

# Arquitectura conceptual de plataforma tecnológica de vigilancia epidemiológica para la COVID-19

Conceptual architecture of the epidemiological surveillance technology platform for COVID-19

Pedro Atencio<sup>1</sup>, German Sánchez-Torres<sup>2</sup>, Rene Iral Palomino<sup>3</sup>, John W. Branch Bedoya<sup>3</sup>, Daniel Burgos<sup>4</sup>

<sup>1</sup> Instituto Tecnológico Metropolitano, Colombia

<sup>2</sup> Universidad del Magdalena, Colombia

<sup>3</sup> Universidad Nacional de Colombia, Sede Medellín, Colombia

<sup>4</sup> Universidad Internacional de La Rioja (UNIR), España

pedroatencio@itm.edu.co , gsanchez@unimagdalena.edu.co , riral@unal.edu.co ,  
jwbranch@unal.edu.co , daniel.burgos@unir.net

**RESUMEN.** Dado que resulta probable que el SARS-CoV-2 se vuelva endémico en muchos países, requerirá no sólo apoyo a corto plazo sino también a largo plazo, ya que las políticas de distanciamiento social no pueden extenderse por mucho tiempo. Por lo tanto, una plataforma tecnológica de vigilancia epidemiológica puede representar una herramienta fundamental. El impacto del proyecto resulta esencial para que los actores relacionados con la salud pública diseñen y evalúen políticas destinadas a la reactivación segura de las actividades sociales después de que se suspendan las políticas de distanciamiento social. Consideramos también este servicio software como una pieza básica en la estrategia de Transformación Digital, ya que permite anticipar comportamientos y recursos necesarios que amolden las necesidades con la provisión de manera dinámica, pero ajustada a la realidad. Este enfoque de anticipación se vuelve un pilar en la estrategia digital de cualquier empresa, Administración y centro de educación. La herramienta incluye un mecanismo basado en Inteligencia Artificial para el análisis de datos con el fin de tener una comprensión dinámica de los síntomas, la evolución, los datos espacio-temporales sociales y las relaciones entre ellos, lo que permitirá a las entidades relevantes optimizar recursos como las pruebas de detección de virus y controles de prueba positivo.

**ABSTRACT.** Since SARS-CoV-2 is likely to become endemic in many countries, it will require not only short-term support but also long-term support, as social distancing policies cannot be extended for long. Therefore, a technological platform for epidemiological surveillance can represent a fundamental tool. The impact of the project is essential for public health actors to design and evaluate policies aimed at the safe reactivation of social activities after social distancing policies are suspended. We also consider this software service as a basic piece in the Digital Transformation strategy, since it allows us to anticipate the behaviors and necessary resources that adapt the needs with the provision in a dynamic way, but adjusted to reality. This anticipation approach becomes a pillar in the digital strategy of any company, Administration and education center. The tool includes a mechanism based on Artificial Intelligence for data analysis in order to have a dynamic understanding of symptoms, evolution, social space-time data and the relationships between them, which will allow the relevant entities to optimize resources such as virus detection tests and positive test controls.

**PALABRAS CLAVE:** COVID-19, Inteligencia artificial, Análisis de datos, Transformación digital.

**KEYWORDS:** COVID-19, Artificial intelligence, Data analytics, Digital transformation.

## 1. Introducción

La Vigilancia Epidemiológica como componente de la Epidemiología es una de las disciplinas básicas de la Salud Pública y se define como: “la recolección, análisis e interpretación sistemática y constante de datos sobre problemas específicos de salud en poblaciones, su procesamiento y análisis para utilizarlos en la planificación, ejecución y evaluación de intervenciones para la prevención y control de los riesgos o daños a la salud”. (Organización Panamericana de Salud, 1974; Guibovich, 2009).

El análisis permanente y evaluación a través de los datos recopilados (Figura 1), puede permitir conocer el comportamiento de las poblaciones afectadas y la manera en que las acciones de prevención e intervención tienen impacto sobre la mitigación o control del evento. Este impacto estará también asociado a las particularidades de las poblaciones en riesgo y la manera en cómo el evento se presenta, en términos de la dinámica de casos y de muertes, si es el caso.

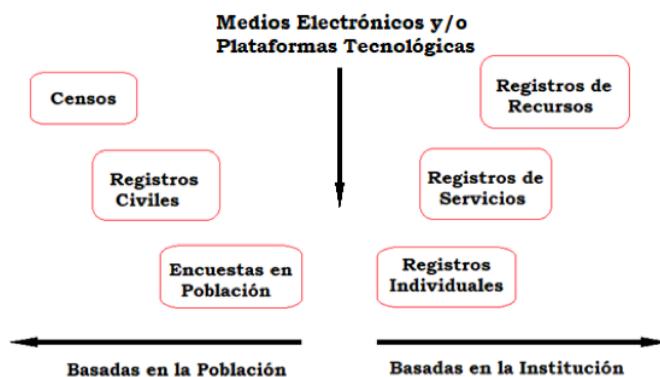


Figura 1. Fuentes de datos. Fuente: (Martínez, 2000; Valenzuela, 2015).

El conocimiento preciso y en tiempo real de ese comportamiento se complementa con la capacidad de anticipación a futuro, lo que conlleva una previsión de recursos, personal y acciones implicadas (Deshpande, Sharma & Peddoju, 2019). Esta anticipación mejora la ratio de adecuación entre necesidades y provisión de las mismas, lo que permite optimizar y reubicar aquellos elementos no necesarios (Kappen et al., 2018). Sin duda, la capacidad de adecuación a escenarios cambiantes, con un análisis de probabilidades y la adaptación de los recursos a dichos escenarios, supone un elemento clave en las estrategias de Transformación Digital de instituciones, empresas, Administraciones y centros educativos (Mergel, Edelman & Haug, 2019; García-Peñalvo & Corell, 2020). En el caso de las Instituciones de Educación Superior (IES) la Transformación Digital se fundamenta, entre otros puntos, en la cultura del dato y en la capacidad de análisis que permite obtener radiografías precisas del momento y, sobre todo, probables escenarios futuros sobre los que actuar con tiempo suficiente (Pappas, 2018; Safiullin & Akhmetshin, 2019; Ramírez-Montoya, 2020; Cabero-Almenara & Llorente-Cejudo, 2020).

En lo que respecta a las fuentes, es importante enfatizar que este es quizás el mayor problema en la recolección de datos, pues debe garantizarse la estandarización de los mismos, sin importar las fuentes que los reporten ya sea de manera obligatoria, o de manera esporádica.

Se hace importante definir qué datos de todos los que se pueden obtener por fuentes como exámenes clínicos o de laboratorios, consultas externas, necropsias, serán tenidos en cuenta en el proceso de vigilancia. En este punto, una plataforma de reporte o un sistema de información, permitiría realizar la respectiva gestión de los datos y generar la información requerida por el sistema de vigilancia. Cabe preguntarse si los sistemas de reporte de eventos de salud con los que podrían contar los países en el mundo, permiten la gestión de los datos allí recopilados. Este punto es clave, ya que, si no hay una correcta estandarización en los reportes y la

manera como se deben recopilar los datos que alimentarán el sistema de vigilancia, serán necesarios procesos de depuración y de calidad de los datos, antes de abordar el entendimiento del evento o los eventos en estudio, lo que al final retrasaría el proceso de toma de decisiones y dificultaría así la medición del impacto de las medidas adoptadas en el marco de la prevención, control y monitoreo.

En relación con los demás procesos de recolección de datos que se hacen de manera rutinaria y que posiblemente no están asociados a un evento de interés particular (encuestas serológicas, entomológicas, morbilidad, tamaños de poblaciones, etc.), es necesario establecer de qué manera se puede acceder a dicha información y asociarla a posibles eventos de salud cuyo monitoreo, prevención y control es parte del proceso de vigilancia epidemiológica. Aquí, nuevamente se debe garantizar la calidad de los datos, su adecuada estandarización en términos de registro y la manera como se podrá acceder a estos, para efecto de realizar las diferentes actividades y acciones del sistema de vigilancia epidemiológica.

En síntesis, aunque se piense en una plataforma tecnológica o en otras herramientas de recolección de datos a través de medios tecnológicos, como herramientas clave que permita soportar procesos de gestión de información, el gran inconveniente que tienen diferentes sistemas de registro de datos en áreas de la salud, es garantizar la calidad de los mismos y el tratamiento de las inconsistencias (procesos de estandarización, validación e imputación, si fuera el caso).

El número exponencialmente creciente de la pandemia del síndrome respiratorio agudo severo de infecciones por SARS-CoV-2 (COVID-19), presenta desafíos urgentes de implementar estrategias de salud pública para dilucidar la epidemiología del nuevo virus y caracterizar su impacto potencial. Comprensión de los factores de riesgo de infección y predictores de los resultados posteriores es crucial para obtener el control de la pandemia SARS-CoV-2 (COVID-19). Sin embargo, la velocidad en el que se desarrolla la pandemia plantea un desafío sin precedentes para la colección de datos. Para caracterizar la amplitud completa de la severidad de la enfermedad, lo que dificulta los esfuerzos en el tiempo de difusión de información precisa, afectando la planificación de la salud pública y la gestión clínica.

Por lo tanto, existe una necesidad urgente de una plataforma adaptable de captura de datos en tiempo real para recoger de forma rápida y prospectiva procesable datos de alta calidad que abarcan el espectro de presentaciones subclínicas y agudas e identificar disparidades en el diagnóstico, tratamiento, y resultados clínicos. Abordar esta prioridad permitirá estimaciones más precisas de la enfermedad incidencia, informar estrategias de mitigación de riesgos, facilitar la asignación de pruebas escasas recursos y alertar la cuarentena adecuada y tratamiento de los afectados.

En este artículo, se establece una Arquitectura de Vigilancia Epidemiológica, para unir a científicos con experiencia en investigación de grandes datos y epidemiología, para desarrollar un producto mínimo viable. Esta aplicación, que ofrece datos sobre factores de riesgo, síntomas predictivos, resultados clínicos y geográficos puntos críticos. Nuestra iniciativa ofrece una prueba de concepto para la reutilización de los enfoques existentes para permitir la recopilación de datos epidemiológicos rápidamente escalables y análisis, que es fundamental para una respuesta basada en datos a este desafío de salud pública.

## 2. Revisión de literatura

Debido a la no disponibilidad de un tratamiento eficaz o una vacuna para enfermedades infecciosas o de naturaleza transmisible, los esfuerzos para hacer frente y mitigar los efectos de las enfermedades, adoptados por las entidades relacionadas con la salud pública, comprenden una amplia gama de estrategias que van desde el uso de la ciencia social y de comportamiento (Bavel et al., 2020) o la psicología (Brady, Crockett & Bavel, 2019; Brooks et al., 2020) para analizar el comportamiento social y definir directrices para una transformación de los hábitos de comportamiento que permitan mitigar la diseminación, hasta la incorporación de herramientas tecnológicas y tecnología digital en las políticas públicas y la atención médica (Whitelaw et al., 2020; Ferretti, 2020; Wang, Chun & Brook, 2020; McCall, 2020). Sin embargo, la estrategia ampliamente utilizada como mecanismo de apoyo en la construcción de políticas, así como el desarrollo de nuevos estudios clínicos

relacionados con eventos críticos en salud, como la aparición o reaparición de pandemias, responden a la construcción de sistemas de vigilancia epidemiológica (Choi, 2012; Nsubuga et al., 2006).

Los sistemas de vigilancia epidemiológica se focalizan en construir herramientas que permitan la detección temprana y la construcción rápida de respuestas para controlar las enfermedades infecciosas (Choi et al., 2012). El desarrollo de la construcción de iniciativas de sistemas de vigilancia ha ido incorporando las tecnologías disponibles para ser cada vez más eficientes, dar respuestas en tiempo real y generar alertas tempranas (Choi et al., 2016). Incluyen además una variedad de tecnologías y fuentes de datos que se agregan y evalúan continuamente.

Antes de la generalización de la construcción de sistemas de vigilancia, las iniciativas estuvieron centradas en la construcción de repositorios digitales de datos clínicos, información clínica, e información médica relacionada. La estrategia consistía en concentrar la información diseminada a través de diversos laboratorios para ofrecerlos públicamente a la comunidad académica, esto como una estrategia de incentivar futuros desarrollos clínicos. Ejemplos de estas aproximaciones son: el sistema del Instituto Nacional de Diabetes y Enfermedades Digestivas y del Riñón NIDDK (Cuticchia et al., 2006) es un repositorio de web de muestras de ADN, muestras biológicas y resultados de ensayos clínicos de enfermedades relacionadas. Reúne un conjunto de datos de laboratorios clínicos quienes mantienen y alimentan el sistema. Es un intento de centralizar información clínica relacionada, de acceso público con información de co-localización e información entre los datos del paciente y las muestras recolectadas. El repositorio MyPubliHealth (Revere, Bugni & Fuller, 2007) es un repositorio de recursos de salud pública que utiliza que organiza información clínica mediante el esquema de metadatos para facilitar la búsqueda y hacerlas eficientes.

De la mano del desarrollo de las tecnologías de la información fue la incorporación de estas en sistemas computacionales que extendía su funcionalidad más allá de la centralización de la información, e incorporan funcionalidades como modelamiento de comportamientos y de políticas públicas (Duke-Sylvester et al., 2008), así como modelos de predicciones (Zixin et al., 2020; Wu, Leung & Leung, 2020). Un número creciente en las últimas dos décadas sobre sistemas de vigilancia o relacionados es evidenciado en los reportes de la comunidad científica. Una búsqueda preliminar para analizar las tendencias de los reportes científicos realizados por la comunidad entre el año 2000 y 2019 sobre tres bases de datos científicas: PubMed, Scopus y Web of science, considerando los términos: epidemiological surveillance, digital disease detection, biosurveillance, medical online surveillance, medical web surveillance system, y relacionados mediante la siguiente ecuación de búsqueda: ( "epidemiological" AND "surveillance" ) OR ( "digital" AND "disease" AND "detection" ) OR ( "biosurveillance" ) OR ( "online" AND "surveillance" AND "medical" ) OR ( "web" AND "surveillance" AND "systems" AND "medical" ), muestra que en las dos últimas décadas la tendencia en el reporte es creciente pasando en promedio de 375 documentos reportados para el año 2000 a 2141 documentos reportados en el año 2019 (Figura 2).

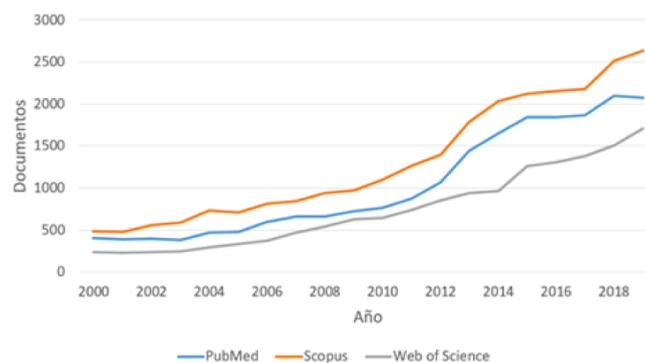


Figura 2. Número de documentos publicados, relacionados con sistema de vigilancia epidemiológica para el periodo 2000-2019.

Fuente: Elaboración propia.



El mayor número de publicaciones, analizando las referencias de PubMed y Scopus, tiene origen en países como Estados Unidos, Reino Unido, China y Francia, entre otros. En Latinoamérica, el país con mayor número de publicaciones es Brasil con un promedio de 1.290 publicaciones, distante del segundo, Colombia, que registra cerca de 250 publicaciones (Figura 3).

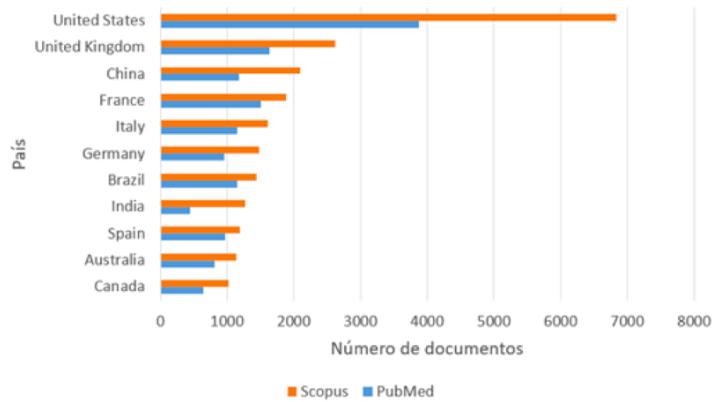


Figura 3. Número de documentos publicados por país para el período 2000-2019. Fuente: Elaboración propia.

### 3. Políticas nacionales

Un fenómeno creciente es la incorporación de estos sistemas de vigilancia a los sistemas nacionales de los países. En estados unidos ejemplo de esto son la Plataforma de datos de vigilancia (SDP) del Center for Surveillance, Epidemiology, and Laboratory Services en Estados Unidos es una herramienta tecnológica creada en 2016 con tecnologías basadas en la nube y orientada al seguimiento de enfermedades para mejorar las capacidades de vigilancia de la salud pública (CSELS, 2016). El Sistema Nacional de Notificación de Vigilancia de Enfermedades (NNDSS) del U.S. Department of Health & Human Services incorpora tecnología para la vigilancia en la recolección, análisis e intercambio de datos de salud relacionando información de cerca de 120 enfermedades diferentes. Esta información es centralizada para que cada entidad o departamento de salud pueda monitorear las tendencias de las enfermedades, analizar factores de riesgo y evaluar políticas de prevención y control (CHHS, 2019).

Desde la perspectiva tecnológica algunos trabajos reportan el detalle del desarrollo de estos sistemas, el modelo de datos, y la arquitectura del software que constituye al sistema de vigilancia. En (Lopes et al., 2010) se describe los requerimientos, arquitectura y el desarrollo de un sistema de información para datos epidemiológicos con el objetivo de la integración con otras plataformas existentes orientadas al modelamiento y la predicción bajo el concepto de herramientas de software libre. En esta dirección (Zamite et al., 2010) describe los requerimientos para el desarrollo de un sistema de recolección de datos epidemiológicos desde diversos tipos de fuentes como consultas a bases de datos de los servicios de información nacional, servicios de noticias, directamente desde los usuarios y mediante fuentes alternativas como redes sociales mediante la minería de texto.

En el contexto específico de la pandemia causada por el coronavirus (Covid-19) un número de aplicaciones tecnológicas han sido diseñadas con el objeto de contrarrestar la expansión, recolectar mayor conocimiento sobre el virus y sus efectos, y apoyar las actividades diarias mientras un tratamiento efectivo es determinado o la vacuna es desarrollada. La organización mundial de la salud ha recomendado el uso de sistemas de vigilancia tecnológica como una estrategia de mitigación (World Health Organization, 2020), principalmente para focalización de grupos de riesgo, aplicación temprana de pruebas de testeo de presencia de virus y seguimiento de contactos de casos positivos.

El parlamento europeo ha recomendado diez tecnologías útiles para la crisis del Covid-19 (Kritikos, 2020)

que han sido incorporadas en los sistemas de vigilancia o herramientas similares, a saber: Inteligencia Artificial (Apostolopoulos, Aznaouridis & Tzani, 2020; Rashid & Wang, 2020; Fong, Dey & Chaki, 2021), Blockchain (Chang & Park, 2020; Resiere, Resiere & Kallel, 2020), tecnologías de código abierto, telemedicina (Contreras et al., 2020), impresión 3D (Tino et al., 2020), tecnologías de edición genética, nanotecnología (Ceylan, Meral & Cetinkaya, 2020), síntesis biológica, drones y robótica.

Sin embargo, aproximaciones tecnológicas focalizadas en esa dirección corresponde a una variedad de aproximación como bases de datos centralizadas de información clínica, científica, etc. Algunos ejemplos son: El repositorio centralizado Covid19-Lake de Amazon Web Services (Amazon, 2020) relacionado con la expansión y las características del virus, de forma similar Google creó el Google Cloud Platform (Google, 2020) para almacenar bases de datos disponibles públicamente centralizando datos de diversos laboratorios y sistemas de salud pública alrededor del mundo. Iniciativas similares son CAS COVID-19 (American Chemical Society, 2020; CDCP, 2020) que reúne información relacionada con casos de contagios e información de vigilancia epidemiológica. Otras iniciativas intentan usar herramientas tecnológicas para crear bases de datos relacionados con la información de investigación clínica e investigación científica alrededor del virus como (Semantic Scholar, 2020; National Institutes of health, 2020).

En el proceso de recolección de la información recientemente se han diversificados las fuentes, además de datos clínicos e información científica, los sistemas han empezado a considerar información proveniente de redes sociales, en las cuales análisis de sentimientos y percepciones psicológicas (Bayer et al., 2020) de la sociedad podrían sugerir estrategias de interés para la salud pública y la salud mental (Chen & Yu, 2020), así como mecanismo eficaz de información ciudadana (Moore et al., 2020).

En general, las tecnologías de la información se constituyen como herramienta imprescindible para hacer frente a crisis de salud pública para los países y organizaciones la amplia variedad de sistemas tecnológicos desarrollados es un indicativo en esa dirección. Si bien los datos y la información son crucial para alimentar sistemas de vigilancia epidemiológica, esto ha generado algunas consideraciones y discusiones frente a la utilización y recolección de datos en el contexto del uso ético (Xafis et al., 2020) y de las garantías de protección de la información de ciudadanos (Zwitter et al., 2020) que limitan la eficacia de la aplicación de los sistemas tecnológicos y que deberán ser concertados para futuras crisis de salud pública.

## 4. Metodología

Construir una plataforma de vigilancia epidemiológica (PVE) requiere considerar múltiples aspectos como: actores, fuentes de información, integraciones y tecnologías asociadas, por mencionar algunos. Es quizá por ello que, si bien la mayoría de los países cuentan con sistemas de vigilancia epidemiológica, son pocos los casos en los que se pueda evidenciar una plataforma de vigilancia epidemiológica con un nivel de desarrollo tecnológico que contemple la integración de los distintos actores a nivel privado y público, y con gran capacidad de brindar información relevante y oportuna a la ciudadanía.

En esta sección se discuten los aspectos básicos que deben ser tenidos en cuenta en la construcción de una PVE que permita a las sociedades, en un futuro cercano, tener un mejor nivel de detección y reacción ante crisis epidemiológicas como las que hoy afrontamos.

### 4.1. Actores y público objetivo

Si bien en principio, los distintos actores del sector salud son los llamados a atender una crisis epidemiológica, la situación que actualmente afrontamos demuestra que este tipo de crisis requieren de la cooperación desde el ciudadano hasta todos los entes gubernamentales, todos con distintos intereses, pero parte de un común. Dicho lo anterior, consideremos los siguientes actores principales de la plataforma tecnológica y sus intereses de información:

- **Epidemiólogos:** Siendo la epidemiología la ciencia que estudia el comportamiento y evolución de la salud en una población, los epidemiólogos requieren principalmente información que les permita: a) evidenciar las



causas y/o factores de riesgo de las enfermedades, b) lanzar predicciones que permitan anticiparse a las dinámicas de propagación, y c) emitir acciones de control que minimicen dichas dinámicas.

- **Personal médico:** De manera general se puede hacer referencia a todos los actores del sector de la medicina, que van desde médicos generales y enfermeros, hasta especialistas directamente relacionados con la enfermedad, como virólogos, bacteriólogos, cardiólogos y nutricionistas. Estos actores requieren información mediante la cual puedan caracterizar la enfermedad y su patógeno, y que posibilite la rápida convergencia de consensos médicos.

- **Centros médicos y laboratorios:** Siendo estos los responsables de la logística de suministros y recursos físicos y de personal, principalmente requieren información que les permita optimizar la efectividad y disponibilidad de sus recursos. Por ejemplo, una mejor caracterización de la enfermedad incide directamente en un mejor triaje o calibrado y por ende una menor tasa de falsos positivos que puedan colapsar la disponibilidad de un recurso crítico.

- **Gobierno:** En medio de una crisis epidemiológica, el gobierno en todos los niveles requiere información que soporte la toma de decisiones que afectan a todos los integrantes de la sociedad, ya sean estas la emisión de comunicados, decretos, políticas locales, licitaciones o convocatorias, entre otras. Usualmente esta información está asociada a la estadística descriptiva y cifras globales de la epidemia misma y su impacto en distintos ámbitos de la sociedad.

- **Ciudadanía:** En la era de la desinformación, tal vez los actores mayormente afectados durante una crisis epidemiológica. En una plataforma de vigilancia epidemiológica, la ciudadanía se constituye como la principal fuente de datos sobre la epidemia, y por otra parte, una ciudadanía informada sobre las cifras globales y comportamientos locales de la epidemia, tendrá una mayor conciencia y autorregulación de su comportamiento.

## 4.2. Componentes

A partir de los actores anteriormente mencionados y sus requerimientos de información, se consideran los siguientes componentes que debe considerar una PVE.

### 4.2.1. Captura e integración de datos

Ya sea mediante la ejecución de estudios o utilización de alguna aplicación web o móvil, el empoderamiento ciudadano a través del ejercicio del auto monitoreo, es la principal vía de captura de datos relacionados con la epidemia: síntomas, signos y pre-existencias (co-morbilidad) de la población. Sin embargo, estos datos por sí solos conforman una estimación del estado de salud poblacional, inicialmente sin relación con la epidemia en concreto. Es por ello que es necesario integrar los datos capturados mediante el auto-seguimiento ciudadano, con los datos signos, síntomas y diagnósticos que puedan ser capturados en centros médicos, así como resultados de laboratorios que permitan realizar confirmaciones sobre dichos datos (Figura 4).

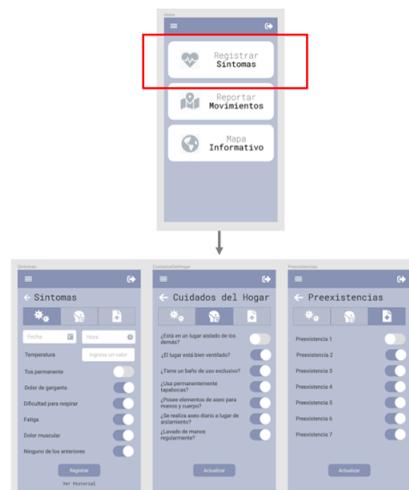


Figura 4. Ejemplo de formulario de captura de datos mediante aplicación móvil. Fuente: Elaboración propia.

- Aplicación móvil & contact-tracing: La captura de datos de los síntomas, signos y pre-existencias de la ciudadanía es comúnmente realizada mediante un formulario de preguntas (ver Figura 4) el cual debe realizarse con una frecuencia diaria. Por otra parte, la utilización de tecnologías de contact-tracing, ha resultado ser muy útil en diversos países como Corea del Sur, China e Israel, donde la combinación de distintas tecnologías como bluetooth, GPS, antenas celulares y wearables, han permitido establecer aplicaciones de pruebas de manera más efectiva y cercos de contagio, una vez identificados ciudadanos con casos positivos de COVID-19 y las personas que tuvieron contacto con estos.

- Integración con centros médicos y laboratorios: Para que una PVE sea realmente efectiva, debe poder integrar datos desde los centros médicos relacionados con la disponibilidad de sus recursos, lo cual generalmente se realiza de manera manual mediante reportes que se entregan de manera periódica a las autoridades. Finalmente, los laboratorios son los encargados de suministrar los datos de exámenes realizados en la población que pueden ser utilizados como confirmaciones positivas o negativas sobre los datos capturados mediante la aplicación móvil.

- Integración con fuentes públicas de datos: Es también importante considerar la inclusión de datos (datasets) públicos que ayuden a afrontar el problema del arranque en frío (cold start) de los modelos de analítica al interior de la PVE. En este punto se hace necesario el diseño y construcción de pipelines de procesamiento de datos estructurados y no-estructurados.

#### 4.2.2. Telemedicina y tele-seguimiento

Debido a que el aislamiento físico de la población es una de las acciones de mayor impacto durante el control de una epidemia por enfermedad contagiosa, la telemedicina en conjunto con el tele-seguimiento o seguimiento remoto de pacientes, se presentan como una alternativa tecnológica que permite descongestionar los centros médicos en pro de optimizar sus recursos instalados, a la vez que se obtienen datos importantes de diversos signos vitales de los pacientes, mientras se controla su estado de salud desde sus casas.

Aunque esta tecnología aún tiene un nivel bajo de implementación en la mayoría de países, es importante considerarla dentro de la arquitectura de una PVE, ya que aporta en diversos objetivos de la plataforma, como lo son, la preservación del aislamiento físico de los ciudadanos, la optimización de recursos instalados en los centros médicos, y en la captura de datos de vital importancia para los modelos de analítica al interior de la PVE.

#### 4.2.3. Analítica de datos

Todos los actores del sistema pueden verse beneficiados por la aplicación de un conjunto de técnicas de analítica de datos que apoyen sistemas de recomendación y predicción (Figura 5).

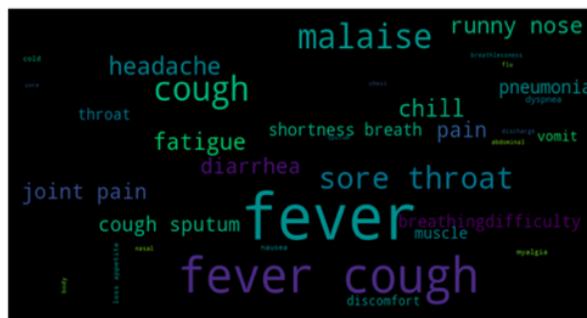


Figura 5. Nube de palabras extraídas mediante el análisis de síntomas del dataset Google Sheets from DXY.cn. Fuente: (Sun, Chen & Viboud, 2020).

Entre los modelos y/o aproximaciones que inicialmente se deben considerar en el diseño y construcción de una PVE, se encuentran:

- Analítica descriptiva: Principalmente orientada a la extracción de información relacionada con cifras

Atencio, P.; Sánchez-Torres, G.; Iral Palomino, R.; Branch Bedoya, J. W.; Burgos, D. (2021). Arquitectura conceptual de plataforma tecnológica de vigilancia epidemiológica para la COVID-19. *Campus Virtuales*, 10(1), 21-34.

globales de la epidemia que apoyen la toma de decisiones para los gobiernos, así como el análisis evolutivo de la enfermedad para los epidemiólogos. En este punto, la implementación de diversas técnicas de visualización mediante dashboards, por ejemplo, mapas de calor e histogramas, son de gran utilidad para muchos de los actores de la PVE.

- **Detección de patrones sintomáticos:** De especial interés para todo el personal médico, los patrones sintomáticos indican los síntomas y signos asociadas a la enfermedad, los cuales podrían cambiar de manera automática a partir de los datos recolectados por la PVE. El sistema debe aprender de manera dinámica a) la ponderación de cada síntoma y/o signo asociado a la enfermedad, así como la identificación o ingreso de nuevos síntomas que surjan de los datos recolectados por la PVE y, b) el sistema de generar reportes sobre scores asociados a síntomas y signos con respecto a la enfermedad. Un ejemplo de este tipo de análisis puede ser observado en la nube de palabras presentada en la Figura 5, obtenida del dataset Google Sheets from DXY.cn de 3396 casos reportados en diferentes países hasta febrero del presente año.

- **Modelos de estimación del riesgo:** Estos modelos permiten estimar una probabilidad de contagio en la población a partir los datos de: a) síntomas, b) signos, c) pre-existencias, d) ubicación y e) contact-tracing. A partir de estos modelos se pueden construir mapas aproximados de distribución de la enfermedad, así como emitir recomendaciones de manera automática a los ciudadanos a través de la aplicación móvil.

- **Modelos de recomendación para aplicación de pruebas:** Generalmente, en los inicios de una epidemia, poco se sabe sobre la sintomatología asociada a la enfermedad, por lo que los consensos médicos son escasos y se van construyendo a partir de la experiencia que pueden ir construyendo médicos en grupos locales y socializaciones con otros grupos. Particularmente en el caso de la COVID-19, su similitud sintomatológica en los estados iniciales con un resfriado común ha dificultado el proceso de triaje o calibrado de los pacientes infectados. Por otra parte, la construcción de los consensos médicos se dificulta en la medida que no exista manera de unificar y estandarizar los datos recopilados en los triajes y aplicaciones de pruebas realizadas a pacientes en distintos centros médicos y en distintas regiones de un país. Es por ello por lo que, la utilización de los modelos de detección de patrones sintomáticos ayudaría a construir consensos médicos de una manera más rápida, así como en mejorar la efectividad en la recomendación de la aplicación de pruebas a pacientes.

- **Modelos de recomendación para tratamientos:** Al igual que en la aplicación de pruebas, los datos recolectados a través de la PVE pueden ser procesados para identificar tratamientos potencialmente efectivos. Sin embargo, esto requerirá necesariamente validación experimental médica.

#### 4.2.4. Planeación y acción

Los modelos anteriores, complementados con herramientas visuales tipo dashboards, resultan de gran utilidad para la planeación de estrategias de mitigación y control de la enfermedad. Por ejemplo, la visualización de focos de contagio y alto riesgo sobre un mapa puede ser utilizado para realizar fumigaciones y/o cercos de contención en determinadas zonas de una ciudad, así como para identificar posibles causas de brotes mediante la tecnología de contact-tracing.

#### 4.2.5. Optimización de recursos

Finalmente, la información obtenida de los modelos de riesgo en la población, así como de la analítica descriptiva y de los datos de la ciudadanía, pueden ser utilizados por los centros médicos para optimizar el uso de sus recursos físicos, así como por las autoridades para realizar una gestión de los territorios más efectiva y basada en datos. Por ejemplo, un municipio puede optimizar recursos importantes de su red de emergencias, tales como ambulancias e infraestructura operativa, mediante modelos de riesgo más precisos.

### 5. Resultados

En la Figura 6 se puede observar un esquema conceptual de una PVE según los actores y componentes descritos anteriormente.

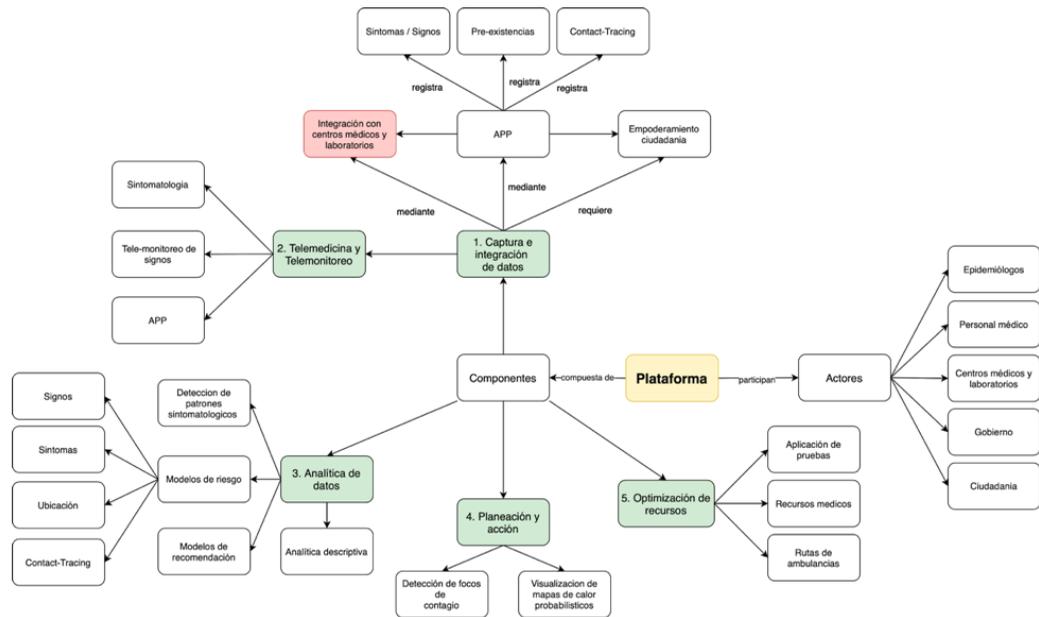


Figura 6. Mapa conceptual de una PVE. Fuente: Elaboración propia.

De cara a la implementación rápida de una PVE es importante considerar una versión inicial o producto mínimo viable de la plataforma, en la cual se consideren las funcionalidades básicas que permitan dar inicio al proceso de captura y análisis de datos para generar información. En la Figura 7 se puede presentar un producto mínimo viable de la plataforma PVE. En esta versión se considera principalmente la interacción de la ciudadanía con la aplicación móvil para el registro y seguimiento de signos, síntomas y pre-existencias, así como del personal médico y laboratorios para confirmar resultados de tests realizados para generar etiquetas reales dentro de los datos recopilados que ayuden a complementar los datos para los modelos de riesgo. Por otra parte, que con el fin de superar el impase que representa el problema del arranque en frío en los sistemas de recomendación, es importante tener estrategias de aprendizaje semi-supervisado para poder extrapolar etiquetas a partir de los casos confirmados y los datos no etiquetados ingresados por la ciudadanía.

Un producto mínimo viable de la PVE debe considerar la visualización de información descriptiva mediante dashboards y de información geoespacial con el fin de que los gobiernos puedan actuar para la mitigación y contención de la epidemia.

En la Figura 8 se observa un ejemplo de un mapa de calor dinámico con una ventana de 40 días utilizando datos públicos y anonimizados, generado mediante el producto mínimo viable de PVE. Este tipo de visualizaciones resulta muy útil para el análisis geoespacial del crecimiento y dinámica epidemiológica de la enfermedad.

Finalmente, para la implementación del producto mínimo viable de la PVE, fue utilizado el stack tecnológico descrito en la Figura 9 compuesto principalmente por: a) una aplicación móvil con el formulario para el registro de signos, síntomas y pre-existencias, b) un pipeline de unificación y transformación de datos tomados de múltiples datasets disponibles a nivel público, c) un motor de analítica para el entrenamiento de modelos de riesgo, expuestos mediante un servicio que utiliza una API-REST, y d) una interfaz web de para la visualización de datos descriptivos y de información geográfica.

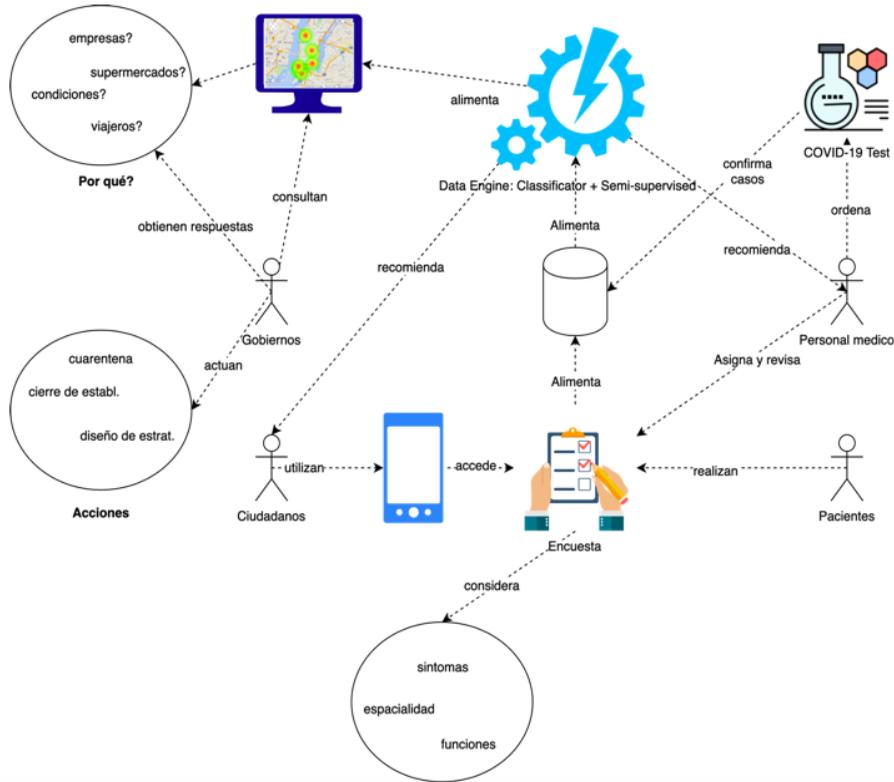


Figura 7. Producto mínimo viable para la PVE diseñada. Fuente: Elaboración propia.

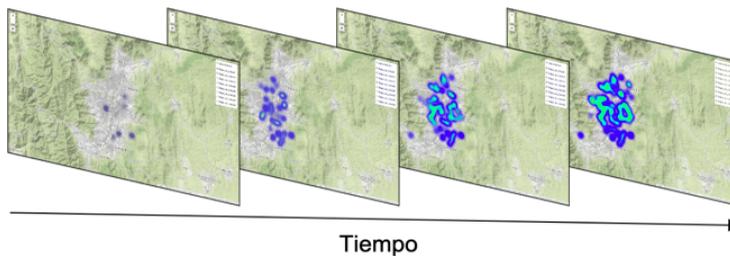


Figura 8. Ejemplo de un mapa de calor dinámico generado con datos públicos de casos positivos de COVID-19 en la ciudad de Medellín, Antioquia, Colombia. Fuente: Elaboración propia.

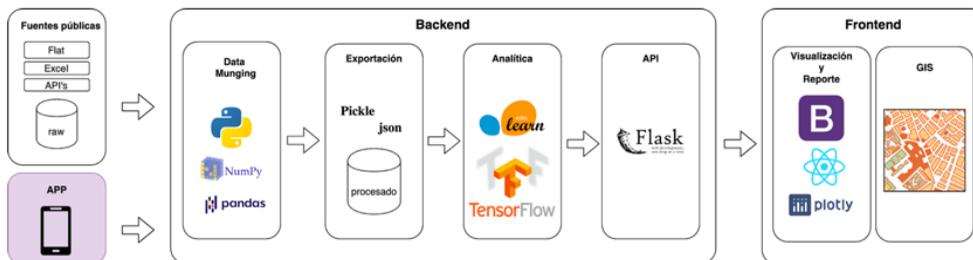


Figura 9. Arquitectura de alto nivel de la analítica de datos al interior de la PVE. Fuente: Elaboración propia.

## 6. Conclusión

La tecnología ha demostrado ser una herramienta útil y necesaria para ayudar a garantizar que los gobiernos locales y regionales en la primera línea de la emergencia continúen brindando servicios públicos

esenciales durante la crisis de la COVID-19. A medida que el coronavirus se propaga por todo el mundo, los gobiernos han establecido restricciones importantes sobre el movimiento de personas, el funcionamiento de los servicios y las normas sobre distanciamiento físico. En este contexto, la tecnología puede causar un efecto profundo en la vida cotidiana de los ciudadanos y garantizarles el acceso a los servicios de salud, el acceso a la información y la comunicación con las autoridades competentes, entre otras cosas. La Plataforma Tecnológica de Vigilancia Epidemiológica presentada en este artículo muestra una arquitectura básica para analizar y predecir la evolución de la pandemia utilizando, por ejemplo, mapas de calor.

#### Cómo citar este artículo / How to cite this paper

Atencio, P.; Sánchez-Torres, G.; Iral Palomino, R.; Branch Bedoya, J. W.; Burgos, D. (2021). Arquitectura conceptual de plataforma tecnológica de vigilancia epidemiológica para la COVID-19. *Campus Virtuales*, 10(1), 21-34. ([www.revistacampusvirtuales.es](http://www.revistacampusvirtuales.es))

## Referencias

- Amazon (2020). Amazon Web Services (AWS) data lake for analysis of COVID-19 data. (<https://djt2aa9i652rf.cloudfront.net/>).
- American Chemical Society (2020). CAS COVID-19 antiviral candidate compounds dataset. (<https://www.cas.org/covid19>).
- Apostolopoulos, I.; Aznaouridis, S.; Tzani, M. (2020). Extracting Possibly Representative COVID-19 Biomarkers from X-ray Images with Deep Learning Approach and Image Data Related to Pulmonary Diseases. *Journal of Medical and Biological Engineering*, 40, 462-469. doi:10.1007/s40846-020-00529-4.
- Bavel, J. J. V.; Baicker, K.; Boggio, P. S.; Capraro, V.; Cichocka, A.; Cikara, M.; ...; Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 1-12. 4, 460-471. doi:10.1038/s41562-020-0884-z.
- Bayer, J. B.; Triệu, P.; Ellison, N. B. (2020). Social Media Elements, Ecologies, and Effects. *Annual Review of Psychology*, 71(10), 471-497.
- Brady, W. J.; Crockett, M.; Van Bavel, J. J. (2019). The MAD Model of Moral Contagion: The role of motivation, attention and design in the spread of moralized content online. *Perspectives on Psychological Science*, 15(4), 978-1010. doi:10.31234/osf.io/pz9g6.
- Brooks, S. K.; Webster, R. K.; Smith, L. E.; Woodland, L.; Wessely, S.; Greenberg, N.; Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The Lancet*, 395(1), 912-920.
- Cabero-Almenara, J.; Llorente-Cejudo, C. (2020). Covid-19: transformación radical de la digitalización en las instituciones universitarias. *Campus Virtuales*, 9(2), 25-34.
- CDCP - Center for Disease Control and Prevention. (2020). CDC COVID-19 Cases, Data, and Surveillance. (<https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/>).
- Ceylan, Z.; Meral, R.; Cetinkaya, T. (2020). Relevance of SARS-CoV-2 in food safety and food hygiene: potential preventive measures, suggestions and nanotechnological approaches. *VirusDisease*, 31, 154-160. doi:10.1007/s13337-020-00611-0.
- Chang, M. C.; Park, D. (2020). How Can Blockchain Help People in the Event of Pandemics Such as the COVID-19?. *Journal of Medical Systems*, 44, 1-2. doi:10.1007/s10916-020-01577-8.
- Chen, X.; Yu, B. (2020). First two months of the 2019 Coronavirus Disease (COVID-19) epidemic in China: real-time surveillance and evaluation with a second derivative model. *Global health research and policy*, 5(1), 1-9.
- Choi, B. C. (2012). The past, present, and future of public health surveillance. *Scientifica*, 2012(875253), 875253. doi:10.6064/2012/875253.
- Choi, J.; Cho, Y.; Shim, E.; Woo, H. (2016). Web-based infectious disease surveillance systems and public health perspectives: a systematic review. *BMC public health*, 16(1), 1238.
- Contreras, C. M.; Metzger, G. A.; Beane, J. D.; Dedhia, P. H.; Ejaz, A.; Pawlik, T. M. (2020). Telemedicine: Patient-Provider Clinical Engagement During the COVID-19 Pandemic and Beyond. *Journal of Gastrointestinal Surgery*, 1. doi:10.1007/s11605-020-04623-5.
- CSELS - Center for Surveillance, Epidemiology, and Laboratory Services (2020). Surveillance Data Platform (SDP) Program Efforts. Estados Unidos: U.S. Department of Health & Human Services. (<https://www.cdc.gov/sdp/index.html>).
- Cuticchia, A. J.; Cooley, P. C.; Hall, R. D.; Qin, Y. (2006). NIDDK data repository: a central collection of clinical trial data. *BMC medical informatics and decision making*, 6(1), 19.
- Deshpande, P. S.; Sharma, S. C.; Peddoju, S. K. (2019). Predictive and Prescriptive Analytics in Big-data Era. In *Security and Data Storage Aspect in Cloud Computing* (pp. 71-81). Singapore: Springer.
- DHHS - Department of Health & Human Services (2020). El Sistema Nacional de Notificación de Vigilancia de Enfermedades (NNDSS), Epidemiology, and Laboratory Services. Estados Unidos: U.S. (<https://www.cdc.gov/nndss/>).
- Duke-Sylvester, S. M.; Perencevich, E. N.; Furuno, J. P.; Real, L. A.; Gaff, H. (2008). Advancing epidemiological science through computational modeling: a review with novel examples. In *Annales Zoologici Fennici* (Vol. 45, No. 5, pp. 385-401). Finnish Zoological and Botanical Publishing Board.



- Ferretti, L.; Wymant, C.; Kendall, M.; Zhao, L.; Nurtay, A.; Abeler-Dörner, L.; ...; Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491), 1-9.
- Fong, S. J.; Dey, N.; Chaki, J. (2020). AI-Enabled Technologies that Fight the Coronavirus Outbreak. In S. J. Fong, N. Dey & J. Chaki (Eds.), *Artificial Intelligence for Coronavirus Outbreak* (pp. 23-45). Singapore: Springer. doi:10.1007/978-981-15-5936-5\_2.
- García-Peñalvo, F. J.; Corell, A. (2020). La COVID-19: ¿enzima de la transformación digital de la docencia o reflejo de una crisis metodológica y competencial en la educación superior?. *Campus Virtuales*, 9(2), 83-98.
- Google (2020). Google Cloud Platform (GCP) Datasets for COVID-19 Research. (<https://console.cloud.google.com/marketplace/browse?filter=category:covid19&pli=1>).
- Guibovich, G. (2009). Sistemas de Información para Vigilancia de la Salud 2009. OPS/OMS. ([https://www.paho.org/venezuela/index.php?option=com\\_docman&view=download&category\\_slug=presentaciones&alias=22-sistemas-de-informacion-para-vigilancia-de-la-salud&Itemid=466](https://www.paho.org/venezuela/index.php?option=com_docman&view=download&category_slug=presentaciones&alias=22-sistemas-de-informacion-para-vigilancia-de-la-salud&Itemid=466)).
- Kappen, T. H.; Van Klei, W. A.; Van Wolfswinkel, L.; Kalkman, C. J.; Vergouwe, Y.; Moons, K. G. (2018). Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagnostic and prognostic research*, 2(1), 11.
- Kritikos, M. (2020). Ten technologies to fight coronavirus, Scientific Foresight Unit (STOA) of European Parliamentary Research Service, Brussels, PE 641.543 – April 2020.
- Lopes, L. F.; Silva, F. A. B.; Couto, F.; Zamite, J.; Ferreira, H.; Sousa, C.; Silva, M. J. (2010). Epidemic Marketplace: An Information Management System for Epidemiological Data. In S. Khuri, L. Lhotská & N. Pisanti (Eds.), *Information Technology in Bio- and Medical Informatics, ITBAM 2010* (pp. 31-44). Springer Berlin Heidelberg.
- Martínez, F. (2000). De la información a la acción: la vigilancia de la salud pública. Centro Nacional de Epidemiología. Instituto de Salud Carlos III. Ministerio de Sanidad y Consumo. *Revista Española de Salud Pública*.
- McCall, B. (2020). COVID-19 and artificial intelligence: protecting health-care workers and curbing the spread. *The Lancet Digital Health*, 2(4), e166-e167.
- Mergel, I.; Edelman, N.; Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, 36(4), 101385.
- Moore, J. H.; Barnett, I.; Boland, M. R.; Chen, Y.; Demiris, G.; Gonzalez-Hernandez, G.; ...; Holmes, J. H. (2020). Ideas for how informaticians can get involved with COVID-19 research. *BioData Mining*, 13, 3.
- National Institutes of health (2020). Open-Access Data and Computational Resources to Address COVID-19. (<https://www.nih.gov/coronavirus>).
- Nsubuga, P.; White, M. E.; Thacker, S. B.; Anderson, M. A.; Blount, S. B.; Broome, C. V.; ...; Trostle, M. (2006). Public health surveillance: a tool for targeting and monitoring interventions. *Disease control priorities in developing countries*, 2, 997-1018.
- Organización Panamericana de Salud (1974). Sistemas de vigilancia epidemiológica. *Boletín de la Oficina Sanitaria Panamericana (OSP)*; 76 (6), jun. 1974.
- Pappas, I. O.; Mikalef, P.; Giannakos, M. N.; Krogstie, J.; Lekakos, G. (2018). Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. *IseB*, 16 (3), 479-491.
- Ramírez-Montoya, M. S. (2020). Transformación digital e innovación educativa en Latinoamérica en el marco del COVID-19. *Campus Virtuales*, 9(2), 123-139.
- Rashid, M. T.; Wang, D. (2020). CovidSens: a vision on reliable social sensing for COVID-19. *Artificial Intelligence Review*. doi:10.1007/s10462-020-09852-3.
- Resiere, D.; Resiere, D.; Kallel, H. (2020). Implementation of Medical and Scientific Cooperation in the Caribbean Using Blockchain Technology in Coronavirus (Covid-19) Pandemics. *Journal of Medical Systems*, 44, 123. doi:10.1007/s10916-020-01589-4.
- Revere, D.; Bugni, P.; Fuller, S. (2007). A Public Health Knowledge Management Repository that Includes Grey Literature. *Publishing Research Quarterly*, 23(1), 65-70. doi:10.1007/s12109-007-9002-6.
- Safullin, M. R.; Akhmetshin, E. M. (2019). Digital transformation of a university as a factor of ensuring its competitiveness. *International Journal of Engineering and Advanced Technology*, 9(1), 7387-7390.
- Semantic Scholar (2020). CORD-19: COVID-19 Open Research Dataset and AI Challenge. (<https://www.semanticscholar.org/cord19>).
- Sun, K. S.; Chen, J.; Viboud, C. (2020). Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study *The Lancet Digital Health*, 2(4), 201-208. doi:10.1016/S2589-7500(20)30026-1.
- Tino, R.; Moore, R.; Antoline, S.; Ravi, P.; Wake, N.; Ionita, C. N.; ...; Chepelev, L. L. (2020). COVID-19 and the role of 3D printing in medicine. *3D Printing in Medicine*, 6(11). doi:10.1186/s41205-020-00064-7.
- Valenzuela, M. T. (2015). Vigilancia Epidemiológica. Universidad de los Andes. ([https://www.sabin.org/sites/sabin.org/files/oct21\\_1000valenzuela.pdf](https://www.sabin.org/sites/sabin.org/files/oct21_1000valenzuela.pdf)).
- Wang, J.; Chun, Y.; Brook, R. (2020). Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *JAMA*, 323(14), 1341-1342.
- Whitelaw, S.; Mamas, M. A.; Topol, E.; Van Spall, H. G. (2020). Applications of digital technology in COVID-19 pandemic planning and response. *The Lancet Digital Health*. doi:10.1016/S2589-7500(20)30142-4.
- World Health Organization (2020). Global surveillance for COVID-19 caused by human infection with COVID-19 virus. WHO. Retrieved 2 April 2020.
- Wu, J. T.; Leung, K.; Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*, 2020; 395(10225), 689-697.
- Xafis, V.; Schaefer, G. O.; Labude, M. K.; Zhu, Y.; Hsu, L. Y. (2020). The perfect moral storm: Diverse ethical considerations in the COVID-19 pandemic. *Asian Bioethics Review*.
- Zamite, J.; Silva, F. A. B.; Couto, F.; Silva, M. J. (2010). MEDCollector: Multisource Epidemic Data Collector. In S. Khuri, L. Lhotská &

- N. Pisanti (Eds.), *Information Technology in Bio- and Medical Informatics, ITBAM 2010* (pp. 16-30). Springer Berlin Heidelberg.
- Zixin, H.; Qiyang, G.; Shudi, L.; Jin, L.; Xiong, M. (2020). Artificial intelligence, forecasting of Covid-19 in China. (<https://arxiv.org/abs/2002.07112>).
- Zwitter, A.; Gstrein, O. J. (2020). Big data, privacy and COVID-19 – learning from humanitarian expertise in data protection. *Journal of International Humanitarian Action*. doi:10.1186/s41018-020-00072-6.

