concepto hacia un entorno digital abre nuevas posibilidades para el diagnóstico, la prognosis y la intervención en diversos trastornos de salud, especialmente en el área de la salud mental [11].

Dimensiones Clave: Biomarcadores y Fenotipos Conductuales

En el fenotipado digital suelen distinguirse dos componentes principales:

1. **Fenotipado Conductual:** Se centra en la información sobre el comportamiento observable (por ejemplo, movimientos, patrones de interacción social y uso de dispositivos). La actividad motora y el historial de mensajes de texto o llamadas telefónicas pueden revelar síntomas conductuales importantes para la identificación de trastornos [11].
2. **Biomarcadores Digitales:** Incluyen variables que pueden reflejar estados y rasgos subyacentes del sistema nervioso, aportando evidencia de riesgo, aparición o progresión

de enfermedades. Tal como se menciona en la misma obra, estos biomarcadores digitales miden rasgos y cambios neuropatológicos que pueden indicar un riesgo de enfermedad o la progresión de la misma [11].

## **4.2. Aplicaciones en Salud Mental**

### Diagnóstico y Evaluación de Trastornos Mentales

Uno de los ámbitos de mayor crecimiento del fenotipado digital es la psiquiatría y la salud mental. La recolección de datos digitales permite caracterizar y comprender mejor trastornos como la depresión, la ansiedad, la esquizofrenia y el trastorno bipolar [6].

Este enfoque facilita la detección temprana de síntomas y signos de alarma. Por ejemplo, una disminución en la cantidad de llamadas telefónicas, la reducción de la movilidad captada por sensores GPS o cambios en la frecuencia de uso de aplicaciones pueden correlacionarse con estados depresivos [6]. En el caso de la fase maníaca del trastorno bipolar, estos mismos parámetros suelen presentar el patrón opuesto (mayor actividad, más uso de llamadas o mensajes).

Además, estudios han demostrado que la recopilación de datos en tiempo real puede mejorar la precisión diagnóstica, dado que la información sobre conductas cotidianas se integra con evaluaciones clínicas tradicionales, reduciendo la subjetividad y ofreciendo contextos más ricos sobre cada individuo [11].

### Monitoreo Continuo y Prevención de Recaídas

Otro de los grandes beneficios del fenotipado digital radica en la monitorización continua. A diferencia de las evaluaciones clínicas puntuales, la tecnología de sensores y los algoritmos de aprendizaje automático posibilitan la recolección de datos en tiempo real, lo cual brinda la oportunidad de detectar cambios sutiles que podrían anticipar una crisis de salud mental.

Cuando estas herramientas se integran en el seguimiento del paciente, se pueden activar intervenciones tempranas (por ejemplo, una notificación al profesional sanitario o al propio usuario) para prevenir recaídas o episodios agudos [6]. De este modo, se fomenta la atención preventiva en vez de la reactiva, ofreciendo una vigilancia más cercana y una respuesta clínica más oportuna.

## Personalización de Tratamientos

La personalización de tratamientos se erige como una de las metas más ambiciosas del fenotipado digital en el campo de la salud mental. Los datos recopilados permiten diseñar intervenciones a la medida de cada paciente, identificando las necesidades individuales y ajustando la terapia en función del comportamiento y estado emocional real del sujeto [6].

Por ejemplo, un paciente que presenta cierto patrón de ansiedad vinculado a momentos específicos del día podría recibir notificaciones de técnicas de relajación justo antes de esos momentos críticos. Esto se basa en la lectura de patrones de uso del teléfono, geolocalización o ritmo cardíaco capturado por un wearable. Esta aproximación, impulsada por la recolección de datos en tiempo real, posibilita intervenciones contextualmente relevantes, aumentando la probabilidad de adherencia y efectividad.

### **4.3. Investigación en Big Data y Aprendizaje Automático**

El auge del Big Data y de técnicas de *machine learning* o *deep learning* ha permitido que el volumen de datos generados a través de teléfonos inteligentes y dispositivos portátiles sea analizado para descubrir patrones complejos que antes permanecían ocultos. Estos métodos analíticos son la pieza esencial que permite avanzar hacia estimaciones de fenotipado cada vez más sofisticadas [11].

Los algoritmos de aprendizaje automático pueden, por ejemplo, detectar correlaciones entre variables de actividad física y estados de ánimo, predecir el riesgo de ideación suicida a partir de cambios sutiles en el comportamiento en línea o estimar la probabilidad de recaída en trastornos como la depresión y la bipolaridad. En este sentido, la integración de técnicas de Big Data con fenotipado digital abre oportunidades para la investigación traslacional, permitiendo el paso de resultados de laboratorio a aplicaciones clínicas concretas.

### **4.4. Psico informática y Nuevos Campos de Estudio**

El concepto de psico informática se refiere a la intersección entre la psicología y la informática, abordando cómo los datos digitales pueden complementar o incluso revolucionar los métodos tradicionales de evaluación psicológica y psiquiátrica. El fenotipado digital puede incluir elementos como el análisis de patrones de voz, la velocidad de escritura o los tiempos de

respuesta ante diversos estímulos en línea, aportando una visión más completa y dinámica de la conducta y la cognición [11].

Estos enfoques no solo se limitan a la salud mental, sino que también pueden emplearse para la evaluación de la personalidad (por ejemplo, la predicción de los rasgos de la personalidad "Big Five" a partir de los datos de uso del *smartphone*), la detección de trastornos en el uso de internet, la identificación de estilos de vida poco saludables y la monitorización de patrones de sueño y vigilia.

#### **4.5. Atención Médica Digitalmente Soportada**

La atención médica digitalmente soportada involucra el uso de tecnología para potenciar la relación médico-paciente, mejorar la toma de decisiones clínicas y anticipar crisis de salud. Las herramientas de fenotipado digital y biomarcadores digitales están siendo consideradas seriamente como medidas clínicas de apoyo al diagnóstico y el seguimiento, con la meta de intervenir antes de que ocurra una crisis [11].

Por ejemplo, en los servicios de crisis en salud mental, la posibilidad de recibir alertas anticipadas cuando un paciente presenta signos de mayor riesgo de desestabilización podría marcar la diferencia entre una hospitalización prolongada y una intervención temprana y ambulatoria.

#### **4.6. Áreas de Expansión: Salud Pública y Más Allá**

Además de las aplicaciones clínicas y de investigación en psiquiatría, el fenotipado digital se está extendiendo a otras áreas:

**Salud Pública:** En el campo de la epidemiología y la salud comunitaria, se han utilizado datos de geolocalización de teléfonos para evaluar la recuperación tras intervenciones médicas e incluso para analizar patrones de movilidad durante emergencias sanitarias, como ocurrió durante la pandemia de COVID-19 [8].

**Recuperación Física:** El fenotipado digital permite monitorear a pacientes que se recuperan de cirugía o de enfermedades crónicas, siguiendo su ritmo de actividad, desplazamientos e interacciones diarias. Esto proporciona una imagen más clara de cómo evoluciona el paciente en su entorno natural.

**Investigación en Conductas Sociales:** Como muestra el estudio HOPES, el análisis de cambios en la movilidad y la actividad física durante periodos de distanciamiento social confirmó la utilidad del fenotipado digital para medir el impacto de intervenciones masivas a nivel poblacional [8].

#### **4.7.Desafíos y Consideraciones Éticas**

A pesar de las amplias posibilidades, el fenotipado digital enfrenta retos importantes. La recolección continua de datos plantea preguntas sobre privacidad e *informed consent*. Existe el riesgo de malinterpretar comportamientos normales como patológicos y, en el extremo opuesto, se requiere un abordaje cuidadoso para evitar la estigmatización de individuos basándose solo en sus rastros digitales [11].

Asimismo, la calidad de los datos y el desarrollo de algoritmos robustos son factores críticos. La interpretación de la información debe llevarse a cabo con criterios científicos y clínicos sólidos, considerando la responsabilidad ética y legal que implica trabajar con datos personales de salud.

#### **4.8.Conclusiones**

El fenotipado digital representa un paso significativo hacia la medicina personalizada y la salud mental de precisión, al integrar datos objetivos y continuos en la evaluación y el tratamiento de diversos trastornos. Su aplicación abarca desde el diagnóstico y el monitoreo en psiquiatría, hasta la definición de programas de atención médica más integrales y la investigación en Big Data y aprendizaje automático.

Tal como se evidencia, el fenotipado digital ofrece:

- **Mayor exactitud diagnóstica**, gracias a la captura de variables conductuales y fisiológicas en tiempo real.
- **Monitoreo continuo y prevención de crisis**, posibilitando intervenciones tempranas en salud mental.
- **Personalización de tratamientos**, adaptando intervenciones a partir de patrones individuales de comportamiento y estados emocionales.
- **Ampliación de la investigación**, habilitando proyectos en *Big Data* y *machine learning* que permitan descubrir nuevos biomarcadores y correlaciones.

- **Integración con la salud pública**, a través de la vigilancia epidemiológica y el seguimiento de conductas colectivas.

No obstante, el campo debe avanzar con cautela, considerando los desafíos éticos y de protección de datos. Solo de esta manera se garantizará que el fenotipado digital contribuya de manera responsable y efectiva al bienestar de los individuos y de la sociedad en general.

## **5. Consideraciones éticas en el fenotipado digital**

El fenotipado digital se ha convertido en una de las áreas de mayor crecimiento dentro de la investigación clínica y las ciencias de la salud mental. Su premisa fundamental es la recopilación sistemática de datos de comportamiento, fisiología y contexto de los individuos a través de sensores portátiles, aplicaciones en teléfonos inteligentes y otras tecnologías digitales [12]. Estas herramientas capturan gran variedad de variables, que van desde la frecuencia de uso de ciertas aplicaciones hasta los patrones de movilidad registrados por el GPS. La intención es traducir estos datos en indicadores relevantes para el diagnóstico, seguimiento y tratamiento de diversas condiciones de salud mental y física.

Sin embargo, en paralelo con el entusiasmo por las oportunidades que brinda el fenotipado digital, han surgido serias preocupaciones éticas y legales. El hecho de recolectar datos de forma pasiva, prácticamente continua y, en muchos casos, muy íntima sobre los hábitos y comportamientos de los usuarios, exige una discusión profunda sobre la privacidad, la autonomía, la validez de la información y las implicaciones para la relación médico-paciente [13]. De igual manera, múltiples autores han señalado el riesgo de que este enfoque pudiera deshumanizar la práctica clínica, someter al profesional a sistemas automatizados de toma de decisiones, o incluso fomentar la manipulación o vigilancia inapropiada, conectándose con fenómenos más amplios como el capitalismo de la vigilancia [14].

En este documento se exponen las principales consideraciones éticas que deben tenerse en cuenta al diseñar, implementar y evaluar soluciones de fenotipado digital. Asimismo, se exploran los aspectos relacionados con el consentimiento, la autonomía de las personas, la propiedad de los datos, la transparencia, la pertinencia clínica, la despersonalización y la posible patologización de comportamientos, entre otras cuestiones de gran relevancia.

### **5.1.Privacidad de los datos y capitalismo de vigilancia**

Uno de los temas más críticos dentro del fenotipado digital es la protección de la privacidad y la seguridad de la información recabada. La capacidad para monitorear de forma pasiva a los usuarios (ya sea a través de aplicaciones instaladas en el teléfono, dispositivos portátiles (*wearables*) o tecnologías de geolocalización) plantea el riesgo de recopilar información extremadamente sensible. Entre estas categorías de datos se incluyen patrones de sueño, niveles

de actividad física, frecuencia de interacciones sociales o incluso aspectos emocionales inferidos a partir de llamadas, mensajes de texto o uso de redes sociales [12].

A este respecto, Shoshana Zuboff (2019) acuñó el término “capitalismo de vigilancia” para describir la tendencia de ciertas industrias tecnológicas que sustentan sus modelos de negocio en recopilar, analizar y comercializar datos de comportamiento de sus usuarios [14]. Aunque la intención original del fenotipado digital sea eminentemente clínica o investigativa, no se puede desconocer que los datos recabados podrían ser eventualmente utilizados con fines comerciales o, peor aún, para la manipulación de conductas y toma de decisiones políticas. En [15], se advierte que la extracción y el análisis masivo de datos biométricos y conductuales se conecta con una tendencia global hacia la vigilancia y la manipulación, por lo que subestimar estos riesgos podría conducir a prácticas invasivas o abusivas.

La cuestión de la privacidad es, por tanto, un desafío transversal. Para abordarlo se requiere, al menos, que las instituciones y equipos de investigación cumplan con estándares sólidos de seguridad de la información, como la familia de normas ISO 27000 [1]. Estas directrices proponen la certificación de políticas, procedimientos y controles internos que garanticen la protección de la información. Sin embargo, la adopción de normas técnicas por sí sola no es suficiente: se precisa de una cultura ética que priorice el manejo responsable de los datos y establezca claramente los límites del uso que se les puede dar, así como la responsabilización de los agentes involucrados.

## **5.2.El consentimiento informado como piedra angular**

La idea de que las personas tengan la posibilidad de decidir qué datos comparten, con qué propósitos y durante cuánto tiempo, resulta central en la ética de la investigación clínica y psicológica. En el caso del fenotipado digital, este principio se complica debido a que, con frecuencia, los datos se recaban de manera pasiva [3]. Es decir, el usuario puede conceder acceso a su GPS o a ciertos sensores de su teléfono en un inicio, pero con el tiempo podría olvidar el alcance real de esa autorización. Por ende, es esencial que el consentimiento sea verdaderamente informado y renovado o revisado cuando se introduzcan cambios en el protocolo o en los métodos de recolección de datos.

En [5], se enfatiza la importancia de un consentimiento explícito para recabar información personal o sensible en intervenciones digitales de salud mental. Además de aclarar la finalidad de la intervención, se debe comunicar a los usuarios de forma transparente cómo se procesarán

sus datos, quién tendrá acceso a ellos y los posibles riesgos asociados. Este proceso de información debe ser continuo, ya que el consentimiento no puede considerarse un trámite burocrático puntual, sino un diálogo permanente entre investigador/terapeuta y participante.

Asimismo, la transparencia se extiende más allá del simple acto de informar; también implica aclarar que, en algunos casos, los usuarios están interactuando con un “agente inteligente” y no con un profesional humano [5]. Este matiz es crucial para evitar engaños, ya que el usuario debe saber si los consejos o análisis que recibe provienen de un sistema algorítmico que actúa basándose en patrones de datos y no en la experiencia clínica directa de un experto humano. Todo esto redundaría en la necesidad de respetar la autonomía y la capacidad de decisión de los participantes, quienes tienen derecho a conocer las limitaciones y alcances de la tecnología que emplean.

### **5.3. Responsabilidad de los comités de ética e instituciones**

Junto al consentimiento informado, otro mecanismo fundamental para resguardar los derechos de los participantes es la revisión y aprobación por parte de comités de ética independientes, como el *Institutional Review Board* (IRB) en el ámbito anglosajón. En [11], se recalca que estos comités tienen el propósito de asegurar que se cumplan los principios éticos básicos (beneficencia, no maleficencia, autonomía y justicia) y que se proteja a las personas que participan en los estudios de riesgos potenciales. No obstante, la revisión de protocolos que involucran fenotipado digital puede ser más compleja que en investigaciones tradicionales, debido a la novedad y la sofisticación de los métodos empleados para recolectar datos.

Además de la revisión inicial, puede ser necesario actualizar las aprobaciones éticas conforme avanza el proyecto y se introducen modificaciones técnicas o analíticas (por ejemplo, el uso de nuevos algoritmos de *machine learning* que extraigan patrones hasta entonces no contemplados). Del mismo modo, se hace patente que la responsabilidad no recae únicamente en los comités, sino también en los investigadores, desarrolladores de software y profesionales de la salud que operan la tecnología. Este círculo virtuoso de supervisión y actualización continua de los riesgos y beneficios se vuelve esencial para mantener la integridad del estudio y el bienestar de los participantes.

#### **5.4. Riesgos de despersonalización, patologización y exclusión**

La implementación de sistemas de fenotipado digital en la práctica clínica puede conllevar efectos no deseados en la dinámica entre el profesional de la salud y el paciente. Existe el temor de que los profesionales de la salud (HCPs) lleguen a depender en exceso de los sistemas automatizados de apoyo a la toma de decisiones, reduciendo su papel a meros validadores de lo que dictan los algoritmos [16]. Ello podría minar la confianza en el criterio clínico o, en el peor de los casos, llevar a la subestimación de factores subjetivos como el relato personal de angustia o el contexto psicosocial del individuo.

Por otra parte, el fenotipado digital también puede propiciar la patologización de ciertos comportamientos normales, al interpretarlos de forma aislada de su contexto o historia de vida. Por ejemplo, un cambio en la rutina de sueño detectado por un dispositivo móvil podría considerarse anómalo según un algoritmo estadístico, pero en realidad responder a cambios temporales en el estilo de vida o a factores culturales. En otras palabras, la digitalización de la conducta corre el riesgo de simplificar en exceso la complejidad humana, contribuyendo a etiquetar como “desviaciones” patrones que, en realidad, podrían ser inofensivos o incluso adaptativos.

Un último punto relacionado es la posible exclusión de ciertas poblaciones que, por razones económicas o culturales, no tienen acceso estable a las tecnologías o no desean ser monitorizadas de manera permanente. Esto conlleva el peligro de que las grandes bases de datos se construyan sobre muestras no representativas de la población, perpetuando desigualdades y limitando la validez externa de los hallazgos.

#### **5.5. Diseño centrado en el usuario y consideraciones de aceptabilidad**

Además de garantizar la privacidad y la ética, la efectividad de las intervenciones basadas en fenotipado digital depende en gran medida de la aceptación que tengan por parte de los usuarios. Las soluciones tecnológicas deben ser clínicamente relevantes y dirigirse a las necesidades y preferencias de los consumidores [17]. Sin un abordaje centrado en las personas (es decir, que considere los contextos en los que viven, sus valores y la facilidad con que pueden adoptar la tecnología), los posibles beneficios podrían retrasarse o diluirse.

La usabilidad de las aplicaciones, la claridad en los mensajes y la posibilidad de que el usuario intervenga en la configuración de los parámetros de recolección de datos son factores

determinantes para que los individuos mantengan el compromiso y se sientan seguros participando en estudios de fenotipado digital. El abandono o la apatía ante la aplicación de estas herramientas puede darse no solo por preocupaciones sobre la privacidad, sino también por el consumo de batería, la complejidad de uso y la falta de retroalimentación significativa [18]. Por ende, los investigadores deben asumir la responsabilidad de generar estrategias que resulten atractivas y transparentes, y que ofrezcan un claro beneficio a los participantes.

## **5.6. Retos técnicos y validez de los datos en escenarios reales**

No menos importante es la cuestión técnica relacionada con la calidad y fiabilidad de los datos. Durante la fase de recolección de datos, se han encontrado problemas tales como deficiencias en los sensores, alto consumo energético y falta de interés de los participantes. Estos factores pueden llevar a un sesgo en la información obtenida, reduciendo su representatividad y, en consecuencia, su aplicabilidad clínica [18].

Además, al trabajar con grandes volúmenes de datos (*big data*), se requieren metodologías robustas para filtrar, procesar y analizar la información de manera precisa. Un manejo inadecuado de los algoritmos de *machine learning* podría conducir a errores de clasificación, correlaciones espurias o resultados que, en última instancia, no sean reproducibles en entornos reales. Esto no solo supone un problema científico, sino también un riesgo ético, pues se podrían tomar decisiones clínicas o de salud pública basadas en datos sesgados o inexactos. De ahí la relevancia de la replicación y la validez externa al probar estos métodos en diferentes poblaciones y contextos.

## **5.7. Perspectivas futuras y líneas de acción**

A pesar de los desafíos éticos, el fenotipado digital representa una oportunidad única para mejorar la comprensión y la gestión de condiciones de salud mental, incluyendo la prevención del suicidio y la intervención temprana en trastornos del ánimo [19]. La clave radica en encontrar un balance que permita aprovechar la gran riqueza de datos y el potencial analítico, a la vez que se respeten los principios éticos fundamentales.

1. **Formulación de marcos normativos y guías de buenas prácticas:** Varios autores coinciden en la necesidad de establecer guías claras y actualizadas que orienten a los investigadores, desarrolladores y profesionales de la salud. Estos marcos deben incluir

aspectos como la gestión de consentimientos, la minimización de riesgos, la transparencia en los usos secundarios de los datos y la equidad en el acceso.

2. **Educación y capacitación:** Tanto los pacientes como los profesionales de la salud deben recibir información y capacitación sobre qué es el fenotipado digital, cuáles son sus ventajas y limitaciones, y cómo mantener una actitud crítica frente a los resultados generados. Esto favorece una cultura de colaboración, en lugar de una mera relación de dependencia de las herramientas tecnológicas.
3. **Diseño centrado en la persona y diversidad cultural:** Para reducir sesgos y asegurar la aceptación social, es fundamental involucrar a los usuarios en el proceso de diseño de las soluciones de fenotipado digital. Esto implica realizar pruebas piloto con poblaciones diversas, reconocer diferencias culturales y ofrecer opciones de personalización y control sobre los datos recabados.
4. **Protección de la autonomía y supervisión humana:** Aunque los algoritmos pueden ofrecer sugerencias diagnósticas o recomendaciones terapéuticas, la decisión final debe descansar en el criterio clínico y la voluntad informada del paciente. La tecnología ha de ser un complemento, no un sustituto de la interacción humano-humano.
5. **Auditorías y evaluación continua:** La implementación de auditorías independientes y mecanismos de revisión periódica por parte de los comités de ética y la comunidad científica fomenta la transparencia y la corrección temprana de desviaciones. Esto también supone la necesidad de publicar los resultados negativos o neutros, evitando sesgos de reporte y manipulaciones de los hallazgos.

## **5.8. Cumplimiento Normativo: GDPR y Otras Regulaciones**

En el espacio europeo, el Reglamento General de Protección de Datos (GDPR) establece reglas claras sobre la obtención y uso de datos personales. Este marco legislativo se ha convertido en un punto de referencia mundial y define principios como la minimización de datos, el derecho al olvido y la necesidad de un consentimiento explícito del titular de la información. En [8] señala la relevancia de contar con equipos interdisciplinarios que incluyan abogados y expertos en privacidad, quienes supervisen la adecuación de los proyectos de fenotipado digital a esta normativa.

Además de la GDPR, otros países han promulgado leyes sobre protección de datos personales, por lo que las iniciativas globales deben adaptarse a múltiples contextos legales. Esto no solo implica la revisión de contratos y términos de uso, sino también el diseño de arquitecturas

tecnológicas capaces de anonimizar y encriptar los datos, junto con la implementación de protocolos de gestión de brechas de seguridad.

## 5.9. Transparencia, Consentimiento Informado y Potenciales Riesgos

La transparencia es un eje fundamental para legitimar el fenotipado digital ante la comunidad científica y la sociedad en su conjunto. Más allá de los documentos de consentimiento, se busca que los participantes comprendan qué información se recopila y con qué propósito. Es conveniente habilitar interfaces que muestren de forma sencilla cómo se procesan sus datos, cómo se guardan y quién puede acceder a ellos.

Entre los riesgos recurrentes se identifican:

- **Potencial vigilancia o control excesivo:** El registro continuo de rutinas podría derivar en usos no deseados, como la discriminación laboral o la restricción de libertades individuales.
- **Comercialización no consentida:** Los datos podrían transferirse a terceros sin aprobación explícita de los participantes, con fines publicitarios o de investigación no relacionados.
- **Falsa seguridad o sobrediagnóstico:** Si los sistemas no están debidamente validados, podrían generar falsas alarmas o interpretaciones erróneas que influyan en la salud mental de los usuarios.

En conjunto, estas consideraciones resaltan la importancia de la visión interdisciplinaria, con la activa participación de bioeticistas, juristas y especialistas en gobernanza de datos para prevenir situaciones de abuso o malas prácticas.

## 5.10. Conclusiones

El fenotipado digital irrumpe con fuerza en el panorama de la salud mental y la investigación clínica, presentando tanto oportunidades extraordinarias como desafíos éticos considerables. Por un lado, promete el desarrollo de diagnósticos más precisos, un monitoreo continuo de los síntomas y la posibilidad de intervenciones personalizadas. Por el otro, al involucrar la recolección masiva de datos íntimos, plantea el riesgo de vulnerar la privacidad de las personas, abusar de la confianza depositada en la tecnología y relegar la experiencia profesional de los clínicos a un segundo plano [13].

La literatura existente insiste en que la clave para el avance responsable del fenotipado digital radica en la construcción de un sólido andamiaje ético, que incluya la validación de protocolos por parte de comités de ética [11], la implementación de prácticas de consentimiento informado y transparencia [5], la adopción de medidas de seguridad y privacidad robustas [1] y la atención a la dimensión humana del cuidado de la salud, evitando la despersonalización de la relación médico-paciente [16].

Por último, no se debe pasar por alto que el fenotipado digital se inserta en un entorno socioeconómico y político en el que la comercialización de datos y la vigilancia masiva son realidades crecientes. Se remarca la urgencia de reflexionar sobre los aciertos y errores del pasado para no reproducir las mismas injusticias y violaciones a la privacidad que han caracterizado otras iniciativas tecnológicas [20]. La tarea, por tanto, es asumir una postura crítica y comprometida, que promueva la colaboración entre profesionales, investigadores, reguladores y usuarios finales, garantizando que el fenotipado digital cumpla su promesa de contribuir a una medicina más preventiva, integral y humana, sin sacrificar los valores éticos y la dignidad de las personas.

## 6. Perspectivas interdisciplinarias

### 6.1. Colaboración entre Ciencia, Tecnología y Sector Comercial

La expansión de los dispositivos móviles, los avances en la ciencia de datos y la creciente disponibilidad de herramientas de aprendizaje automático han generado un contexto idóneo para la cooperación entre el ámbito académico, el sector industrial y las organizaciones de salud. Distintos investigadores señalan que el desarrollo de tecnologías asociadas al fenotipado digital es resultado de un cruce de intereses entre profesionales de la psicología, psiquiatría, ingeniería de software y ciencia de datos.

“Es crucial establecer una colaboración saludable entre la academia y entidades comerciales en el desarrollo de tecnologías de fenotipado digital” [8].

Esta colaboración puede tomar diversas formas, por ejemplo:

1. **Proyectos de investigación conjuntos:** Donde universidades y centros de salud se asocian con empresas tecnológicas para obtener financiación y acceso a infraestructuras de almacenamiento y análisis de datos a gran escala.
2. **Consolidación de ‘start-ups’ o spin-offs:** Profesionales formados en grupos académicos deciden lanzar sus propias empresas para aplicar el conocimiento generado, consolidando soluciones que puedan comercializarse y llegar a entornos clínicos o directamente a la población general.
3. **Contratos de transferencia tecnológica:** El sector privado financia la investigación sobre fenotipado digital con miras a generar productos o servicios (aplicaciones, dispositivos vestibles, plataformas de monitorización), mientras que los académicos proporcionan rigor metodológico, validación de resultados y publicaciones científicas que respaldan la eficacia de los desarrollos.

No obstante, surgen algunas tensiones inherentes a la colaboración. Entre los desafíos más frecuentes se encuentra la divergencia de objetivos: el sector comercial busca a menudo la rentabilidad económica y la proyección de mercado, en tanto que la academia persigue la generación de conocimiento y el cumplimiento de estándares éticos y científicos. Para evitar conflictos, se requieren estructuras de gobernanza que clarifiquen la propiedad intelectual, la privacidad de los datos y los límites de uso comercial. Asimismo, es esencial que exista una

supervisión ética y legal que prevenga abusos o desvíos en la finalidad de los proyectos, sobre todo cuando se trabaja con datos sensibles relacionados con la salud mental.

Otro aspecto para subrayar es la calidad de los desarrollos tecnológicos. Las entidades comerciales suelen contar con equipos dedicados a la ingeniería y la optimización de plataformas escalables. Al colaborar con grupos de investigación, se aprovecha el conocimiento que estos tienen sobre los fundamentos científicos de la conducta, la psicopatología o las estadísticas avanzadas. Esta combinación de esfuerzos conlleva un salto cualitativo en la robustez de las herramientas de fenotipado digital, permitiendo que se trasladen a una fase de uso en clínicas y consultas.

A mediano y largo plazo, la capacidad de aunar recursos financieros y talento multidisciplinario propicia que los proyectos de fenotipado digital no se queden en prototipos de laboratorio, sino que pasen a ser soluciones estables y adaptadas a la realidad de diversos entornos (hospitales públicos, clínicas privadas, programas de prevención, etc.). Sin embargo, dicha viabilidad está supeditada a la correcta gestión de las relaciones entre los múltiples actores y a la transparencia en la definición de objetivos compartidos.

## **6.2. Avances en Psiquiatría y Ciencias del Comportamiento**

La psiquiatría y, en general, las ciencias del comportamiento constituyen uno de los campos donde el fenotipado digital ha tenido mayor impacto e interés. Investigadores y clínicos han visto en esta metodología un potente aliado para estudiar fenómenos que tradicionalmente se abordaban únicamente a través de autoinformes, evaluaciones puntuales o test psicológicos aplicados en entornos controlados.

“El fenotipado digital ha demostrado su potencial en psiquiatría para clasificar trastornos y desarrollar biomarcadores digitales” [9].

Estos avances se sustentan en varias características diferenciales:

1. **Continuidad de la medición:** A diferencia de la evaluación clínica en gabinete, que se limita a un momento preciso (por ejemplo, una vez al mes o al año), el fenotipado digital recoge datos de manera constante y puede detectar variaciones diarias o incluso horarias en el estado de ánimo, el nivel de actividad o los patrones de sueño.

2. **Captura de la conducta en el entorno cotidiano:** Al basarse en dispositivos que acompañan al individuo en su vida diaria (*smartphones*, pulseras, relojes inteligentes), se obtiene información más espontánea y menos sujeta a sesgos de deseabilidad social.
3. **Análisis de factores contextuales:** No solo se miden variables intrapersonales, sino también elementos del entorno físico y social, como la frecuencia de interacción con otros, los lugares visitados o el tiempo dedicado a ciertas actividades.

Esta riqueza de datos permite desarrollar modelos explicativos que conectan el aspecto biológico (por ejemplo, ritmos circadianos) con dinámicas relacionales y comportamentales (contacto social, ocio, etc.). De modo simultáneo, se abren oportunidades para identificar patrones de alerta que preceden a crisis o recaídas de trastornos mentales. Si un paciente con trastorno bipolar modifica drásticamente sus hábitos de sueño y actividad, las aplicaciones y algoritmos podrían adelantar la posibilidad de que se acerque un episodio maníaco o depresivo, permitiendo acciones preventivas por parte del equipo de salud.

En el plano práctico, muchos servicios de psiquiatría y psicología clínica han comenzado a incorporar estrategias de fenotipado digital en la monitorización remota de pacientes. Esto cobra relevancia en contextos de atención domiciliaria o en lugares con acceso limitado a servicios especializados. Además, la combinación del fenotipado digital con consultas presenciales u online permite al terapeuta contar con una visión longitudinal que complementa y enriquece la relación terapéutica, convirtiéndose en un insumo valioso para la toma de decisiones clínicas.

### **6.3. Ampliación de Campos de Aplicación**

La atención original al ámbito de la salud mental se ha visto rápidamente complementada por exploraciones del fenotipado digital en otras disciplinas médicas y científicas.

“...aunque actualmente el enfoque se centra en la psiquiatría, el fenotipado digital podría emplearse en otras disciplinas para cuantificar el comportamiento humano y proporcionar nuevos *insights* significativos” [9].

#### Enfermedades Crónicas y Vigilancia Epidemiológica

El monitoreo continuo de variables fisiológicas (por ejemplo, frecuencia cardíaca, presión arterial, nivel de actividad física) resulta esencial para la detección temprana de

descompensaciones en enfermedades crónicas como la diabetes, la hipertensión o las dolencias cardíacas. Una aplicación que registre la adherencia al tratamiento, los patrones de alimentación o el nivel de glucosa en sangre a través de sensores específicos puede enviar alertas cuando se identifique una desviación crítica [9].

Asimismo, en una perspectiva más amplia, la geolocalización y los datos de movilidad podrían ayudar a agencias de salud pública en tareas de vigilancia epidemiológica, especialmente en brotes infecciosos. El cruce de datos anónimos sobre ubicación y sintomatología declarada puede detectar zonas de mayor incidencia y coadyuvar en la toma de decisiones gubernamentales [9].

### Rehabilitación Física y Neurológica

En entornos de rehabilitación (ya sea para pacientes con accidentes cerebrovasculares, lesiones medulares o trastornos neurológicos degenerativos) el fenotipado digital posibilita un seguimiento detallado de los progresos y la adhesión al plan de ejercicios. Por ejemplo, el análisis de la calidad y frecuencia de la marcha, la fuerza en extremidades o la postura se integra con datos recogidos por sensores portátiles, generando marcadores objetivos del avance del paciente [9].

Este escenario implica la participación conjunta de fisioterapeutas, neurólogos, ingenieros biomédicos y desarrolladores de software, quienes han de acordar cómo traducir variables cinéticas y cinemáticas en indicadores clínicamente útiles [9].

### Educación y Contextos Organizacionales

Aunque menos explorado, el ámbito de la educación empieza a ver aplicaciones de fenotipado digital, por ejemplo, para medir los niveles de atención, los hábitos de estudio o el *engagement* en entornos virtuales. Este tipo de monitoreo, combinado con informes de docentes y psicopedagogos, permitiría diseñar intervenciones adaptativas para estudiantes con problemas de concentración u otras dificultades [9].

En el sector organizacional, se han delineado proyectos para detectar signos de estrés laboral, burnout o insatisfacción de los empleados a partir de patrones de uso del correo electrónico o la actividad en dispositivos corporativos. Sin embargo, esta aplicación es muy sensible desde el punto de vista ético, pues se deben evitar modalidades de vigilancia excesiva que vulneren la privacidad y los derechos laborales [9].

## **6.4.Perspectiva Crítica e Interdisciplinaria**

### Riesgo de Reduccionismo y Sesgos Algorítmicos

La capacidad de recopilar datos conductuales y fisiológicos de forma masiva no implica, por sí misma, que se logre explicar el comportamiento humano en su totalidad. El *digital gaze* la mirada que se confía excesivamente en la objetividad de los datos) puede invisibilizar aspectos culturales, históricos y relacionales que constituyen la esencia de las conductas humanas [15].

En este sentido, la colaboración interdisciplinaria es crucial para prevenir posturas de reduccionismo, entendidas como la simplificación excesiva de problemas complejos a meras correlaciones estadísticas. Además, las ciencias sociales y las humanidades pueden aportar enfoques que adviertan sobre sesgos algorítmicos: los conjuntos de datos entrenados con poblaciones específicas podrían no generalizarse a otros grupos, reforzando estereotipos o propiciando la discriminación [15].

### Co-creación y Ontologías Múltiples

Para sortear estos riesgos, se sugiere la adopción de ontologías colaborativas que reflejen la pluralidad de perspectivas. La co-creación de marcos conceptuales y categorías diagnósticas involucra a médicos, psicólogos, estadísticos y miembros de la comunidad destinataria (pacientes, familiares, cuidadores). Esto fomenta la transparencia epistémica y la inclusividad, de modo que las categorías no se definan unilateralmente por un único campo del saber, sino que emerjan del diálogo entre múltiples disciplinas [5].

## **6.5.Plataformas, Herramientas y Equipos Multidisciplinarios**

### Retos Técnicos y Heterogeneidad de Dispositivos

La implementación práctica del fenotipado digital exige un trabajo complejo de desarrollo de software y gestión de la diversidad tecnológica. Hoy en día, las personas utilizan multitud de dispositivos con sistemas operativos distintos, versiones de firmware y capacidades de hardware variables, lo que dificulta la estandarización de las aplicaciones y la garantía de una calidad homogénea de los datos.

Los equipos de investigación afrontan dificultades al crear aplicaciones para cada sensor y cuestionario, topándose con desafíos relacionados con la compatibilidad, el rendimiento y la

seguridad. Para subsanar estos problemas, es frecuente que los grupos multidisciplinarios incluyan ingenieros especializados en sistemas móviles y en computación en la nube, así como diseñadores de interacción que optimicen la experiencia de uso para que los participantes puedan instalar y ejecutar las aplicaciones sin complicaciones [18].

### Visualización de Datos y Herramientas de Soporte

Otro reto es la interpretación de la gran cantidad de información generada. Aparte de la ingeniería para la recolección y el procesamiento, resulta vital crear interfaces o *dashboards* que faciliten la comprensión de las métricas. Un profesional de la salud, por ejemplo, necesita rápidamente distinguir si un paciente presenta cambios preocupantes en sus hábitos de sueño o actividad, sin tener que examinar manualmente miles de datos de localización o de pasos diarios [18].

Estos paneles de control, por ende, requieren la colaboración con expertos en visualización de la información, diseño centrado en el usuario y, por supuesto, con los propios clínicos que especifican qué variables son relevantes. De esta forma, los datos brutos se convierten en indicadores claros (por ejemplo, “variabilidad del sueño en la última semana” o “nivel de interacción social frente al promedio mensual”) que faciliten la toma de decisiones [18].

## **6.6. Aplicaciones en la Prevención de Conductas de Riesgo**

### Prevención del Suicidio

Dentro de los ámbitos más prominentes de aplicación del fenotipado digital se encuentra la prevención del suicidio. Dada la magnitud de esta problemática a nivel global y la dificultad de identificar señales tempranas en el día a día de los pacientes, distintos proyectos han buscado capitalizar el análisis continuo de datos para anticipar crisis suicidas [19].

El registro de cambios en la rutina, disminución de la interacción social o alteraciones del ciclo de sueño y vigilia pueden funcionar como marcadores de alerta [19]. El fin último es permitir a familiares, cuidadores o profesionales de la salud intervenir a tiempo, proponiendo apoyo, modificando medicaciones o reforzando la red de contención [19].

No obstante, la eficacia y seguridad de estas estrategias dependen fuertemente de:

- **Modelos de *machine learning*** bien entrenados, capaces de discriminar entre variaciones normales y patrones de riesgo real.
- **Canales de comunicación** inmediatos con personas de referencia (p. ej., un cuidador o terapeuta designado) para activar protocolos de emergencia.
- **Ética y responsabilidad** en la notificación de alertas, evitando sobre alarmar o estigmatizar al individuo monitoreado.

#### Adicciones y Otras Conductas de Riesgo

El fenotipado digital también se explora en la prevención de recaídas en adicciones (sustancias, ludopatía, etc.). A través de los datos de geolocalización, por ejemplo, se podrían detectar patrones que indiquen visitas a lugares asociados con el consumo. Sumado a la monitorización de estados anímicos o estrés, estos sistemas podrían emitir alertas que recordaran estrategias de afrontamiento o contactaran con un profesional. La colaboración de especialistas en psicología de la adicción, trabajo social y terapia ocupacional con equipos de ingenieros y analistas de datos resulta indispensable para diseñar aplicaciones que vayan más allá de la detección, ofreciendo una respuesta integral y respetuosa de la dignidad de la persona [19].

### 6.7. Desarrollo Metodológico, Machine Learning y Poblaciones Diversas

#### Big Data, Modelado Predictivo y Validación Científica

El fenotipado digital genera una avalancha de datos: señales de movimiento, ubicaciones, uso de aplicaciones, frecuencia de interacción social, patrones de habla en mensajes de texto, etc. Este volumen y variedad de información se aproxima al paradigma de *big data*. Las iniciativas en fenotipado digital recogen datos de manera continua, lo que exige la adopción de algoritmos de *machine learning*, estadística avanzada y técnicas de minería de datos [6].

Sin embargo, la validez de los hallazgos depende de:

1. **Calidad de los datos:** Minimizar ruidos, datos faltantes, mediciones erróneas o duplicadas.
2. **Diseño de investigación:** Emplear metodologías que permitan aislar variables relevantes y considerar factores de confusión.

3. **Replicabilidad y transparencia:** Facilitar la revisión de código, procedimientos de análisis y modelos, para que otros grupos puedan reproducir resultados en diferentes contextos.

#### Perspectiva Transcultural y Diversidad en la Muestra

La generalización de los resultados constituye un punto neurálgico. Muchos estudios iniciales se han realizado con muestras limitadas geográficamente o compuestas por voluntarios con alta alfabetización digital. Para extender la utilidad del fenotipado digital a poblaciones rurales, personas mayores o individuos con discapacidades, se requiere una adaptación cultural y tecnológica que tenga en cuenta modos de uso del teléfono diferentes, niveles de acceso a Internet y otras particularidades [16].

Por esa razón, distintos investigadores han impulsado proyectos en naciones con menor infraestructura digital, tomando en consideración variables socioeconómicas y culturales que podrían afectar la interpretación de los datos (por ejemplo, el significado de un bajo número de llamadas entrantes varía si se considera una cultura donde las interacciones se dan más cara a cara que mediante dispositivos electrónicos).

### **6.8.Hacia la Psiquiatría de Precisión y Nuevas Categorías Diagnósticas**

#### Revisión de Manuales Diagnósticos y Nuevas Tipologías

Desde distintas corrientes teóricas se ha criticado el carácter estático y a veces poco matizado de manuales diagnósticos como el DSM o la CIE (la CIE tiene como objetivo ser de utilidad a los países miembros de la OMS, así como a los servicios de atención sanitaria de primera línea, mientras que el DSM está mayormente dirigido al campo de la Psiquiatría), sobre todo en trastornos psiquiátricos con elevada heterogeneidad clínica. El fenotipado digital ofrece la posibilidad de refinar las categorías diagnósticas, pues aporta datos de alta granularidad y enfoque longitudinal.

“El fenotipado digital sostiene la promesa de proveer conocimientos sobre condiciones y estados, así como la posibilidad de crear potencialmente nuevas categorías psicológicas” [13].

La hipótesis subyacente es que la observación micro temporal de la conducta podría descubrir subtipos dentro de un diagnóstico mayor (por ejemplo, diferentes trayectorias de la depresión o

ansiedad) que hoy no se distinguen en los manuales. Esto podría traducirse en tratamientos más específicos y pronósticos más precisos [13].

### Modelos Dimensionales y Psiquiatría de Precisión

Sumado a lo anterior, han cobrado fuerza los modelos dimensionales de la psicopatología. Estos modelos proponen que los trastornos no se reduzcan a “presencia” o “ausencia” de un síndrome, sino a grados y variaciones continuas de síntomas. Los datos continuos generados por el fenotipado digital sintonizan muy bien con esta visión, ya que permiten ver cómo un síntoma fluctúa en intensidad a lo largo del tiempo [13].

En perspectiva, la psiquiatría de precisión busca integrar esta información con datos genéticos, de neuroimagen y de historia de vida, a fin de diseñar intervenciones verdaderamente personalizadas. El fenotipado digital constituye, así, uno de los pilares de dicha aproximación, complementando las escalas clínicas tradicionales con un seguimiento dinámico y personalizado [13].

## **6.9. Perspectiva de la Psico informática y el Estudio de la Personalidad**

### Psico informática y el Encuentro entre la Psicología y la Computación

La psico informática, entendida como un área que fusiona la psicología con las ciencias de la computación, encuentra en el fenotipado digital un terreno fértil para generar conocimiento.

“Este apartado del psicodiagnóstico puede considerarse parte de una nueva disciplina de investigación llamada Psico informática, donde psicólogos y científicos de la computación cooperan para comprender mejor la mente humana” [14].

El interés clave radica en el uso de modelos computacionales y el análisis de grandes bases de datos para identificar correlaciones entre rasgos de personalidad (ejemplo., los Big Five: apertura, responsabilidad, extraversión, amabilidad y neuroticismo) y el comportamiento digital cotidiano (mensajería, llamadas, patrones de visita a redes sociales, etc.) [14].

### Aplicaciones Prácticas en Salud y Entornos Laborales

La personalidad condiciona múltiples esferas, incluyendo el cumplimiento de tratamientos médicos y el desempeño profesional. Conocer los rasgos de personalidad de un individuo puede ayudar a personalizar intervenciones de bienestar o planes de ejercicio, así como a predecir el

éxito en determinadas áreas laborales. No obstante, su aplicación comercial (por ejemplo, en la selección de personal) exige cuidados éticos adicionales para evitar prácticas discriminatorias o invasión de la privacidad [14].

El fenotipado digital, alineado con la psico informática, brinda la oportunidad de elaborar modelos predictivos más precisos y de diseñar herramientas que ayuden a las personas a conocer mejor sus fortalezas y áreas de mejora, siempre y cuando se respeten los lineamientos de consentimiento y se garantice la no manipulación de la información.

## **6.10. Ética y Empoderamiento del Paciente**

### Enfoque Centrado en la Persona

La visión tradicional del sistema de salud se ha basado en una relación asimétrica, donde el profesional es el principal depositario del conocimiento. Con la adopción de soluciones digitales, se abre la posibilidad de que los propios usuarios y pacientes tengan un rol más activo, accediendo a sus datos y entendiendo cómo sus hábitos influyen en su salud.

El fenotipado digital podría “empoderar a los individuos a definir su propia salud mental digital y establecer sus propias normas digitales” [16]. Esto supone un cambio de paradigma: la persona no se limita a recibir directrices desde la clínica, sino que coparticipa en el monitoreo y evalúa su propia evolución.

### Transparencia y Autogestión

La implementación de proyectos orientados al empoderamiento precisa la creación de interfaces y reportes comprensibles, evitando tecnicismos que confundan al usuario. Se han sugerido sistemas donde el individuo recibe retroalimentaciones periódicas sobre su estado de ánimo estimado, su nivel de actividad física o la regularidad de sus ciclos de sueño, lo cual puede motivarle a ajustar conductas. Con todo, resulta fundamental que este flujo informativo se maneje con suma delicadeza, de modo que no genere alarmas innecesarias o una ansiedad excesiva [16].

Por otro lado, existe el debate sobre el derecho a la desconexión o la libre decisión de no participar en sistemas que registran permanentemente la conducta. En ese sentido, la mayoría de los expertos concuerdan en que la participación debe ser voluntaria, con la opción de

suspenderla sin repercusiones en la atención médica, garantizando así la autonomía de cada persona [16].

### **6.11. Convergencia Metodológica y Coexistencia con Evaluaciones Tradicionales**

La proliferación de mediciones digitales no anula la importancia de los métodos clásicos de evaluación en psicología y psiquiatría, sino que se posiciona como un complemento poderoso. En [19] se hace hincapié en la necesidad de integrar autoinformes y observaciones clínicas tradicionales con la analítica de datos pasivos, para enriquecer la exactitud diagnóstica.

1. **Triangulación de información:** Cuando las métricas digitales coinciden con los reportes de familiares y los resultados de cuestionarios validados, se incrementa la confianza en el diagnóstico.
2. **Contextualización:** La labor de psicólogos y psiquiatras es situar los datos en la narrativa vital del paciente, entendiendo sus circunstancias personales, su historia de enfermedad y otros factores que los algoritmos no captan por sí solos.
3. **Identificación de perfiles:** La convergencia metodológica podría descubrir perfiles específicos de usuarios, por ejemplo, personas que subreportan síntomas en autoinformes pero cuyos datos digitales revelan inactividad social o problemas de sueño.

Para alcanzar este ideal, se precisa capacitación constante de los profesionales. Muchos clínicos, formados en marcos tradicionales, desconocen las posibilidades y limitaciones del análisis de datos masivos. De igual forma, los analistas e ingenieros han de interiorizar el razonamiento clínico y la complejidad del sufrimiento psíquico, para no caer en interpretaciones simplistas basadas exclusivamente en correlaciones estadísticas [19].

### **6.12. Movilización General del Conocimiento y Sigüientes Pasos**

#### *Difusión y Estándares de Buenas Prácticas*

La adopción responsable del fenotipado digital requiere la movilización del conocimiento. Esto implica la organización de congresos, cursos de formación, talleres y guías de buenas prácticas para que los investigadores y clínicos conozcan las últimas metodologías y se unan a redes

colaborativas. En [3] sugiere establecer proyectos piloto de bajo riesgo que sirvan como modelos de referencia, facilitando la transición a escalas mayores.

Crear estándares de interoperabilidad y protocolos de intercambio seguro de datos también favorece que distintos centros de investigación compartan y comparen resultados, acelerando la validación y la confianza en las herramientas desarrolladas. Varias asociaciones profesionales y organismos gubernamentales están empezando a elaborar lineamientos internacionales para normalizar los enfoques de recolección, almacenamiento, procesamiento y reporte de información fenotípica digital [3].

### Políticas Públicas y Financiamiento

Otro eslabón esencial es la voluntad política y el respaldo de entidades de financiamiento, tanto públicas como privadas. Los proyectos de fenotipado digital pueden requerir inversiones considerables en infraestructura tecnológica (almacenamiento en la nube, servidores de análisis), formación de profesionales y ensayos clínicos de larga duración. Cuando los gobiernos o agencias internacionales de salud promueven y financian tales iniciativas, se facilita el surgimiento de consorcios que incluyan hospitales, universidades y empresas tecnológicas, multiplicando el impacto y la calidad de la investigación [3].

Aunado a ello, las políticas públicas pueden regular el uso comercial de datos de fenotipado digital para fines publicitarios o no consensuados, asegurando que la sociedad se beneficie de los avances científicos sin sacrificar la privacidad y la dignidad de las personas [3].

### 6.13. Conclusiones

El fenotipado digital encarna una oportunidad sin precedentes para profundizar en el estudio del comportamiento humano y mejorar la calidad de la atención en salud mental y otras áreas médicas. Al mismo tiempo, su despliegue a gran escala expone la necesidad imperiosa de un abordaje verdaderamente interdisciplinario, donde converjan:

- **Psiquiatras y psicólogos**, que aportan la perspectiva clínica y teórica sobre los trastornos mentales y el comportamiento.
- **Ingenieros y científicos de datos**, encargados de desarrollar aplicaciones robustas y modelos algorítmicos que puedan procesar la avalancha de información recolectada.

- **Juristas y expertos en ética**, que velan por el cumplimiento de regulaciones y protegen la privacidad, la autonomía y los derechos fundamentales de los usuarios.
- **Sociólogos, antropólogos y humanistas**, que recuerdan la complejidad cultural y la diversidad de realidades que subyacen a los datos cuantitativos, evitando reduccionismos y sesgos.

La colaboración entre ciencia y sector comercial describe un escenario promisorio, pues la industria puede acelerar la transferencia de los avances del laboratorio a la práctica, siempre que se establezcan marcos de responsabilidad y transparencia. En paralelo, la ampliación a campos como la medicina interna, la rehabilitación o la educación evidencia la versatilidad del fenotipado digital para capturar aspectos conductuales y fisiológicos de manera continua.

Sin embargo, la adopción masiva de estas tecnologías exige cautela. Es fundamental vigilar la calidad de los datos, la eficacia real de los algoritmos y, sobre todo, la protección de la autonomía e intimidad de los individuos. El fenotipado digital puede volverse una poderosa herramienta de prevención y diagnóstico si se enmarca en principios éticos sólidos y no se transforma en un mecanismo de control o discriminación.

En definitiva, el futuro del fenotipado digital depende de cómo y quién lo implemente. Las sinergias interdisciplinarias, las políticas responsables y la evolución constante de metodologías colaborativas crearán las condiciones para que esta aproximación se consolide como un pilar de la salud y el bienestar en el siglo XXI, al tiempo que se respete la diversidad y la dignidad de las personas.

## 7. Tecnologías para el fenotipado digital

### 7.1. Herramientas de desarrollo de software

La fenotipificación digital es un campo en rápida evolución, todavía en su fase experimental, donde el desarrollo de software juega un papel crucial [8]. Las herramientas de desarrollo de software en este ámbito están diseñadas para capturar y analizar datos complejos de comportamiento humano en tiempo real mediante dispositivos personales como teléfonos inteligentes y wearables [8]. Sin embargo, uno de los desafíos más importantes es la necesidad de producir hallazgos replicables que puedan ser utilizados en aplicaciones clínicas [8]. Esto implica que el software debe ser robusto, preciso y capaz de manejar grandes volúmenes de datos [8]. Además, debe ser flexible para adaptarse a las necesidades cambiantes de los investigadores y las especificaciones de los estudios clínicos [8]. Las herramientas también deben cumplir con altos estándares de seguridad y privacidad, especialmente cuando se manejan datos sensibles de salud [8]. En resumen, el desarrollo de software en fenotipificación digital no solo se trata de la creación de aplicaciones y plataformas, sino de diseñar soluciones sostenibles, escalables y seguras, capaces de integrarse de manera efectiva en el entorno clínico y de investigación [8].

En [8] se mencionan herramientas utilizadas para la recolección y análisis de datos en el contexto del fenotipado digital.

- **BEHAPP:** Es una herramienta que permite la monitorización pasiva del comportamiento humano mediante dispositivos móviles, principalmente *smartphones* [8]. Esta plataforma no solo recopila datos de comportamiento en tiempo real, sino que también los analiza para proporcionar información relevante sobre patrones de comportamiento y otros fenómenos clínicos [8].
- **MindLamp:** Es una plataforma desarrollada por la División de Psiquiatría Digital del Beth Israel Deaconess Medical Center [8]. Al igual que BEHAPP, MindLamp permite la monitorización pasiva del comportamiento, centrándose en la recolección de datos a través de dispositivos móviles para estudiar el comportamiento humano en tiempo real [8].
- **Beiwe:** Es una plataforma desarrollada por el Onnela Laboratory en la Harvard T.H. Chan School of Public Health [8]. Esta herramienta se utiliza para la monitorización

digital del comportamiento humano y es conocida por su capacidad de recolectar múltiples flujos de datos a través de teléfonos inteligentes [8].

- **AWARE Framework:** Es un *framework* para la instrumentación del contexto móvil que permite a los investigadores recopilar datos de comportamiento en tiempo real desde dispositivos móviles [8]. Es conocido por su flexibilidad y capacidad para integrarse con otros sistemas de investigación [8].
- **Behapp-data-kit:** Este es un kit de herramientas internas desarrollado como parte del proyecto BEHAPP, destinado a facilitar la carga y exploración de datos para investigadores mientras mantiene altos niveles de seguridad [8].

Estas herramientas y plataformas son ejemplos clave de cómo se están utilizando tecnologías avanzadas en la fenotipificación digital para recopilar, analizar y aplicar datos de comportamiento en contextos clínicos y de investigación [8]. Cada una de estas herramientas tiene características únicas que las hacen adecuadas para diferentes tipos de estudios y aplicaciones en el ámbito de la salud mental y el comportamiento humano [8].

Por su parte [21] menciona el uso de MATLAB como una herramienta clave para la implementación de modelos y el análisis en el contexto de sensores móviles. Específicamente, se hace referencia a MATLAB en la siguiente cita:

“El libro apoya los trabajos propuestos proporcionando teoría y derivaciones de ecuaciones detrás de los conceptos, pasos de diseño y consejos para la implementación de la programación de teléfonos inteligentes, y el uso de MATLAB para la presentación del análisis de rendimiento.” [21].

MATLAB es una plataforma de programación y un entorno de desarrollo diseñado específicamente para el cálculo numérico, el análisis de datos y la visualización. Es ampliamente utilizado en áreas como ingeniería, matemáticas, física e informática para el desarrollo de algoritmos, la creación de modelos y la realización de simulaciones [21].

En el contexto de la detección consciente del contexto en dispositivos móviles, MATLAB se emplea para desarrollar y analizar algoritmos y modelos, proporcionando una base sólida para el cálculo numérico y la simulación de estos sistemas [21].

Según [3], la reproducibilidad en la investigación puede facilitarse mediante el uso de herramientas de gestión de software que permitan documentar y replicar código y resultados de

manera estructurada. En este sentido, se destacan plataformas como Docker, Packrat y GitHub, que ofrecen soluciones para asegurar que los estudios puedan reproducirse en diferentes entornos sin problemas de compatibilidad [3].

Además, [3] menciona que los Sistemas de Información Geográfica (SIG o GIS, por sus siglas en inglés) juegan un papel clave en el procesamiento y análisis de datos espaciales en la investigación. Existen diversas herramientas en esta área, que van desde soluciones comerciales como ArcGIS y ENVI, hasta opciones de código abierto como QGIS, GRASS GIS y GeoDa. También se pueden utilizar lenguajes de programación con módulos geográficos, como Python y R, así como bases de datos espaciales como PostgreSQL/PostGIS y Oracle Spatial, que permiten gestionar grandes volúmenes de información geoespacial [3].

Herramientas clave en el desarrollo y análisis de datos:

- **Docker:** Plataforma de contenedores que permite a los desarrolladores empaquetar aplicaciones junto con todas sus dependencias en un contenedor que puede ejecutarse de manera consistente en distintos entornos, ya sea en servidores locales o en la nube. En este contexto, se usa para facilitar la reproducibilidad de los estudios, asegurando que el entorno de desarrollo sea el mismo para todos los investigadores y permitiendo replicar experimentos y análisis sin discrepancias [3].
- **Packrat:** Sistema de gestión de paquetes y dependencias para R, diseñado para mantener un entorno de desarrollo consistente, asegurando que los análisis sean reproducibles. Al congelar las versiones de los paquetes utilizados en un proyecto, permite que los resultados puedan replicarse en el futuro sin problemas de compatibilidad [3].
- **GitHub:** Plataforma de control de versiones basada en Git, que facilita la colaboración en proyectos de software. Se menciona como una herramienta clave para documentar y compartir código, asegurando que otros investigadores puedan acceder a él y replicar estudios, promoviendo la transparencia y reproducibilidad en la investigación [3].
- **Geographic Information Systems (GIS):** Los Sistemas de Información Geográfica (SIG) son programas diseñados para capturar, almacenar, analizar y visualizar datos geoespaciales. En la investigación, permiten estudiar patrones de comportamiento en relación con la ubicación y el entorno, proporcionando herramientas avanzadas para el procesamiento de datos espaciales [3].

- **ArcGIS y ENVI:** ArcGIS es un sistema comercial de SIG desarrollado por ESRI, ampliamente utilizado para la creación de mapas y análisis geospaciales. ENVI, por otro lado, es un software especializado en procesamiento de imágenes geospaciales, empleado en investigaciones que requieren análisis detallados de datos espaciales, como el impacto del entorno físico en el comportamiento humano [3].
- **QGIS:** Aplicación de código abierto que ofrece herramientas para visualizar, editar y analizar datos geospaciales, siendo una alternativa gratuita a ArcGIS [3].
- **GRASS GIS:** Sistema de información geográfica de código abierto, utilizado para la gestión y análisis de datos geospaciales y la generación de mapas. Es especialmente útil para el análisis de grandes volúmenes de datos espaciales y estudios que requieren procesamiento avanzado [3].
- **GeoDa:** Software de código abierto diseñado para realizar análisis espaciales estadísticos. Permite explorar patrones de datos geográficos y realizar análisis de autocorrelación espacial [3].
- **Python y R:** Ambos lenguajes de programación cuentan con módulos especializados para el análisis geoespacial. Python ofrece bibliotecas como *GeoPandas* y *Shapely*, mientras que R incluye paquetes como *sf* y *raster*, los cuales permiten la manipulación y análisis de datos espaciales en estudios científicos [3].
- **PostgreSQL/PostGIS:** PostgreSQL es un sistema de gestión de bases de datos relacional de código abierto, mientras que PostGIS es una extensión que añade soporte para datos geospaciales. Se utiliza para almacenar, consultar y procesar información geográfica de manera eficiente [3].
- **Oracle Spatial:** Extensión del sistema de bases de datos Oracle, diseñada para manejar grandes volúmenes de datos geospaciales en entornos de alto rendimiento. Se emplea en aplicaciones que requieren análisis avanzados de datos espaciales, integrando modelos SIG en bases de datos relacionales [3].

## 7.2. Técnicas de análisis y procesamientos de datos

El procesamiento y análisis de datos son fundamentales para extraer información valiosa a partir de grandes volúmenes de datos [5]. Estos procesos incluyen la transformación de datos brutos en información útil, la detección de patrones, tendencias y correlaciones, y la generación de *insights* accionables para la toma de decisiones y la resolución de problemas [5].

Los pasos iniciales en el procesamiento de datos incluyen la recolección y almacenamiento, los cuales pueden provenir de sensores, bases de datos, redes sociales y otras plataformas digitales [5]. Posteriormente, la preparación de los datos es una fase crítica para garantizar su calidad y fiabilidad, lo que implica la limpieza de datos con el fin de eliminar inconsistencias, errores y valores atípicos [5].

En el contexto de la fenotipificación digital, el procesamiento de datos abarca varias etapas críticas [6]. En primer lugar, se lleva a cabo la recolección de datos en bruto desde diversas fuentes, como sensores de dispositivos móviles y portátiles, registros de ubicación, actividades en redes sociales y eventos de interacción con el dispositivo [6]. Aunque estos datos representan la base fundamental de la fenotipificación digital, en su estado inicial no son directamente inteligibles ni útiles sin una adecuada transformación y contextualización [6].

Para convertir los datos en bruto en información procesable, se combinan registros de diferentes fuentes, organizándolos en categorías más comprensibles y relevantes. Por ejemplo, los datos de sensores como acelerómetros y giroscopios pueden ser transformados en métricas de actividad física, mientras que los registros de ubicación permiten analizar patrones de movilidad. De manera similar, los registros de llamadas y el uso de redes sociales pueden proporcionar información sobre la sociabilidad virtual y física de los pacientes [6].

Finalmente, se emplean herramientas computacionales avanzadas para investigar y fusionar estos datos transformados, permitiendo detectar patrones de ánimo y comportamiento que contribuyen a la construcción del fenotipo digital del paciente [6]. Este enfoque no solo mejora la precisión del diagnóstico y el monitoreo de los trastornos del ánimo, sino que también abre nuevas oportunidades para personalizar las intervenciones terapéuticas [6].

## **Etapas en el análisis y procesamiento de datos**

Según [1], el proceso de sensorización móvil se compone de tres etapas principales:

- **Recolección de datos:** El primer paso en el proceso de sensorización móvil es la recolección de datos crudos provenientes de los sensores [1]. Aunque alguna información puede no ser recopilada exclusivamente por un solo sensor, casi toda la información puede obtenerse cuando se combina con suficientes sensores relevantes. La recolección de datos a menudo se realiza utilizando software instalado en el dispositivo, como AWARE o INSIGHTS [1].

- **Preprocesamiento de datos:** Una vez que los datos han sido recolectados exitosamente, es necesario prepararlos para su análisis posterior [1]. Esto incluye la limpieza de los datos, ya que, en el ámbito de la salud, los datos suelen ser altamente heterogéneos, ruidosos, inconclusos y fragmentarios. En el siguiente paso, se pueden extraer características mediante estadísticas descriptivas de series temporales de sensores [1]. Por ejemplo, los datos crudos sobre uso del teléfono pueden convertirse en métricas como número y duración de llamadas entrantes, mensajes de texto enviados, tiempo promedio de uso del dispositivo o la frecuencia con la que la pantalla se enciende o apaga. El propósito de este paso es reducir el ruido y mejorar la precisión, aunque también puede conllevar una pérdida de sensibilidad en los datos [1].
- **Análisis de datos:** En la última etapa, los datos preprocesados se someten a análisis avanzado [1]. La minería de datos se encarga de extraer conocimiento a partir de grandes volúmenes de datos, utilizando técnicas de aprendizaje automático (ML) y aprendizaje profundo (DL) [1]. El objetivo de estos métodos es identificar patrones, tendencias o correlaciones en grandes conjuntos de datos y emplear los patrones detectados para la construcción de modelos predictivos en el ámbito de la salud [1].

En [3] se menciona el concepto de capas y resalta que es fundamental en el procesamiento de información espacial. Permite trabajar en diferentes capas con sus respectivos modelos de datos y visualizar diferentes fuentes de datos al mismo tiempo.

- **Capa Inferior (Realidad Compleja):** En esta capa se encuentran los datos en bruto recopilados desde múltiples fuentes, tales como sensores de teléfonos inteligentes, GPS y registros de actividad en redes sociales. Estos datos representan la información sin procesar que constituye la base del análisis en fenotipificación digital [3].
- **Capas Intermedias:** Esta sección comprende la transformación de los datos en bruto en información procesada, lo que permite identificar patrones significativos. Entre estos se incluyen patrones de movimiento, uso del dispositivo, patrones de sueño, entre otros, los cuales pueden correlacionarse con el bienestar mental y proporcionar indicios sobre el estado emocional del individuo [3].
- **Capa Superior (Análisis de NDVI):** En esta capa se lleva a cabo un análisis avanzado que permite calcular un "índice de bienestar", análogo al NDVI (Índice de Vegetación de Diferencia Normalizada) en el análisis de la vegetación. Este índice tiene como

objetivo medir factores del estado mental, como la interacción social y el nivel de actividad física, los cuales pueden estar relacionados con el bienestar emocional [3].

Por último, en [8] se mencionan las siguientes etapas:

- **Recolección y encriptación de datos:** El proceso de datos comienza con la recolección de información a través de dispositivos personales, como *smartphones*. Para garantizar la seguridad de la información desde su origen, estos datos se encriptan directamente en el dispositivo antes de su transmisión y almacenamiento [8]. La encriptación se realiza mediante técnicas asimétricas, en las cuales cada participante posee un par de claves (una pública y una privada), utilizadas para encriptar y desencriptar los datos de manera segura [8].
- **Transmisión y almacenamiento seguro:** Una vez encriptados, los datos se transmiten a un servidor central, donde se almacenan en bases de datos separadas para cada estudio. Este enfoque de almacenamiento segmentado garantiza que los investigadores solo tengan acceso a los datos específicos de los estudios en los que están trabajando, minimizando así el riesgo de exposición de datos sensibles [8].
- **Análisis de datos en la zona privada:** El procesamiento y análisis de los datos se lleva a cabo en una "zona privada" del sistema, diseñada para garantizar un entorno seguro [8]. Durante esta fase, los datos son desencriptados temporalmente para su análisis, permitiendo la ejecución de algoritmos avanzados que extraen características comportamentales. Este método permite que la información sea utilizada de manera eficiente sin comprometer la privacidad y seguridad de los participantes [8].

### **Técnicas de preprocesamiento, procesamiento y análisis de los datos**

El análisis de sensibilidad y la gestión de la incertidumbre son aspectos fundamentales en el campo de la computación *context-aware*. Estos enfoques permiten evaluar cómo las variaciones en los datos de entrada pueden afectar los resultados de un modelo, lo que es crucial en aplicaciones donde la precisión y confiabilidad de los datos contextuales son esenciales [2].

Por ejemplo, en bases de datos contextuales, el análisis de sensibilidad puede ayudar a determinar la robustez de las respuestas de consulta frente a datos imperfectos o inciertos [2]. Además, el manejo de grandes volúmenes de datos contextuales requiere estrategias eficientes de almacenamiento, compresión y discretización para mantener la precisión de los datos y

reducir la incertidumbre inherente. Estas técnicas permiten que los sistemas respondan proactivamente a cambios en el contexto, adaptándose dinámicamente a las necesidades del usuario [2].

#### Técnicas de preprocesamiento de datos

Estas técnicas se aplican antes del análisis, con el objetivo de optimizar la calidad y estructura de los datos para mejorar su utilidad en el procesamiento posterior. Esto incluye la limpieza de datos, normalización para escalarlos a un rango común, imputación de valores faltantes, extracción de características relevantes y suavizado para reducir el ruido. La segmentación divide los datos en segmentos significativos, y la reducción de dimensionalidad simplifica el análisis eliminando características redundantes [5].

**Selección de atributos:** Esta técnica se utiliza para identificar qué atributos de los datos son más importantes o relevantes para un análisis específico. Al reducir el número de atributos a considerar, se mejora la eficiencia del análisis y se reduce la complejidad computacional [2]. En sistemas contextuales, los datos pueden provenir de diversas fuentes con múltiples atributos. La selección de atributos permite identificar los más relevantes, reduciendo la dimensionalidad del problema y previniendo el sobreajuste del modelo [2]. Por ejemplo, en un sistema que monitorea el estado de salud de un usuario a través de múltiples sensores (ritmo cardíaco, temperatura corporal, nivel de actividad, etc.), no todos estos datos son igualmente relevantes para predecir una condición específica [2]. La selección de atributos ayuda a determinar qué datos son más predictivos para detectar episodios de estrés, permitiendo al sistema enfocarse en ellos para mejorar la precisión y eficiencia [2].

**Provenance (procedencia) y linaje de los datos:** Se refiere a la capacidad de rastrear el origen de los datos y cómo estos han sido modificados o procesados a lo largo del tiempo. Esto es crucial para entender cómo las variaciones en los datos de entrada pueden influir en los resultados obtenidos [2]. Saber la procedencia de un dato puede ayudar a diagnosticar problemas, como si un resultado anómalo en el sistema es causado por un error en la fuente original del dato o una transformación incorrecta en algún punto del proceso [2].

#### Técnicas de procesamiento de datos

Estas técnicas se encargan de estructurar y transformar los datos, haciéndolos más adecuados para el análisis posterior.

**Causalidad y responsabilidad:** Consiste en identificar relaciones causales entre distintos elementos de datos y determinar qué partes de los datos son responsables de ciertos resultados o comportamientos observados en el sistema [2]. En un sistema de recomendación basado en el contexto, si una recomendación es ineficaz o incorrecta, el análisis de causalidad puede ayudar a identificar qué factores contextuales (ubicación del usuario, hora del día, preferencias previas, etc.) contribuyeron a la falla [2]. Este enfoque permite ajustar el sistema para mejorar futuras recomendaciones, asegurando que sean más precisas y personalizadas [2].

**Modelo de Mezcla Gaussiana (GMM):** Este es un modelo probabilístico que asume que los datos pueden ser representados como una combinación de varias distribuciones gaussianas. Es utilizado para modelar la distribución de los datos de entrada (en este caso, las lecturas del acelerómetro) y para ayudar en la clasificación de las posturas [21].

**Búsqueda de k-Vecinos más Cercanos (k-NN):** Este es un algoritmo de clasificación que clasifica un punto de datos basado en los datos más cercanos a él. En este contexto, se podría utilizar para identificar la postura actual del usuario comparando las lecturas del acelerómetro con las lecturas de referencia de posturas conocidas [21].

**Análisis Discriminante Lineal (LDA):** Este es un método de reducción de dimensionalidad que se utiliza para encontrar una combinación lineal de características que mejor separa dos o más clases. En el caso de la detección de posturas, LDA puede ayudar a distinguir entre diferentes tipos de movimientos del usuario [21].

**Gaussian Naive Bayes:** Un estudio exploró el uso de aprendizaje automático aplicado a datos de resonancia magnética funcional (fMRI) para investigar las representaciones neuronales de conceptos relacionados con la vida y la muerte en jóvenes [6]. El objetivo era clasificar a individuos con ideación suicida, utilizando un algoritmo Gaussian Naive Bayes en una muestra de 17 participantes con ideación suicida y 17 controles sanos [6]. El modelo alcanzó una precisión del 91%, demostrando una alta efectividad en la clasificación de los participantes [6].

**Naive Bayes:** Otro estudio tuvo como objetivo identificar una firma textual de riesgo de suicidio, empleando un algoritmo de aprendizaje automático Naive Bayes para la clasificación de textos provenientes de cartas y diarios de Virginia Woolf. El modelo logró una precisión equilibrada de aproximadamente 80% en la predicción del comportamiento suicida, validando su utilidad en el análisis de textos con posibles señales de ideación suicida [6].

**Random Forest:** Un estudio de caso-control entrenó un modelo *Random Forest* utilizando datos de Twitter con el objetivo de predecir el estado de ideación suicida [6]. Para ello, el modelo analizó un total de 512,526 tuits de 283 casos de ideación suicida y 3,518,494 tuits de 2655 controles [6]. El rendimiento del modelo fue evaluado mediante el área bajo la curva ROC (AUC), alcanzando un AUC de 0.88, lo que indica una alta capacidad predictiva [6].

**Regresión:** Una investigación longitudinal desarrolló modelos de aprendizaje automático para predecir la muerte por suicidio tras una hospitalización psiquiátrica, empleando árboles de regresión y algoritmos de regresión penalizada [6]. El análisis se basó en datos administrativos de 40,820 soldados del ejército de EE.UU., alcanzando AUCs de aproximadamente 0.85, lo que sugiere una elevada precisión en la predicción del riesgo de suicidio en esta población [6].

**Método de Clasificación en Línea para la Detección de Posturas:** El estudio propone un método de clasificación en línea, diseñado para ser computacionalmente ligero y operar sin necesidad de datos previos o información a priori [21]. Esto significa que el sistema puede clasificar las posturas del usuario en tiempo real, sin requerir un conjunto de datos de entrenamiento inicial [21]. Este enfoque resulta especialmente útil en escenarios donde el acceso a grandes volúmenes de datos previamente etiquetados es limitado o inexistente. El método propuesto tiene la capacidad de operar de manera independiente, diferenciando entre distintos estados del usuario, como estar sentado, de pie, caminando o corriendo, sin necesidad de intervención externa [21]. Esta característica es fundamental para aplicaciones móviles que requieren funcionar en tiempo real y de manera continua, sin la necesidad de ajustes o intervención constante por parte del usuario [21]. Además de su funcionalidad autónoma, este método también puede asistir a otros modelos de clasificación supervisada. Es capaz de generar clases de datos de entrenamiento y matrices de entrada, lo que contribuye a mejorar la precisión y eficiencia de algoritmos que requieren una fase de entrenamiento [21]. Esto lo convierte en una herramienta versátil dentro del ecosistema de análisis de datos, ya que facilita la creación de conjuntos de datos más robustos para modelos de aprendizaje supervisado [21].

#### Técnicas de análisis de datos

Estas técnicas se enfocan en la evaluación de los datos para extraer información significativa, patrones y tendencias. Medición y cálculo de la sensibilidad: Evalúa cuán susceptible es un modelo o un sistema a cambios en sus datos de entrada [2]. En bases de datos y sistemas contextuales, esto es vital para garantizar que el sistema maneje incertidumbres en los datos sin

producir resultados incorrectos o inestables [2]. Por ejemplo, si un sistema de monitoreo ambiental utiliza sensores para medir la calidad del aire, la medición de la sensibilidad puede revelar cómo pequeñas variaciones en las lecturas de los sensores (ruido o errores) afectan las alertas generadas por el sistema. Si el sistema es demasiado sensible, podría generar falsas alarmas, mientras que, si es poco sensible, podría no detectar peligros reales. Ajustar esta sensibilidad es crucial para el rendimiento y la fiabilidad del sistema [2].

### **7.3.Seguridad, privacidad y protección de datos**

#### **Introducción a la Privacidad y Seguridad en actividades en línea**

En la actualidad, la privacidad y seguridad juegan un papel crucial en la protección de la información personal y en el mantenimiento de la confianza de los usuarios en entornos digitales [5]. Es esencial que las personas adopten medidas para salvaguardar su información, tales como el uso de contraseñas seguras, el cifrado de datos y la concienciación sobre la información que se comparte en línea [5].

La protección de información sensible se ha vuelto más relevante que nunca debido al incremento de ataques cibernéticos y violaciones de datos [5]. Datos como números de seguridad social, registros financieros e información médica pueden ser utilizados con fines fraudulentos si llegan a manos equivocadas [5]. Por ello, es fundamental que tanto individuos como organizaciones implementen estrategias para garantizar la seguridad de estos datos [5].

Los datos recopilados a través de teléfonos inteligentes y otros dispositivos con sensores, como rastreadores de actividad física, están directamente vinculados o pueden ser relacionados con individuos [1]. En este sentido, los profesionales encargados de recopilar, procesar y almacenar esta información tienen obligaciones tanto morales como legales para manejarla de manera responsable [1]. Esto es especialmente crítico cuando se trata de datos relacionados con la salud o trastornos mentales, ya que cualquier mal manejo de la información podría afectar la privacidad y el bienestar del individuo [1].

El derecho a la protección de datos personales ha sido reconocido como un derecho humano fundamental, y está formalmente incluido en la Carta de los Derechos Humanos de la Unión Europea [1].

Por último, [2] menciona que la privacidad y la protección de datos son aspectos cruciales en los sistemas contextualmente conscientes, donde se recopila y maneja una gran cantidad de

datos personales y sensibles. Se tratará este tema en profundidad, destacando varias técnicas y enfoques que se utilizan para garantizar que la información del usuario esté protegida mientras se mantiene la funcionalidad del sistema.

## **Privacidad como concepto legal**

Muchos países han regulado diferentes aspectos de la privacidad a través de sus leyes. La protección legal de los datos personales se conoce como protección de datos. Uno de los marcos más completos en esta materia es el Reglamento General de Protección de Datos (GDPR) de Europa, el cual se aplica en toda la Unión Europea desde 2018 y ha unificado los regímenes de protección de datos que anteriormente variaban entre los distintos países europeos [1]

El GDPR, o "Reglamento (UE) 2016/6791", regula el procesamiento de datos personales de individuos dentro de la UE, ya sea realizado por una persona, una empresa u otra organización. Según lo descrito en su Artículo 4, los datos personales incluyen cualquier información que pueda estar directa o indirectamente relacionada con una persona. Además, el procesamiento de datos abarca actividades como la recopilación, almacenamiento y análisis de estos datos [1].

El GDPR impone una serie de requisitos tanto a quienes procesan los datos (denominados "procesadores de datos") como a quienes son responsables del procesamiento (denominados "controladores de datos"). Asimismo, otorga a los sujetos de datos, es decir, las personas a quienes se refieren los datos, una serie de derechos, como el derecho a ser informados, el derecho a acceder a su información, el derecho a la eliminación de datos procesados y el derecho a oponerse al procesamiento de datos [1].

Para que el procesamiento de datos personales sea legal, debe basarse en una de las bases jurídicas establecidas por el GDPR, como la obligación legal, el cumplimiento de un contrato o el consentimiento informado del sujeto de los datos. En los estudios psicológicos, este último suele ser el principal fundamento para el procesamiento de datos personales. No obstante, en algunos casos también puede invocarse el interés público en la investigación. Sin embargo, ni el consentimiento ni el interés público eximen al controlador o procesador de datos del cumplimiento de todas las obligaciones establecidas en el GDPR. Entre estas obligaciones se incluyen los derechos de información de los sujetos de datos, como el acceso a su información en un formato portátil y la posibilidad de revocar su consentimiento en cualquier momento [1].

Este marco normativo plantea ciertos conflictos con prácticas como la gestión de datos de investigación y los principios de Ciencia Abierta, que promueven el archivo y el intercambio de datos para mejorar la reproducibilidad de los estudios científicos. La necesidad de equilibrar la protección de datos con la accesibilidad de la información en el ámbito académico sigue siendo un desafío clave en la actualidad [1].

### **Privacidad como concepto técnico**

Desde una perspectiva técnica, la privacidad se refiere a la implementación de medidas y prácticas destinadas a proteger los datos personales contra accesos no autorizados y posibles abusos [1]. Para lograrlo, se aplican medidas técnicas y organizativas de última generación, estas medidas incluyen diversas tecnologías y enfoques, como el cifrado de datos, el control de acceso, la minimización de datos y la anonimización o seudonimización de la información personal [1].

El cifrado es una técnica fundamental que convierte los datos en un formato ininteligible para cualquier persona que no posea la clave de descifrado correspondiente [1]. Gracias a este mecanismo, incluso si los datos son interceptados por terceros, no pueden ser fácilmente comprendidos ni utilizados, garantizando así un nivel adicional de seguridad [1].

El control de acceso regula quién puede acceder a los datos personales, limitando su disponibilidad únicamente a usuarios autorizados [1]. Este sistema puede reforzarse mediante autenticación multifactor, un método que exige a los usuarios validar su identidad utilizando múltiples factores de verificación antes de permitir el acceso a la información [1].

La minimización de datos es otro principio clave en la protección de la privacidad. Consiste en recopilar únicamente la cantidad mínima de datos necesarios para cumplir con un propósito específico [1]. Esta estrategia no solo reduce la cantidad de información en riesgo, sino que también facilita el cumplimiento de regulaciones legales que limitan la recolección de datos innecesarios [1].

Por otro lado, la anonimización y la seudonimización son técnicas esenciales en la gestión de datos personales. La anonimización implica eliminar o modificar permanentemente cualquier información que pueda permitir la identificación directa o indirecta de los individuos [1]. En contraste, la seudonimización reemplaza los identificadores personales por seudónimos, lo que

disminuye los riesgos de privacidad, aunque aún permite vincular los datos a un individuo bajo ciertas condiciones controladas [1].

Desde una perspectiva técnica, la implementación de estas medidas es crucial para mitigar los riesgos inherentes al procesamiento de datos personales. Además, adoptar prácticas robustas de privacidad no solo refuerza la seguridad de la información, sino que también fomenta la confianza de los usuarios en los sistemas y aplicaciones que manejan sus datos [1].

## **Medidas de protección y prevención**

Para proteger la información sensible, es esencial emplear métodos como el cifrado, los controles de acceso y la realización periódica de auditorías de seguridad. Las aplicaciones y plataformas deben garantizar la seguridad de los datos mediante la implementación de medidas robustas que prevengan tanto el acceso no autorizado como los ataques cibernéticos [(Kuldeep Singh Kaswan, 2024)].

En el contexto de la sensorización móvil, la privacidad adquiere un papel fundamental, especialmente en relación con los principios establecidos por el GDPR. Se han identificado diversas estrategias de diseño de privacidad propuestas por Hoepman (2014), como Minimizar, Separar, Agregar, Ocultar, Informar, Controlar, Hacer cumplir y Demostrar, las cuales proporcionan un marco para la implementación de medidas técnicas y organizativas que garantizan la protección de los datos personales durante su procesamiento [1].

Un ejemplo práctico de estas estrategias es la de Agregación, la cual sugiere procesar los datos al nivel más alto de agregación y con el menor detalle posible. Este enfoque resulta particularmente útil para la gestión de datos sensibles, como perfiles personales o información relacionada con la salud, ya que no solo protege la privacidad del usuario, sino que también optimiza el uso de los recursos, al reducir la cantidad de información que necesita ser transmitida y almacenada [1].

El consentimiento del usuario y la gestión de controles de acceso son factores clave en la privacidad de los datos. Los sistemas deben contar con mecanismos para obtener el consentimiento informado de los usuarios y, además, restringir el acceso a los datos sensibles exclusivamente al personal autorizado [5].

Las estrategias de retención y eliminación de datos son fundamentales para una gestión segura de la información. Es necesario establecer directrices claras sobre el tiempo de almacenamiento

de los datos y su eventual eliminación, con el fin de minimizar los riesgos de violaciones de datos y accesos no autorizados [5].

La anonimización es otra técnica importante que se utiliza para proteger la privacidad del usuario. Consiste en eliminar o modificar los identificadores personales de los datos para que no se pueda rastrear la información de vuelta a un individuo específico. Esta técnica es especialmente útil en el manejo de grandes volúmenes de datos en aplicaciones de análisis donde es necesario proteger la identidad del usuario, pero aún se necesita utilizar los datos para extraer información útil [2].

Las bases de datos hipocráticas integran principios de privacidad directamente en su diseño. Estas bases de datos están diseñadas para cumplir con las políticas de privacidad desde el momento en que los datos se recopilan hasta su eventual eliminación. Se enfocan en la autodeterminación informativa, donde los usuarios pueden especificar cómo se deben manejar sus datos, quién puede acceder a ellos y bajo qué condiciones [2].

La gestión del ciclo de vida de la información es una estrategia para manejar la información sensible desde su creación hasta su eliminación. En los sistemas contextualmente conscientes, es esencial asegurar que los datos ya no necesarios se eliminen de manera segura para evitar riesgos de seguridad. Esto incluye la implementación de políticas de retención de datos, donde los datos se mantienen solo por el tiempo que sean necesarios, y luego se eliminan o anonimizan [2].

La búsqueda encriptada como una técnica avanzada para proteger la privacidad durante las consultas de datos. Esto permite realizar búsquedas en bases de datos encriptadas sin necesidad de desencriptar los datos, protegiendo así la privacidad del usuario incluso durante la manipulación activa de los datos [2]. Ejemplo: Un sistema de búsqueda de historial médico podría permitir a los médicos buscar condiciones específicas en registros encriptados, garantizando que la información sensible no sea revelada más allá de lo necesario para la consulta específica [2].

El cumplimiento de regulaciones como el GDPR y el CCPA es fundamental para garantizar un manejo responsable de los datos y proteger los derechos de privacidad de los individuos. Para ello, las organizaciones deben realizar evaluaciones de impacto de privacidad y llevar a cabo auditorías periódicas, con el fin de identificar riesgos potenciales y detectar posibles vulnerabilidades [5].

Cuando se trabaja con Información de Salud Protegida (PHI, por sus siglas en inglés), es imprescindible aplicar un enfoque sistemático que garantice la seguridad de los datos. Esto implica la participación de personas, procesos y sistemas de TI, los cuales deben ser auditados anualmente de manera independiente. Existen diversas organizaciones de certificación, entre ellas la Organización Internacional de Normalización (ISO) y The Health Information Trust Alliance (HITRUST) [1].

Dentro de la familia de estándares ISO 27000, se certifican las políticas, procedimientos y controles de una organización para garantizar la seguridad de los activos de información. Por su parte, HITRUST es un estándar de certificación privado en los Estados Unidos, similar a ISO 27001, que certifica organizaciones en su marco de seguridad para la creación, acceso, almacenamiento o intercambio de información sensible. De hecho, obtener una certificación ISO 27001 o HITRUST suele ser un requisito previo antes de compartir información sensible, como PHI, entre distintas organizaciones [1].

## **8. Prueba de conceptos, prototipo funcional**

### **8.1. Introducción**

Este capítulo describe el desarrollo y resultados del proyecto cuyo objetivo fue crear un prototipo funcional de una aplicación que recolecta datos de sensores de un teléfono móvil. Este trabajo forma parte de una investigación sobre fenotipado digital, con énfasis en probar la viabilidad técnica de la recolección y almacenamiento de datos.

Motivación: Este proyecto se desarrolla como parte de una tesis académica, cuyo propósito es demostrar la viabilidad técnica de la recolección de datos.

Impacto esperado: Probar que es posible implementar una solución para recolectar datos de sensores y sentar las bases para trabajos futuros.

### **8.2. Objetivo del proyecto**

El objetivo fundamental de este proyecto es demostrar la viabilidad de capturar datos de al menos tres sensores de un dispositivo móvil con sistema operativo Android, transmitirlos a un sistema central y almacenarlos para su posterior análisis. A través de esta implementación, se busca validar la capacidad de recolección, procesamiento y persistencia de la información obtenida desde un entorno móvil, garantizando su trazabilidad y disponibilidad para futuras consultas.

Para ello, se desarrollará una aplicación Android capaz de registrar información proveniente del sensor de luz, la ubicación geográfica (latitud y longitud) y los eventos de bloqueo y desbloqueo de pantalla del dispositivo. Estos datos serán enviados de manera periódica a un *backend*, donde serán procesados y almacenados en una base de datos estructurada, permitiendo su posterior acceso y explotación.

El proyecto también tiene como propósito evaluar la efectividad y estabilidad de la infraestructura tecnológica utilizada, asegurando que los datos puedan ser recolectados de manera confiable y transmitidos de forma segura mediante una arquitectura robusta. Asimismo, se analizará la capacidad de la base de datos para gestionar estos registros de manera eficiente y optimizar su consulta y uso en reportes analíticos.

Otro aspecto clave es la validación de la integración de los datos almacenados con herramientas de visualización como Power BI, permitiendo generar reportes y representaciones gráficas que faciliten su interpretación y análisis. La definición de las métricas y visualizaciones quedará a criterio del equipo de desarrollo, priorizando aquellas que mejor representen la información obtenida.

En términos generales, este proyecto tiene como objetivo central comprobar la factibilidad técnica de implementar un sistema capaz de recopilar, almacenar y analizar información sensorial proveniente de dispositivos móviles. A partir de los resultados obtenidos, se podrán extraer conclusiones sobre la viabilidad de ampliar el alcance de la solución, optimizar su desempeño y extender su funcionalidad en futuras iteraciones o aplicaciones prácticas.

### **8.3. Alcance**

Este proyecto abarca el desarrollo e implementación de un sistema completo para la recolección, almacenamiento, procesamiento y visualización de datos obtenidos desde dispositivos móviles con sistema operativo Android. La solución contempla múltiples componentes tecnológicos que interactúan de manera integrada para garantizar la eficiencia en la captura de información, su procesamiento seguro y la posterior generación de reportes para su análisis.

En primer lugar, la aplicación Android será responsable de capturar y enviar datos provenientes de diferentes fuentes del dispositivo, incluyendo información del sensor de luz, coordenadas de ubicación geográfica (latitud y longitud) y eventos de bloqueo y desbloqueo de pantalla. El usuario tendrá la capacidad de habilitar o deshabilitar la transmisión de estos datos según su preferencia, permitiendo un mayor control sobre la información que se comparte con el sistema. La comunicación entre la aplicación y el *backend* se realizará a través de un canal seguro utilizando el protocolo HTTPS, garantizando la privacidad e integridad de los datos transmitidos. La frecuencia de envío de la información será cada 15 minutos, asegurando una actualización periódica de los registros en el sistema central.

El *backend* será desarrollado en Node.js y tendrá la responsabilidad de recibir, procesar y almacenar los datos enviados por la aplicación móvil. Para ello, se implementará una base de datos MySQL alojada en la nube de Azure, que permitirá la persistencia de la información de manera estructurada y eficiente. Esta base de datos deberá garantizar la integridad y disponibilidad de los datos, facilitando su posterior consulta y análisis. Además, el *backend*

estará desplegado en AWS, lo que permitirá una infraestructura escalable y confiable para soportar la carga del sistema.

En cuanto al procesamiento y explotación de los datos, el sistema incluirá la generación de reportes dinámicos utilizando Power BI. Estos reportes permitirán visualizar los datos recolectados, aplicando filtros personalizados para realizar análisis según diferentes criterios, como rangos de fechas o tipos de información recopilada. La definición de qué datos serán visualizados y en qué formato se presentarán quedará a criterio del equipo de desarrollo, priorizando la claridad y utilidad para los usuarios.

El alcance del proyecto también contempla la implementación de medidas de seguridad y control de acceso, asegurando que cada componente del sistema cumpla con los estándares adecuados para la protección de la información. Esto incluye la gestión segura de la autenticación de usuarios y la restricción del acceso a los datos según los permisos correspondientes.

En términos de infraestructura y despliegue, la solución se diseñará para operar de manera eficiente en entornos de producción, garantizando alta disponibilidad y rendimiento óptimo en cada uno de sus componentes. Se priorizará la escalabilidad del sistema, permitiendo su crecimiento y adaptación a futuras necesidades sin comprometer la estabilidad ni la seguridad de la información procesada.

En resumen, este proyecto tiene como objetivo proporcionar una solución integral para la recolección, almacenamiento, análisis y visualización de datos obtenidos desde dispositivos móviles Android. A través de una arquitectura robusta y tecnologías modernas, se busca garantizar la confiabilidad, seguridad y usabilidad del sistema, permitiendo que la información capturada sea utilizada de manera efectiva para la toma de decisiones y el análisis de tendencias en los datos recolectados.

## **8.4.Requerimientos**

### **Funcionales**

**RF1.** Envío de datos desde la aplicación Android

**Actor:** Aplicación Android

La aplicación debe capturar y enviar datos del sensor de luz, ubicación geográfica (latitud y longitud) y eventos de bloqueo y desbloqueo de pantalla del dispositivo a un sistema central.

El envío de estos datos debe realizarse de manera periódica o en tiempo real, asegurando que cada dato contenga una marca de tiempo para su correcta trazabilidad.

#### **RF2.** Almacenamiento de datos en la base de datos

**Actor:** No aplica

El *backend* debe recibir los datos enviados por la aplicación Android y almacenarlos en una base de datos. Cada registro debe incluir la información capturada, junto con su marca de tiempo, para garantizar su trazabilidad.

#### **RF3.** Habilitación y deshabilitación del envío de datos

**Actor:** Usuario

La aplicación debe permitir al usuario habilitar o deshabilitar el envío de datos del sensor de luz, ubicación geográfica (latitud y longitud) y eventos de bloqueo y desbloqueo de pantalla.

El usuario podrá gestionar esta configuración desde la interfaz de la aplicación, aplicando los cambios en tiempo real sin necesidad de reiniciar la aplicación.

#### **RF4.** Generación de reportes con los datos obtenidos

**Actor:** No aplica

El sistema debe generar un reporte que muestre los datos obtenidos desde la aplicación móvil. La información presentada y las visualizaciones utilizadas serán definidas por el equipo de desarrollo, asegurando que sean claras y relevantes para el análisis.

El reporte debe permitir la consulta de datos históricos y ofrecer opciones de filtrado según los criterios que el equipo considere más útiles para la interpretación de la información.

### **No funcionales**

#### **RNF1.** Tecnología del reporte

El sistema de reportes debe utilizar Power BI para la visualización y análisis de los datos obtenidos desde la aplicación móvil.

**RNF2.** Tecnología del *backend*

El *backend* debe ser desarrollado utilizando Node.js, asegurando compatibilidad con los demás componentes del sistema.

**RNF3.** Tecnología de la base de datos

El almacenamiento de datos debe realizarse en una base de datos MySQL, garantizando integridad y eficiencia en la gestión de la información.

**RNF4.** Infraestructura de la base de datos

La base de datos debe estar desplegada en la nube de Azure, asegurando disponibilidad y escalabilidad del servicio.

**RNF5.** Infraestructura del *backend*

El *backend* debe estar desplegado en la nube de AWS, asegurando alto rendimiento y disponibilidad del sistema.

**RNF6.** Plataforma de la aplicación móvil

La aplicación debe ser desarrollada para dispositivos con sistema operativo Android, garantizando compatibilidad con los dispositivos objetivo.

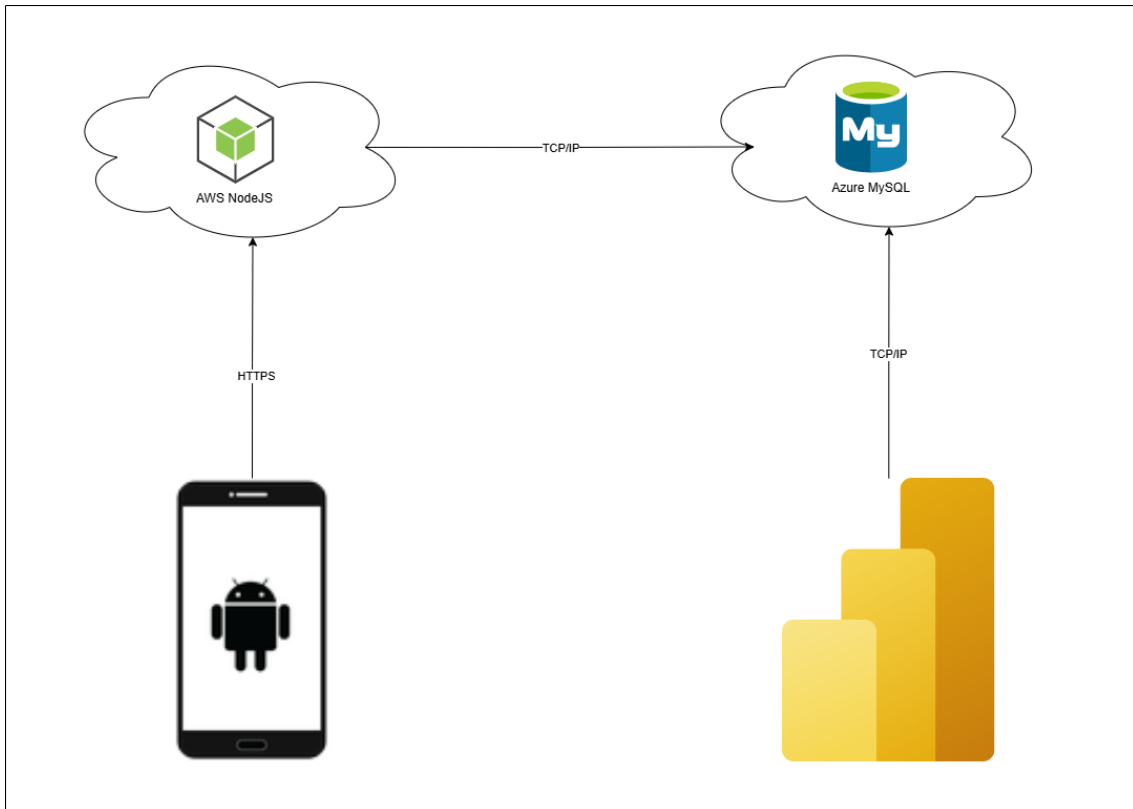
**RNF7.** Protocolo de comunicación

El envío de datos entre la aplicación y el *backend* debe realizarse utilizando el protocolo HTTPS, asegurando la seguridad y privacidad de la información transmitida.

**RNF8.** Frecuencia de envío de datos

La aplicación debe enviar los datos al *backend* cada 15 minutos, garantizando una actualización periódica de la información en el sistema.

## 8.5. Diagrama de despliegue



El diagrama muestra la arquitectura del sistema y la comunicación entre los distintos componentes. Explicación detallada:

### 1. Aplicación Android → *Backend* (AWS Node.js)

**Protocolo:** HTTPS

**Descripción:**

La aplicación Android captura datos del sensor de luz, ubicación geográfica (latitud y longitud) y eventos de bloqueo/desbloqueo de pantalla.

Estos datos se envían al *backend* alojado en AWS mediante peticiones HTTPS.

La comunicación es segura, garantizando la privacidad e integridad de los datos transmitidos.

### 2. *Backend* (AWS Node.js) → Base de Datos (Azure MySQL)

**Protocolo:** TCP/IP (MySQL sobre TLS)

### **Descripción:**

El *backend* recibe los datos de la aplicación Android y los almacena en una base de datos MySQL alojada en Azure.

La conexión entre el *backend* y la base de datos se establece a través de TCP/IP, con cifrado TLS para garantizar la seguridad.

### **3. Power BI → Base de Datos (Azure MySQL)**

**Protocolo:** TCP/IP (MySQL sobre TLS)

### **Descripción:**

Power BI se conecta directamente a la base de datos en Azure para extraer los datos almacenados y generar reportes analíticos.

La conexión se realiza mediante consultas SQL sobre MySQL, asegurando que la información esté siempre actualizada.

Power BI opera en *Import* (carga de datos previa).

## **8.6. Modelo de gestión y ciclo de vida**

El desarrollo de este proyecto siguió un modelo en cascada, basado en un ciclo de vida de desarrollo tradicional, donde cada fase se completa antes de pasar a la siguiente. Este enfoque permitió estructurar el proceso de manera clara y ordenada, asegurando que cada etapa estuviera bien definida antes de avanzar.

Fases del ciclo de vida

Requisitos

Se identificaron las necesidades del sistema, incluyendo la captura de datos desde la aplicación móvil, su almacenamiento en la base de datos y la generación de reportes.

Se definieron los requerimientos funcionales y no funcionales, asegurando una visión clara del proyecto antes de iniciar el desarrollo.

Diseño

Se diseñó la arquitectura del sistema, definiendo la interacción entre la aplicación Android, el *backend* en AWS, la base de datos en Azure y Power BI.

Se seleccionaron las tecnologías adecuadas: Kotlin para la aplicación móvil, Node.js para el *backend*, MySQL para la base de datos y Power BI para los reportes.

Se elaboraron diagramas para visualizar la estructura y el flujo de datos del sistema.

### Implementación

Se desarrolló la aplicación Android en Kotlin, asegurando la captura de datos de sensores y su envío al *backend*.

Se programó el *backend* en Node.js, integrándolo con la base de datos MySQL en Azure.

Se configuró la base de datos para almacenar los registros de forma segura y estructurada.

Se estableció la conexión con Power BI para la generación de reportes basados en los datos recolectados.

### Pruebas

En lugar de realizar pruebas unitarias formales, el equipo verificó manualmente el funcionamiento del sistema.

Se comprobó que la aplicación móvil enviara correctamente los datos de los sensores al *backend*.

Se revisó que el *backend* recibiera, procesara y almacenara la información en la base de datos.

Se validó que los datos estuvieran accesibles desde Power BI y que los reportes se generaran correctamente.

Se ajustaron configuraciones y se corrigieron errores detectados durante estas pruebas.

### Despliegue

La aplicación fue desplegada en un entorno de prueba para evaluar su funcionamiento en condiciones reales.

Se configuraron los servidores en AWS y la base de datos en Azure, asegurando la estabilidad del sistema.

### Mantenimiento y ajustes

Se realizaron mejoras y ajustes en base a los resultados de las pruebas y la retroalimentación del equipo.

Se optimizó el rendimiento de la aplicación y la base de datos para asegurar un correcto funcionamiento.

El enfoque en cascada permitió desarrollar el proyecto de manera ordenada y estructurada, garantizando que cada fase estuviera bien definida antes de pasar a la siguiente. Aunque este modelo no es tan flexible como otros enfoques ágiles, resultó adecuado para un proyecto con un alcance claro y definido, donde los requisitos y objetivos no cambiaron significativamente a lo largo del desarrollo.

## 8.7. Tecnologías utilizadas

Utilizamos una combinación de tecnologías específicas para el desarrollo de cada uno de sus componentes. A continuación, presentamos una comparativa con otras opciones disponibles, justificando las elecciones realizadas.

### Aplicación Móvil: Kotlin vs otras tecnologías

Tecnología	Ventajas	Desventajas
Kotlin (Usada en el proyecto)	Lenguaje oficial para Android, más conciso que Java, mejor manejo de <i>null safety</i> , interoperable con Java.	Solo funciona en Android, curva de aprendizaje inicial para desarrolladores acostumbrados a Java.
Java	Amplio soporte, madurez, gran comunidad y documentación.	Código más verboso y menos moderno en comparación con Kotlin.
Flutter (Dart)	Permite desarrollo multiplataforma (Android e iOS), UI declarativa potente, buen rendimiento.	No es nativo, mayor consumo de recursos, menos soporte específico para Android que Kotlin.
React Native	Permite desarrollo multiplataforma con JavaScript, buena comunidad.	Peor rendimiento que Kotlin en aplicaciones con alto uso de sensores.

Tabla 1 Comparación entre Kotlin y otras tecnologías

Se eligió Kotlin porque es el lenguaje recomendado por Google para Android y permite un desarrollo más rápido y seguro en comparación con Java. Otras opciones como Flutter o React Native fueron descartadas porque el objetivo del proyecto no requería compatibilidad con iOS ni un enfoque multiplataforma.

### Backend: Node.js vs otras tecnologías

Tecnología	Ventajas	Desventajas
Node.js (Usado en el proyecto)	Basado en JavaScript, alta escalabilidad, buen manejo de concurrencia con eventos asíncronos.	No es ideal para tareas intensivas en CPU, requiere configuración avanzada para alto rendimiento.
Python (Django/Flask)	Código más legible, buena integración con AI/ML y análisis de datos.	Menor rendimiento en aplicaciones de alto tráfico.
Java (Spring Boot)	Alto rendimiento, robustez, buena opción para sistemas empresariales.	Mayor consumo de memoria, más complejo de desarrollar en comparación con Node.js.
Go (Golang)	Excelente rendimiento y concurrencia, bajo consumo de recursos.	Menos comunidad y librerías disponibles en comparación con Node.js.

Tabla 2 Comparación entre Node.js y otras tecnologías

Se seleccionó Node.js porque es el lenguaje que más tenemos conocimiento, además es rápido, escalable y eficiente en la gestión de solicitudes concurrentes, además de ser fácil de integrar con otros servicios en la nube. Alternativas como Spring Boot o Django habrían sido más robustas, pero Node.js ofreció mayor velocidad de desarrollo y flexibilidad para este tipo de aplicación.

### Base de datos: MySQL vs otras tecnologías

Tecnología	Ventajas	Desventajas
MySQL (Usado en el proyecto)	Amplio soporte, buen rendimiento en consultas estructuradas, estabilidad.	Menos eficiente en datos no estructurados, requiere optimización para grandes volúmenes de datos.

PostgreSQL	Más potente en consultas complejas, soporte para JSON y datos espaciales.	Puede ser más lento que MySQL en consultas simples.
MongoDB (NoSQL)	Manejo eficiente de datos no estructurados y escalabilidad horizontal.	No es ideal para datos altamente relacionales como los de este proyecto.
Firebase Firestore	Fácil integración con aplicaciones móviles, sin necesidad de <i>backend</i> propio.	No es adecuado para grandes volúmenes de datos con estructuras relacionales.

Tabla 3 Comparación entre MySQL y otras tecnologías

Se eligió MySQL porque dominamos más las bases de datos relacionales y además los datos recolectados tienen una estructura bien definida, lo que facilita su gestión con bases de datos relacionales. Alternativas como MongoDB o Firebase Firestore fueron descartadas porque este proyecto no requiere una base NoSQL, y PostgreSQL habría sido más útil en escenarios con consultas analíticas complejas, lo cual no era un requerimiento clave en esta fase.

### Plataforma de visualización: Power BI vs Otras herramientas

Tecnología	Ventajas	Desventajas
Power BI (Usado en el proyecto)	Integración con Azure, buena interfaz de usuario, potente en análisis de datos.	Licenciamiento costoso en versiones avanzadas, curva de aprendizaje.
Tableau	Visualizaciones más personalizables, mejor manejo de Big Data.	Más costoso que Power BI, curva de aprendizaje mayor.
Google Data Studio	Gratuito, integración con productos de Google.	Menos potente en análisis avanzado y procesamiento de grandes volúmenes de datos.
Metabase	Open-source, fácil de usar, bueno para reportes simples.	Menos funcionalidades avanzadas que Power BI.

Tabla 4 Comparación entre Power BI y otras tecnologías

Se optó por Power BI debido a su fuerte integración con Azure, facilidad de uso y potencia en la generación de reportes. Alternativas como Tableau habrían ofrecido mayor personalización, pero con un costo más alto y sin ventajas significativas en este caso.

## 8.8. Decisiones clave

A la hora de diseñar este prototipo de proyecto de fenotipado digital, se decidió utilizar una combinación de Node.js, TypeScript, el patrón *Pipe and Filters* y una base de datos relacional, en vista de las necesidades de escalabilidad, mantenibilidad y confiabilidad que exige el manejo continuo de datos de sensores de múltiples dispositivos móviles.

En primer lugar, Node.js ofrece un modelo de E/S asíncrono que resulta esencial cuando se reciben datos de varios *smartphones* en tiempo real. Gracias a su *event loop* no bloqueante, se pueden procesar grandes volúmenes de información sin que el rendimiento del sistema se vea afectado. Además, Node.js cuenta con un ecosistema amplio y maduro que proporciona librerías y herramientas consolidadas para construir servicios web, lo cual facilita la implementación de APIs y la integración con diferentes plataformas.

Por otro lado, el uso de TypeScript aporta robustez y mantenibilidad al proyecto. Al ser un *superset* tipado de JavaScript, TypeScript permite detectar inconsistencias en tiempo de compilación y proporciona una documentación implícita a través de las definiciones de tipos. Esto es especialmente valioso en un dominio como el fenotipado digital, donde cada sensor puede generar datos muy distintos (como coordenadas de geolocalización, valores de iluminación o acciones de pantalla), lo que hace necesaria una mayor confiabilidad en la manipulación de la información.

El patrón *Pipe and Filters* se integra de forma natural en este contexto porque separa el procesamiento de la información en pasos o “filtros” que van encadenados: validar la existencia de un sensor, comprobar la coherencia de los valores que llegan y, finalmente, persistirlos. Esta división en etapas bien definidas favorece la extensibilidad (se pueden añadir nuevos filtros con facilidad) y la mantenibilidad (cada filtro tiene una responsabilidad clara y enfocada). A medida que el proyecto crezca o se requieran nuevas validaciones, bastará con insertar o modificar un filtro en la cadena sin alterar la lógica completa.

Por último, se eligió una base de datos relacional para garantizar la consistencia y fiabilidad de la información almacenada. Los datos de múltiples sensores, usuarios y dispositivos pueden

requerir relaciones complejas y, gracias a la integridad referencial y las transacciones ACID que ofrecen las bases de datos relacionales, se asegura que la información permanezca coherente incluso ante errores o altas cargas de trabajo. Además, la posibilidad de realizar consultas SQL avanzadas (*joins*, agregaciones, filtros) facilita el análisis posterior, clave para extraer correlaciones en un proyecto de fenotipado digital.

En conjunto, esta arquitectura (Node.js y TypeScript para la ejecución escalable y robusta, el patrón *Pipe and Filters* para la organización modular del flujo de datos, y una base de datos relacional para la persistencia confiable) proporciona una base sólida y flexible para manejar, validar y almacenar eficientemente los datos de sensores a lo largo de todo el proceso de fenotipado digital.

## 8.9. Análisis de riesgos

### Riesgos técnicos

**Riesgo:** Fallos en la recolección de datos debido a limitaciones del hardware.

**Impacto:** Los datos recolectados podrían ser incompletos o no representativos.

**Mitigación:** Implementar verificaciones de estado para asegurar que los sensores están funcionando correctamente.

**Riesgo:** Problemas de compatibilidad con versiones antiguas de Android.

**Impacto:** Usuarios con dispositivos desactualizados no podrían utilizar la aplicación.

**Mitigación:** Definir un mínimo de compatibilidad claro y optimizar para las versiones soportadas.

### Riesgos éticos

**Riesgo:** Uso indebido de los datos recolectados en aplicaciones futuras.

**Impacto:** Podría dañar la confianza de los usuarios y llevar a implicaciones legales.

**Mitigación:** Asegurar que los datos sean anonimizados y limitar el acceso a ellos.

**Riesgo:** Falta de consentimiento informado por parte de los usuarios.

**Impacto:** Recolección de datos no ética que podría afectar la reputación del proyecto.

**Mitigación:** Diseñar una interfaz que explique claramente qué datos se recolectan y obtener consentimiento explícito.

### **Riesgos operativos**

**Riesgo:** Pérdida de datos debido a fallos en la base de datos o infraestructura.

**Impacto:** La información recolectada podría no estar disponible para su análisis.

**Mitigación:** Configurar copias de seguridad automáticas y redundancia en la infraestructura.

**Riesgo:** Retrasos en la sincronización de datos entre la aplicación y el *backend*.

**Impacto:** Los datos podrían llegar de forma tardía, afectando el análisis en tiempo real.

**Mitigación:** Optimizar el protocolo de envío de datos para asegurar una sincronización eficiente.

## **8.10. Ilustraciones del prototipo y reporte**

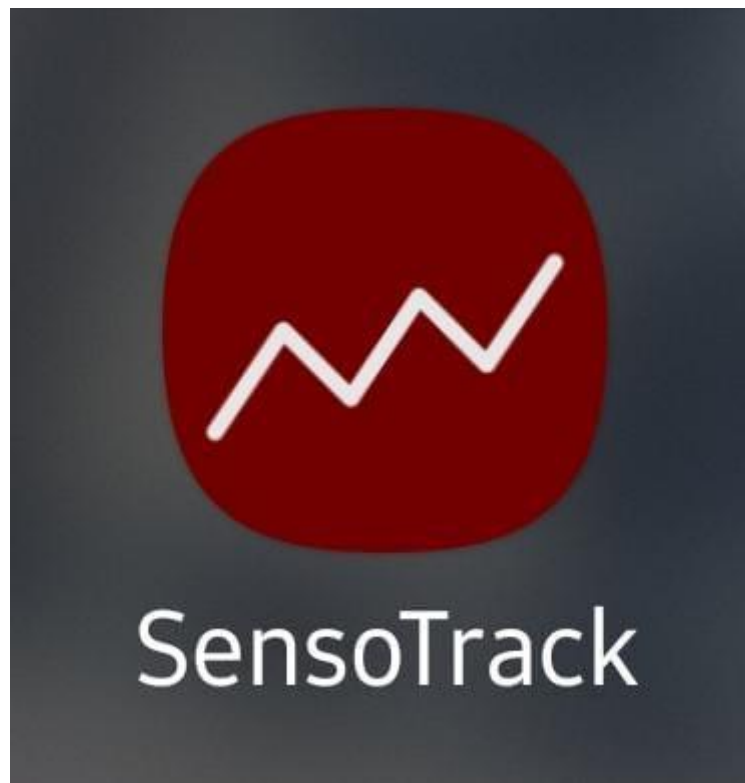


Ilustración 2 Icono de SensoTrack



Ilustración 3 Pantalla principal de SensoTrack

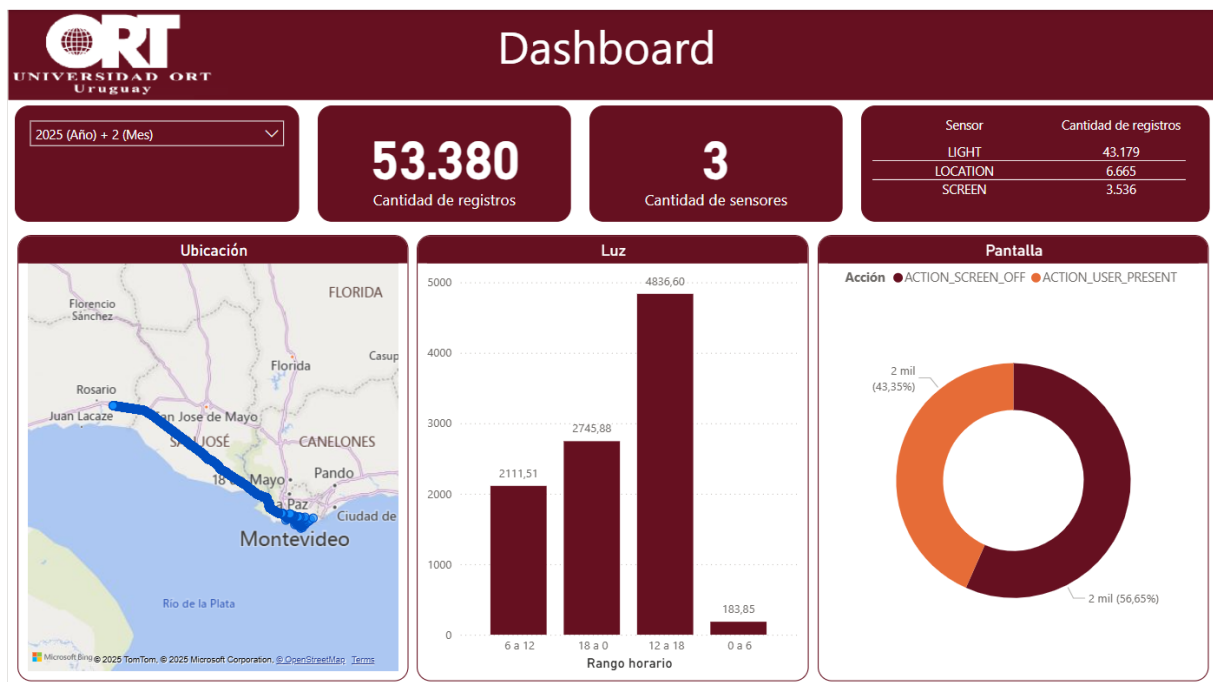


Ilustración 4 Hoja principal del reporte

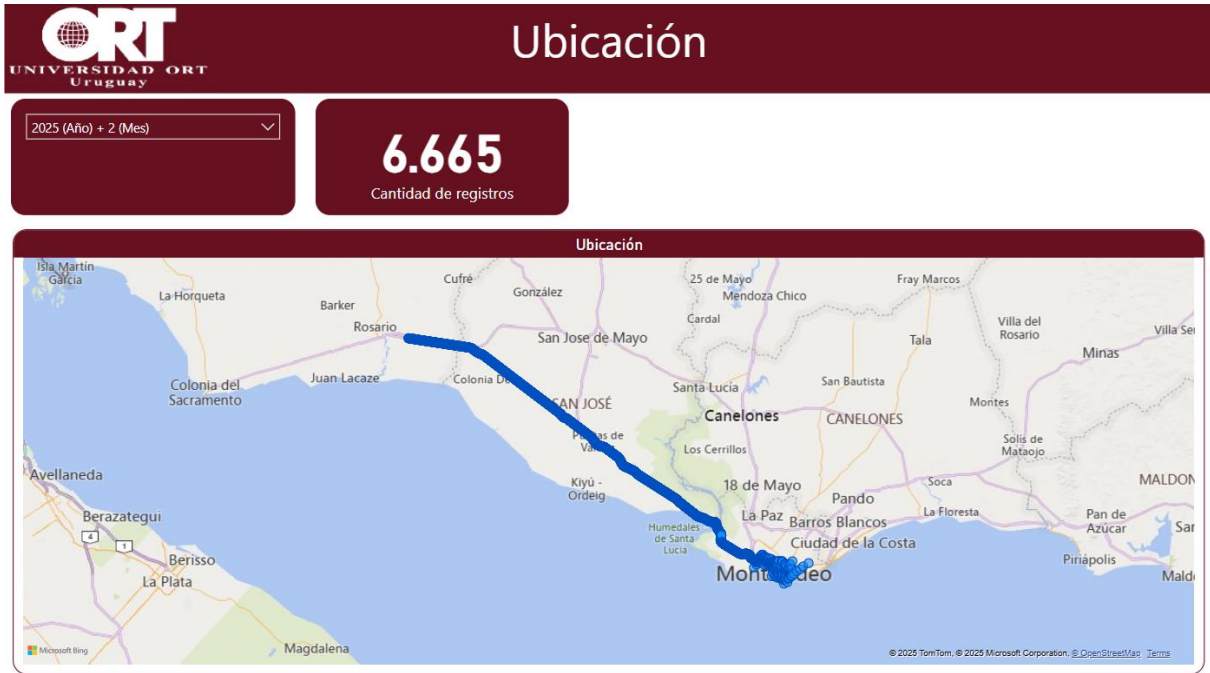


Ilustración 5 Hoja del sensor de ubicación del reporte

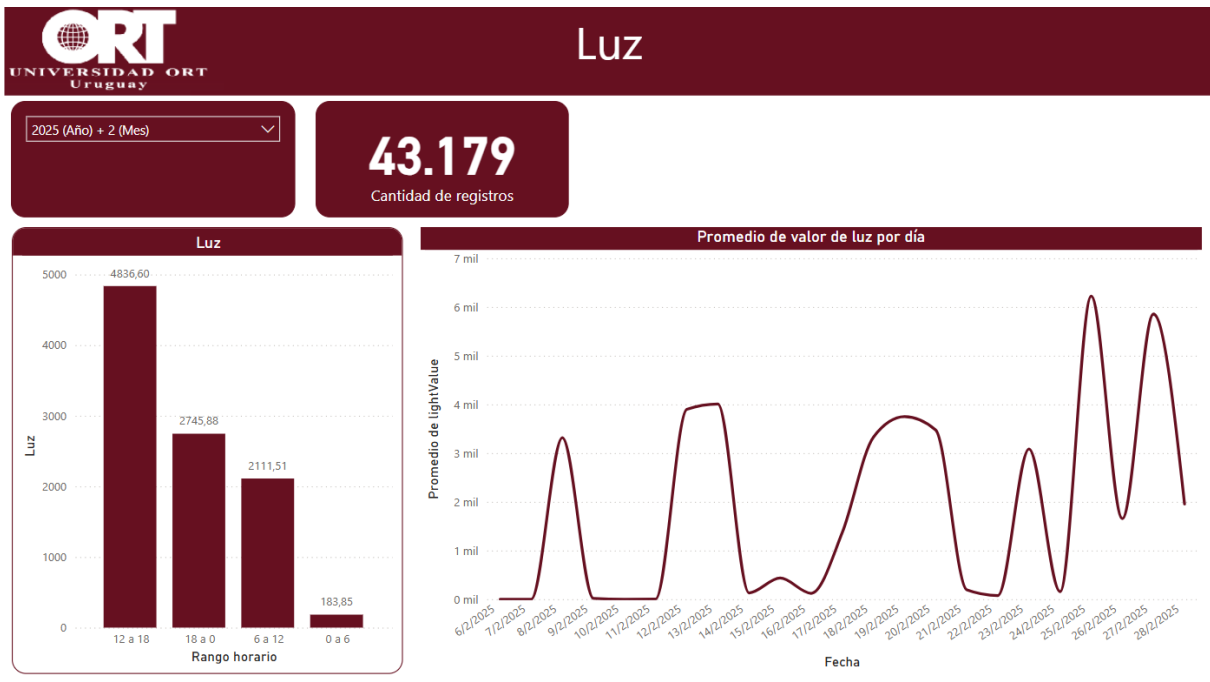


Ilustración 6 Hoja del sensor de luz del reporte

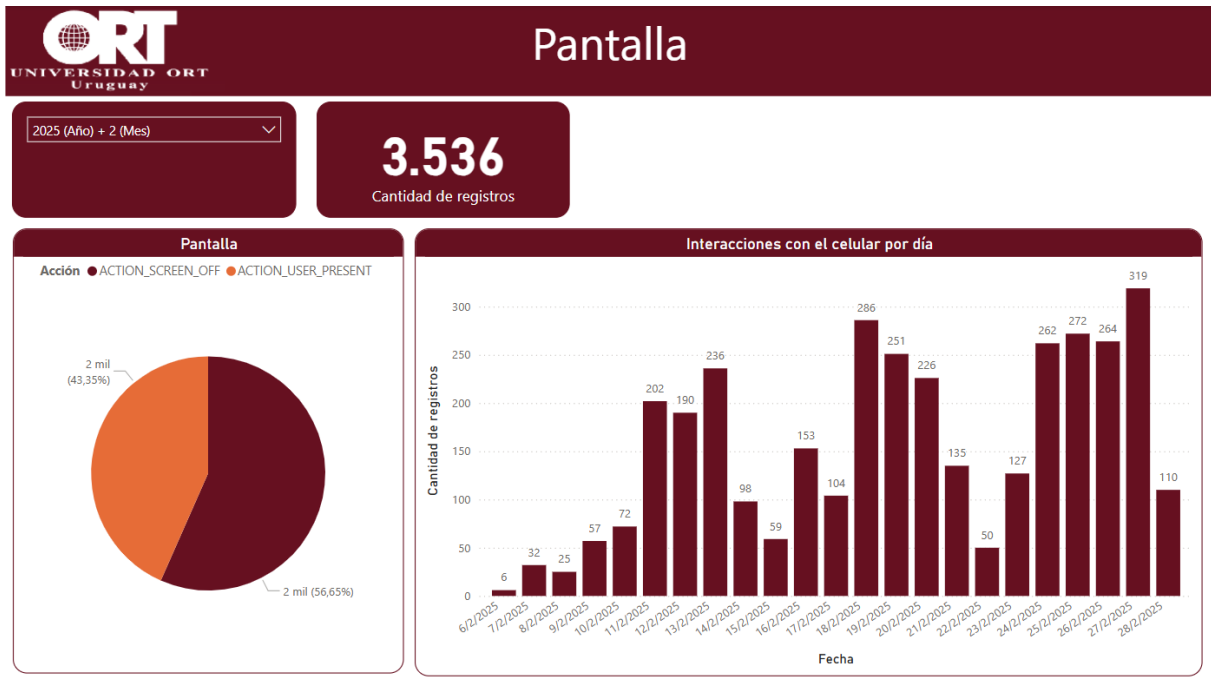


Ilustración 7 Hoja del sensor de bloqueo del reporte

**Datos**

2025 (Año) + 2 (Mes)

LIGHT  
 LOCATION  
 SCREEN

Sensor	Acción	Fecha y hora	Longitud	Latitud	Luz
LIGHT		14/02/2025 07:46:54 p.m.	0,00	0,00	13
LIGHT		14/02/2025 07:50:34 p.m.	0,00	0,00	2
LIGHT		14/02/2025 07:51:21 p.m.	0,00	0,00	234
LIGHT		14/02/2025 07:51:23 p.m.	0,00	0,00	17
LIGHT		14/02/2025 07:51:36 p.m.	0,00	0,00	256
LIGHT		14/02/2025 07:51:40 p.m.	0,00	0,00	13
LIGHT		14/02/2025 08:26:06 p.m.	0,00	0,00	228
LIGHT		14/02/2025 08:26:07 p.m.	0,00	0,00	18
LIGHT		14/02/2025 08:26:22 p.m.	0,00	0,00	11
LIGHT		14/02/2025 08:26:22 p.m.	0,00	0,00	251
LIGHT		14/02/2025 08:26:31 p.m.	0,00	0,00	212
LIGHT		14/02/2025 08:26:32 p.m.	0,00	0,00	11
LIGHT		14/02/2025 08:27:30 p.m.	0,00	0,00	223
LIGHT		14/02/2025 08:27:32 p.m.	0,00	0,00	8
LIGHT		14/02/2025 08:27:41 p.m.	0,00	0,00	249
LIGHT		14/02/2025 08:27:42 p.m.	0,00	0,00	24
LIGHT		14/02/2025 08:28:18 p.m.	0,00	0,00	20
LIGHT		14/02/2025 08:28:18 p.m.	0,00	0,00	245
LIGHT		14/02/2025 08:28:19 p.m.	0,00	0,00	8
LIGHT		14/02/2025 08:28:19 p.m.	0,00	0,00	267
LIGHT		14/02/2025 08:28:32 p.m.	0,00	0,00	228
LIGHT		14/02/2025 08:28:33 p.m.	0,00	0,00	8
LIGHT		14/02/2025 08:28:42 p.m.	0,00	0,00	256
LIGHT		14/02/2025 08:28:43 p.m.	0,00	0,00	49
LIGHT		14/02/2025 08:28:57 p.m.	0,00	0,00	252
LIGHT		14/02/2025 08:28:59 p.m.	0,00	0,00	7
LIGHT		14/02/2025 08:29:10 p.m.	0,00	0,00	270
LIGHT		14/02/2025 08:29:11 p.m.	0,00	0,00	13

Ilustración 8 Hoja de datos del reporte

## 8.11. Conclusiones

El desarrollo de este proyecto resultó una experiencia muy buena, no solo por haber logrado cumplir los objetivos planteados, sino también por la oportunidad de trabajar con tecnologías y servicios en la nube que no habíamos utilizado previamente. A lo largo del proceso, adquirimos nuevos conocimientos sobre la captura y transmisión de datos desde dispositivos móviles, la integración de múltiples plataformas en la nube y la visualización de datos en reportes dinámicos.

Uno de los logros más importantes del proyecto fue validar que es posible capturar datos de al menos tres sensores de un dispositivo Android (sensor de luz, ubicación geográfica y eventos de bloqueo/desbloqueo de pantalla), transmitirlos a un *backend*, almacenarlos en una base de datos y visualizarlos en Power BI. Este flujo de datos se implementó de manera efectiva, asegurando que la información se mantuviera disponible y estructurada para su análisis posterior.

Además, el proyecto nos permitió explorar el desarrollo e integración de servicios en las nubes más utilizadas del mercado, como AWS y Azure. Aprendimos cómo desplegar un *backend* en AWS utilizando Node.js, cómo gestionar bases de datos en Azure con MySQL, y cómo conectar estos servicios de manera segura y eficiente. También fue interesante comprobar cómo Power BI se integra con Azure para la generación de reportes, permitiendo visualizar y analizar los datos de manera intuitiva.

Otro aspecto positivo del proyecto fue la posibilidad de trabajar con un modelo de arquitectura distribuida, donde cada componente del sistema tenía una función específica y bien definida. Esto nos dio una mejor comprensión sobre cómo los diferentes servicios interactúan entre sí y qué consideraciones de seguridad y conectividad son necesarias al trabajar con múltiples entornos en la nube.

Si bien el desarrollo fue exitoso, también identificamos áreas de mejora, como la necesidad de incorporar pruebas automatizadas, optimizar la transmisión de datos para mejorar la eficiencia y documentar mejor algunos procesos técnicos. Estos aprendizajes serán valiosos para futuros proyectos en los que trabajemos con arquitecturas similares.

Finalmente, este proyecto no solo permitió validar la factibilidad de capturar y procesar datos desde un dispositivo móvil, sino que también sirvió como una excelente oportunidad para

experimentar con tecnologías modernas y fortalecer conocimientos sobre la implementación de servicios en la nube, bases de datos escalables y herramientas de análisis de datos. Fue una experiencia interesante y desafiante, que nos dejó mejor preparados para abordar desarrollos más complejos en el futuro.

Repositorio GIT: <https://github.com/jorgekino/SensoTrack.git>

## 9. Conclusiones generales

El fenotipado digital ha surgido como una herramienta clave para la investigación y monitoreo en diversas áreas, particularmente en la salud mental y el bienestar de los individuos. Su capacidad para recopilar y analizar grandes volúmenes de datos en tiempo real ofrece oportunidades para mejorar el diagnóstico, tratamiento y seguimiento de condiciones psicológicas y médicas. A través del uso de sensores móviles y técnicas avanzadas de análisis de datos, el fenotipado digital permite obtener información precisa y continua sobre el comportamiento, la movilidad y otros indicadores relevantes que antes solo podían ser evaluados mediante métodos tradicionales, como encuestas o entrevistas clínicas.

En este trabajo, logramos con éxito la recolección de datos de sensores clave en dispositivos móviles, específicamente GPS, luz ambiental y bloqueo/desbloqueo de pantalla. Estos datos proporcionan información valiosa sobre la movilidad del usuario, su exposición a la luz y los patrones de uso del dispositivo, lo que puede contribuir significativamente a estudios sobre hábitos de sueño, actividad diaria y bienestar general. La integración de estos sensores demuestra el potencial del fenotipado digital para capturar información en tiempo real y ofrecer análisis más detallados sobre el comportamiento humano.

Dado que el éxito de cualquier plataforma de fenotipado digital depende en gran medida de la infraestructura tecnológica utilizada, en este proyecto hemos seleccionado herramientas con las cuales tenemos experiencia previa, asegurando así un desarrollo eficiente y confiable. Optamos por tecnologías como Node.js para el *backend* debido a su capacidad para manejar grandes volúmenes de solicitudes de manera eficiente, AWS y Azure por su infraestructura escalable y segura, y MySQL como sistema de base de datos debido a su estabilidad y flexibilidad en la gestión de información estructurada.

En cuanto al desarrollo de la aplicación móvil, elegimos Kotlin como lenguaje de programación por diversas razones estratégicas. Kotlin es el lenguaje oficial para el desarrollo de Android, lo que garantiza compatibilidad y soporte continuo por parte de Google. Además, su sintaxis concisa y segura reduce la probabilidad de errores, lo que se traduce en un código más mantenible y robusto. También ofrece una excelente interoperabilidad con Java, lo que facilita la integración con bibliotecas y *frameworks* ya existentes en el ecosistema de desarrollo móvil.

Para finalizar, el fenotipado digital representa una oportunidad única para avanzar en la comprensión y monitoreo de la salud y el comportamiento humano. Gracias al uso de sensores móviles y técnicas de análisis de datos, es posible recolectar información detallada y continua que antes era inaccesible. Con la implementación de herramientas sólidas y tecnologías bien establecidas, este proyecto ha demostrado que es factible obtener datos clave de los dispositivos móviles para su análisis en el contexto del fenotipado digital. Así, sentamos las bases para el desarrollo de plataformas más avanzadas, escalables y seguras, con el potencial de contribuir significativamente en el campo de la investigación digital y la salud móvil.

## 10. Trabajos futuros

Nuestro trabajo ha logrado con éxito la implementación de un prototipo funcional capaz de recolectar y enviar datos de sensores móviles cada 15 minutos, específicamente de GPS, luz ambiental y bloqueo/desbloqueo de pantalla. Si bien esta etapa ha demostrado la viabilidad técnica del fenotipado digital mediante dispositivos móviles, aún existen múltiples áreas por explorar para profundizar en el análisis e interpretación de los datos obtenidos.

Uno de los principales enfoques para trabajos futuros es el procesamiento y análisis avanzado de los datos recopilados. Actualmente, contamos con información en bruto sobre la ubicación del usuario, las condiciones de iluminación en su entorno y los patrones de uso del dispositivo. No obstante, es necesario aplicar técnicas de procesamiento de datos, estadística avanzada y *machine learning* para extraer patrones, correlaciones y conclusiones significativas.

Algunas líneas de investigación incluyen:

- Patrones de movilidad y rutinas diarias: A través del análisis de los datos de GPS, es posible identificar hábitos de desplazamiento, evaluar niveles de actividad y detectar cambios en la rutina del usuario. Esto podría ser útil para estudios sobre movilidad urbana, monitoreo de bienestar y detección de signos de aislamiento social o disminución de actividad física.
- Impacto de la luz ambiental en el bienestar: La exposición a diferentes niveles de luz natural y artificial está relacionada con la calidad del sueño, los ritmos circadianos y el estado de ánimo. Un análisis futuro podría evaluar correlaciones entre los datos del sensor de luz ambiental y factores como la productividad, la fatiga y la regulación emocional.
- Hábitos digitales y bienestar: El monitoreo del bloqueo y desbloqueo de pantalla permite identificar patrones de uso del dispositivo, como frecuencia y duración de interacciones. Esto podría ayudar a evaluar el impacto del uso del celular en la concentración, la calidad del sueño y la regulación del estrés.

Otro punto para desarrollar es la implementación de modelos predictivos que permitan analizar tendencias en los datos recopilados y realizar predicciones sobre el comportamiento y estado del usuario. Aplicaciones de *machine learning* podrían detectar patrones anómalos que actúen como señales de alerta para el monitoreo de la salud digital.

En conclusión, nuestro proyecto ha demostrado la viabilidad del fenotipado digital mediante sensores móviles, pero los próximos pasos deben enfocarse en la interpretación de los datos obtenidos, su validación y la exploración de nuevas aplicaciones. Con un análisis riguroso y herramientas avanzadas, el sistema podría evolucionar hacia una solución más completa que contribuya al bienestar y la investigación en salud digital.

## 11. Lecciones aprendidas

A lo largo del desarrollo de nuestro trabajo, hemos adquirido una serie de habilidades, conocimientos y experiencias de gran valor tanto en el ámbito técnico como en el metodológico. El trabajo realizado no solo nos permitió profundizar en el fenotipado digital y su aplicación en dispositivos móviles, sino que también nos brindó la oportunidad de enfrentar y superar diversos desafíos.

Uno de los aprendizajes más significativos fue el desarrollo de una aplicación móvil funcional. Antes de este proyecto, no contábamos con mucha experiencia en desarrollo móvil, por lo que fue necesario aprender a utilizar Kotlin y entender el ecosistema de desarrollo de Android. Esto nos permitió conocer buenas prácticas de programación para móviles, optimización del consumo de batería y gestión de permisos de sensores en sistemas operativos modernos.

Además, el proyecto requirió una revisión exhaustiva de la literatura sobre fenotipado digital, privacidad, seguridad de datos y procesamiento de información sensorial. Esto nos permitió familiarizarnos con un campo de investigación en crecimiento y entender sus implicaciones tanto a nivel técnico como ético y legal. Aprendimos a evaluar fuentes científicas, sintetizar información relevante y construir un marco teórico sólido que respaldara la parte práctica del trabajo.

Otro aspecto clave fue la variedad interdisciplinaria del proyecto. Combinar conocimientos de desarrollo de software, análisis de datos, ciencias de la salud y privacidad digital nos hizo valorar la importancia de integrar diferentes enfoques para abordar problemas complejos. La implementación en AWS, Azure y MySQL también nos permitió fortalecer nuestras habilidades en infraestructura en la nube y bases de datos, asegurando un sistema escalable y seguro.

Desde el punto de vista organizacional, el proyecto nos enseñó a trabajar en equipo de manera efectiva. Coordinar tareas, definir responsabilidades y tomar decisiones conjuntas fue fundamental para avanzar en cada etapa del desarrollo. Además, aprendimos a gestionar imprevistos, como dificultades técnicas o cambios en los requerimientos, manteniendo siempre un enfoque flexible y orientado a soluciones.

Finalmente, este trabajo nos permitió adentrarnos en un tema completamente nuevo y comprender su potencial para futuras aplicaciones. La posibilidad de recolectar datos de

sensores móviles abre distintas oportunidades en investigación y desarrollo, desde la salud digital hasta el análisis de comportamiento humano. Nuestro trabajo sentó las bases para continuar explorando el fenotipado digital, con la posibilidad de profundizar en el análisis de datos y expandir su funcionalidad en el futuro.

En conclusión, más allá del producto final obtenido, este proyecto nos dejó un valioso aprendizaje en desarrollo móvil, investigación, trabajo en equipo y análisis de datos, lo que nos permitirá abordar futuros desafíos con una base más sólida y una visión más amplia.

## 12. Referencias bibliográficas

- [1] C. Montag y H. Baumeister, Eds., *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics*. en *Studies in Neuroscience, Psychology and Behavioral Economics*. Cham: Springer International Publishing, 2023. doi: 10.1007/978-3-030-98546-2.
- [2] L. Feng, «Preface», en *Context-Aware Computing*, De Gruyter, 2017, p. VII-VIII. doi: 10.1515/9783110556674-202.
- [3] M. R. Mehl, Ed., *Mobile sensing in psychology: methods and applications*. New York ; London: The Guilford Press, 2024.
- [4] F. Beierle, *Integrating Psychoinformatics with Ubiquitous Social Networking: Advanced Mobile-Sensing Concepts and Applications*. en *T-Labs Series in Telecommunication Services*. Cham: Springer International Publishing, 2021. doi: 10.1007/978-3-030-68840-0.
- [5] K. S. Kaswan, J. S. Dhatteerwal, y A. Nayyar, *Digital Personality: A Man Forever: Volume 1: Introduction*, 1.<sup>a</sup> ed. Boca Raton: CRC Press, 2024. doi: 10.1201/9781032628431.
- [6] I. C. Passos, F. D. Rabelo-da-Ponte, y F. Kapczynski, Eds., *Digital Mental Health: A Practitioner's Guide*. Cham: Springer International Publishing, 2023. doi: 10.1007/978-3-031-10698-9.
- [7] I. Perez-Pozuelo, «Digital phenotyping through multimodal, unobtrusive sensing».
- [8] R. Jagesar, «Digital Phenotyping: Operation and Execution», University of Groningen, 2022. doi: 10.33612/diss.218110466.
- [9] N. Jongs, «Passive Digital Phenotyping: objective quantification of human behaviour through smartphones», University of Groningen, 2021. doi: 10.33612/diss.171368248.
- [10] E. Sükei, «A Step Towards Advancing Digital Phenotyping In Mental Healthcare».
- [11] H. Baumeister y C. Montag, Eds., *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics*. en *Studies in Neuroscience, Psychology and Behavioral Economics*. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-31620-4.
- [12] R. Maatoug *et al.*, «Digital phenotype of mood disorders: A conceptual and critical review», *Front. Psychiatry*, vol. 13, p. 895860, jul. 2022, doi: 10.3389/fpsy.2022.895860.
- [13] S. Coghlan y S. D'Alfonso, «Digital Phenotyping: an Epistemic and Methodological Analysis», *Philos. Technol.*, vol. 34, n.º 4, pp. 1905-1928, dic. 2021, doi: 10.1007/s13347-021-00492-1.
- [14] D. Marengo, J. D. Elhai, y C. Montag, «Predicting Big Five personality traits from smartphone data: A meta-analysis on the potential of digital phenotyping», *J. Pers.*, vol. 91, n.º 6, pp. 1410-1424, dic. 2023, doi: 10.1111/jopy.12817.
- [15] S. Mulinari, «Short-circuiting biology: Digital phenotypes, digital biomarkers, and shifting gazes in psychiatry», *Big Data Soc.*, vol. 10, n.º 1, p. 20539517221145680, ene. 2023, doi: 10.1177/20539517221145680.
- [16] A. Oudin *et al.*, «Digital Phenotyping: Data-Driven Psychiatry to Redefine Mental Health», *J. Med. Internet Res.*, vol. 25, p. e44502, oct. 2023, doi: 10.2196/44502.
- [17] K. Huckvale, S. Venkatesh, y H. Christensen, «Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety», *Npj Digit. Med.*, vol. 2, n.º 1, p. 88, sep. 2019, doi: 10.1038/s41746-019-0166-1.
- [18] J. M. O. Ceja, A. G. Arenas, C. D. G. Romero, S. R. Rodríguez, y G. L. M. Luna, «Visual Characterization of Gathered Data for Digital Phenotyping», en *Information Technology and Systems*, vol. 932, Á. Rocha, C. Ferrás, J. Hochstetter Diez, y M. Diéguez Rebolledo,

Eds., en *Lecture Notes in Networks and Systems*, vol. 932. , Cham: Springer Nature Switzerland, 2024, pp. 203-211. doi: 10.1007/978-3-031-54235-0\_19.

- [19] S. F. M. Pizzoli, D. Monzani, L. Conti, G. Ferraris, R. Grasso, y G. Pravettoni, «Issues and opportunities of digital phenotyping: ecological momentary assessment and behavioral sensing in protecting the young from suicide», *Front. Psychol.*, vol. 14, p. 1103703, jun. 2023, doi: 10.3389/fpsyg.2023.1103703.
- [20] R. Potier, «The Digital Phenotyping Project: A Psychoanalytical and Network Theory Perspective», *Front. Psychol.*, vol. 11, p. 1218, jul. 2020, doi: 10.3389/fpsyg.2020.01218.
- [21] O. Yurur y C. H. Liu, *Generic and Energy-Efficient Context-Aware Mobile Sensing*, 0 ed. CRC Press, 2015. doi: 10.1201/b18058.