IMMUNOWAR

Mejorando la salud de las personas con
el mejor equipo

Índice

Índice **1**

Introducción **4**

Misión 4

Visión 4

Valores 4

El equipo 4

Situación del sector en España / Europa **5**

Situación actual y tendencias 5

Actores 7

Referencias de proyectos BigData en el sector salud 9

Ámbito de actuación del proyecto 9

Definición de problema **10**

Problema 10

Entrevistas a clientes 13

Solución Funcional 17

Competencia 19

¿En qué somos diferentes? 21

DAFO 22

Modelo Normativo 22

Modelo de Negocio **24**

Definición del reto 24

Mapa de empatía 24

Análisis CAME 25

Integración del Caso de Uso con los Procesos de Negocio 25

Modelo Canvas 27

Análisis de riesgos de la empresa. **30**

Solución técnica **31**

Release 1.0 32
Release 2.0 33
Release 3.0 33
Release 4.0 34
Visión 360º del paciente (Analítica de datos) 34
ETL: Talend 36
Anonimización de datos 39
Datos reales para pruebas 40
Tableau 41
API Rest AWS modelos Machine Learning 42
Protocolo estándar medicina API Rest 44
Recomendaciones de instalación del conjunto de herramientas 45
Diagrama arquitectura completa 46
Modelos 46
Fuentes de datos 47
Clasificación de notas médicas y diagnóstico del doctor en lenguaje natural 49
Recomendador de antibióticos (análisis prescriptivo) 49
Modelo Predicción Antibiótico sin antibiograma 51
Modelo predicción Antibiótico con antibiograma 53
Modelo Predicción antibiótico tras test rápido molecular 55
Papers consultados 56
Planes operativos 57
Plan de implantación y recursos necesarios. 57
Sistemas on-premise. 57
Etapa de descubrimiento y entrenamiento de modelos. 58
Métricas. 59
Plan de expansión 59
Plan de precios 61
Plan de marketing 62

Estudio económico financiero **66**

Plan de Ingresos 66

Plan de Costes 67

Punto muerto 68

Cuenta de resultados. 68

Subvenciones 69

Valoración de la empresa y necesidades de fondos. 72

Líneas de Negocio Futuras y Nuevos Mercados **73**

Conclusiones **74**

Introducción

Misión

Nuestra empresa enmarca su propósito en uno de los pilares fundamentales de los Objetivos de Desarrollo Sostenible de la Organización de las Naciones Unidas, "**contribuir a un tratamiento certero para mejorar la salud de las personas**".

Visión

Ser una herramienta necesaria en la diagnosis clínica.

Valores

Sostenibles

Sólo la mejor prescripción. Reducirá el impacto en las personas haciendo al ser humano menos autoinmune a los fármacos y garantizará la mejor respuesta del cuerpo a las enfermedades.

Responsables

Eficaces en el asesoramiento. Recomendamos donde aportamos valor al profesional médico.

Innovadores

Mediante las últimas tecnologías en BigData ofrecemos recursos hasta ahora desconocidos que mejoran la salud de las personas.

El equipo

	Manuel E. Romero Navarro https://www.linkedin.com/in/manuel-romero-navarro-16896a30/ C.M.O.	Ingeniero informático con más de 10 años de experiencia como Consulting Health Service. / Certificación PMP / Certificación HL7 CDA. Numerosos proyectos de sistemas de información sanitarios implantados por todo el territorio nacional.
	Rocío Domínguez Jiménez https://www.linkedin.com/in/rdomjim/ C.O.O.	Ingeniera Superior Industrial con amplia experiencia en el desarrollo e, implementación de proyectos I+D+i, así como en búsqueda de convocatorias de ayudas y vías de financiación.

	María Dolores Paredes Garrido https://www.linkedin.com/in/mar%C3%ADa-dolores-paredes-garrido-58613518/ C.D.O	Ingeniera informática. Más de 10 años de experiencia gestionando proyectos en Unidad de Sistemas de Información y coordinación de área de desarrollo (integración hardware & software, compliance, cumplimiento de requisitos legales...).
	Miguel Saborido del Castillo Luna https://www.linkedin.com/in/miguel-saborido-ab5565163/ C.T.O	Ingeniero de Telecomunicaciones con más de 8 años de experiencia marcadas por la innovación y búsqueda continua de soluciones tecnológicas. Experto en lenguaje de programación Python, creando varias aplicaciones de escritorio en PyQt5 servidas por APIs REST. Amplios conocimientos en backend, base de datos y web services.
	Miguel Magaña https://www.linkedin.com/in/miguel-maga%C3%B1a-su%C3%A1rez/ C.F.O..	Ingeniero Industrial por la U. de Sevilla y Alta Dirección por el Instituto San Telmo, con experiencia en proyectos digitales y aplicaciones en movilidad. Con dilatada experiencia en Dirección de Operaciones de una importante empresa pública.
	José Á. Costa Jiménez https://www.linkedin.com/in/joseangelcosta/ C.E.O	Empresario ejecutivo con experiencia en empresas de tecnología y servicios de valor añadido. Sus empresas integran 130 personas y una facturación consolidada de 20.000.000 € en 2.017. Roles actuales: CIO, KAM, COO, QoS.

Situación del sector en España / Europa

Situación actual y tendencias

A mediados del año pasado la Unión Europea emitió un informe donde analizaba el sector sanitario en los países de la unión. (State of Health in the EU España Perfil Sanitario del país 2017). Algunos aspectos destacables del informe son los siguientes:

Una de las conclusiones del estudio para la sanidad española es que, como uno de los retos para el futuro, España tendrá que invertir más en el Sistema Nacional de Salud, lo que podría traducirse en un **gasto** público **mayor** o en más impuestos o requerimientos de

copago sanitario. Los dos escollos a los que hacer frente en los próximos años serán el **innegable progreso tecnológico** y el no menos preocupante envejecimiento de la población.

Una población cada vez más envejecida que plantea retos importantes para el futuro. El envejecimiento de la población hace que una de las conclusiones del estudio diga así: “Será necesario lograr una mejora de la eficiencia en la prestación de asistencia sanitaria y cuidados de larga duración para abordar de manera asequible las crecientes necesidades de una población que envejece”.

El gasto sanitario per cápita en España está por debajo de la media de la U.E, sin embargo, el número de médicos está por encima de la media de la U.E. Los tiempos de espera de los distintos servicios sanitarios representan un antiguo problema en España.

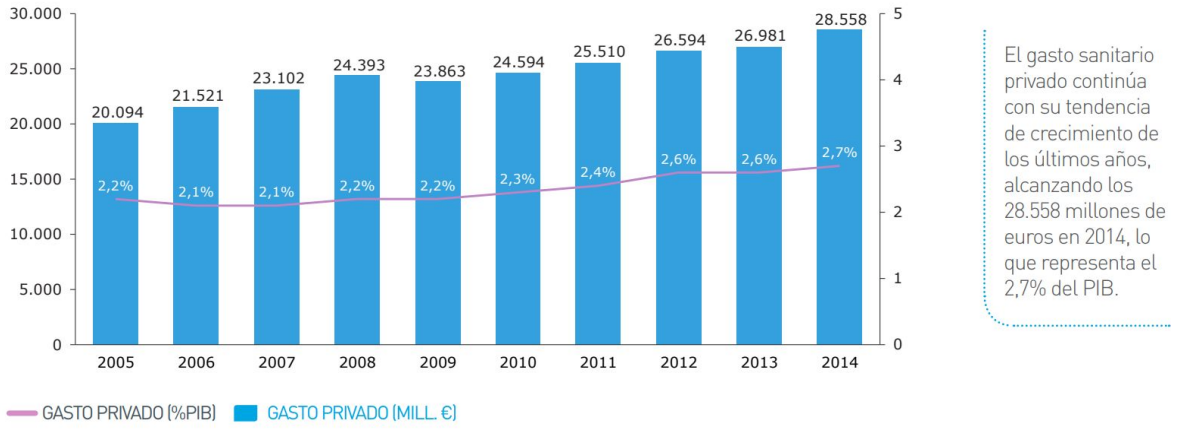
Desde 2013/2014 el **gasto público en sanidad se está incrementando** recuperando ratios pre-crisis.

El **gasto farmacéutico** descendió de media un 1 % al año en términos reales entre 2009 y 2014, pero **empezó a subir de nuevo en 2015**.

España está reasignando recursos desde una asistencia de escaso valor hacia otra más eficaz.

Según informe de la fundación IDIS sobre la sanidad pública y privada en 2017 en España se puede extraer los siguientes datos interesantes para nuestro proyecto.

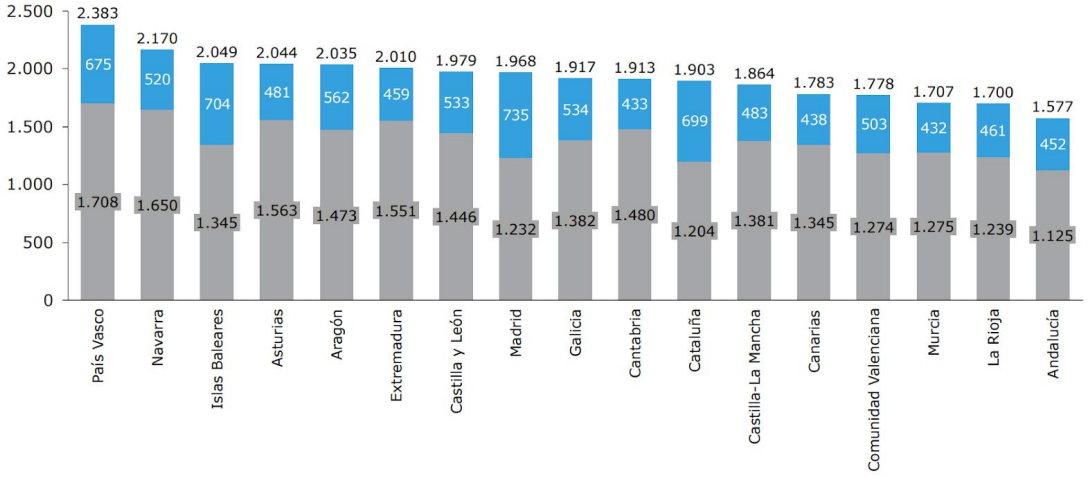
Gráfico 4 Evolución del gasto sanitario privado en Mill. € y en relación al PIB, 2005-2014



El gasto sanitario privado continúa con su tendencia de crecimiento de los últimos años, alcanzando los 28.558 millones de euros en 2014, lo que representa el 2,7% del PIB.

Nota: datos históricos actualizados en base a la última actualización disponible de OCDE.
Fuente: OCDE, Health Data 2016.

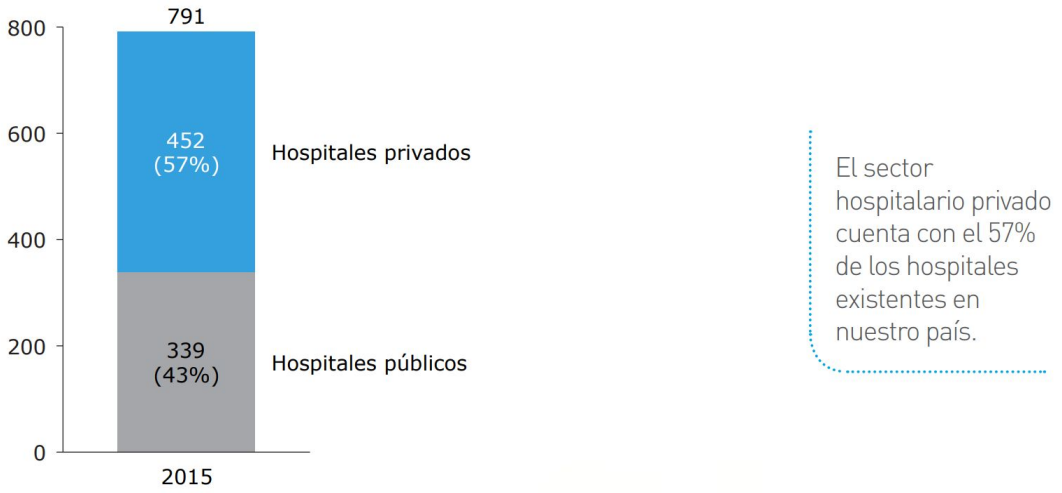
Gráfico 12 Estimación del gasto sanitario público y privado per cápita por comunidad autónoma, 2016



Actores

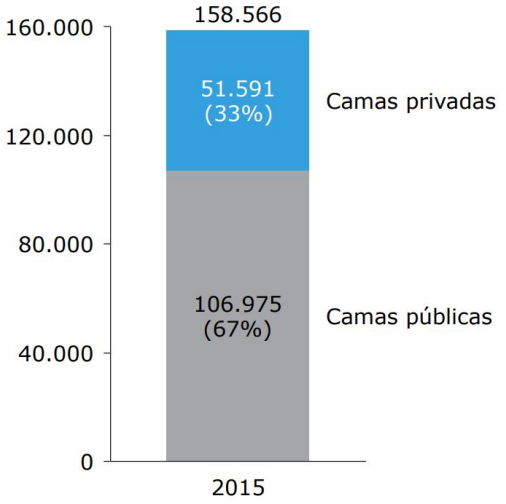
En nuestro país existen aproximadamente 3.000 centros médicos. La presencia de este tipo de centros en todo el territorio nacional asegura la accesibilidad de la población a la asistencia sanitaria, mejorando de este modo la atención y la satisfacción de los pacientes.

Gráfico 18 Número de hospitales en España, 2015



Nota: los complejos hospitalarios se contabilizan como un único hospital.
Fuente: Ministerio de Sanidad, Servicios Sociales e Igualdad, Catálogo Nacional de Hospitales 2016.

Gráfico 19 Número de camas en España, 2015



El sector hospitalario privado cuenta con el 33% de las camas existentes en nuestro país.

Fuente: Ministerio de Sanidad, Servicios Sociales e Igualdad, Catálogo Nacional de Hospitales 2016.

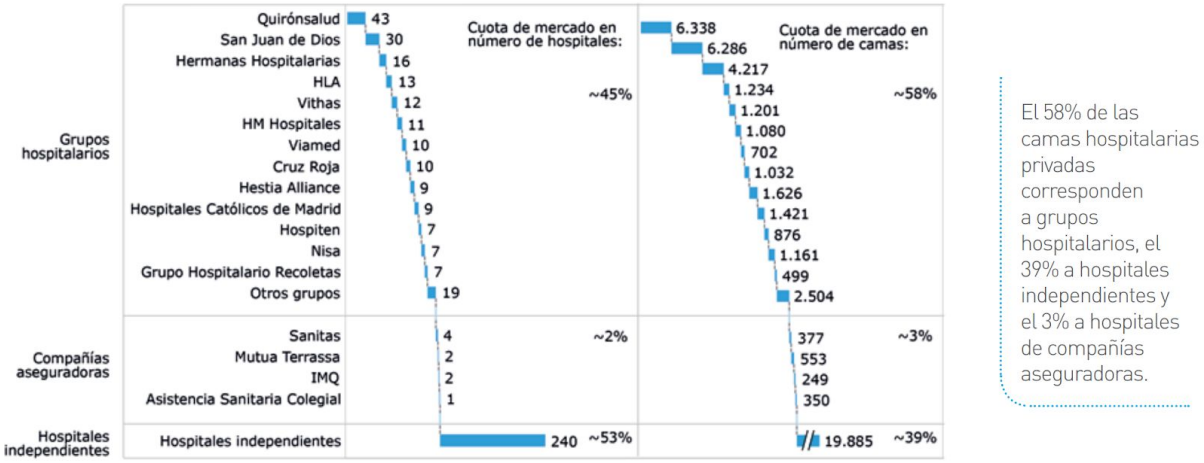
Tabla 8 Distribución de hospitales y camas por comunidad autónoma (ordenadas en función del número de camas privadas), 2015

Comunidad autónoma	Hospitales		Camas	
	Públicos	Privados	Públicas	Privadas
Cataluña	67	147	15.299	19.239
Madrid	34	47	14.473	7.005
Andalucía	46	62	15.089	5.889
Canarias	15	23	4.966	2.680
Comunidad Valenciana	36	25	10.899	2.446
País Vasco	17	24	5.613	2.347
Galicia	14	24	7.444	2.341
Castilla y León	15	21	7.213	2.262
Murcia	11	16	3.275	1.575
Islas Baleares	11	13	2.584	1.332
Asturias	9	11	2.764	998
Aragón	20	9	4.367	926
Navarra	4	7	1.405	883
Cantabria	3	4	1.419	658
Castilla-La Mancha	19	9	5.247	472
Extremadura	11	8	3.587	401
La Rioja	5	2	911	137
Ceuta	1	0	252	0
Melilla	1	0	168	0
Total	339	452	106.975	51.591

Cataluña, Madrid y Andalucía son las comunidades autónomas que cuentan con un mayor número de hospitales y camas privadas.

Fuente: elaboración propia a partir de datos del Ministerio de Sanidad, Servicios Sociales e Igualdad, Catálogo Nacional de Hospitales 2016.

Gráfico 27 Distribución de hospitales y camas en función de los principales agentes, 2016



Fuente: elaboración propia a partir de datos del Ministerio de Sanidad, Servicios Sociales e Igualdad, Catálogo Nacional de Hospitales 2016.

Referencias de proyectos BigData en el sector salud

Luchar contra el cáncer con Inteligencia Artificial y Big Data

https://www.bbvaopenmind.com/tecnologia/inteligencia-artificial/luchar-contra-el-cancer-con-inteligencia-artificial-y-big-data/

"….. toda la experiencia de lo que le pasa al 97% de pacientes no está siendo usada ….. "

Proyecto Big Data Enfermedades Infecciosas

https://cio.economictimes.indiatimes.com/news/case-studies/apollo-hospitals-uses-big-data-analytics-to-control-hospital-acquired-infections/55023676?redirect=1

Ámbito de actuación del proyecto

ESTRATEGIA RELATIVA A LA RESISTENCIA A LOS ANTIMICROBIANOS (RAM)

En 2014, España introdujo un plan estratégico de acción de cuatro años para reducir el riesgo de resistencia a los antibióticos, siguiendo la recomendación de la Comisión Europea a los Estados miembros de la UE. Dicho plan se articula en torno a seis ámbitos de acción principales: vigilancia, investigación, prevención, control, formación y comunicación. El objetivo general es frenar la resistencia a los antimicrobianos (RAM) fomentando un uso más racional de los antibióticos. A nivel internacional, España es miembro activo de la iniciativa de programación conjunta sobre la resistencia a los antimicrobianos y de la Agenda para la seguridad sanitaria global. Ambas iniciativas pretenden coordinar los planes sanitarios para abordar la RAM.

Otros artículos de interés:

La OMS publica lista para las bacterias para las que se necesitan urgentemente nuevos medicamentos, y afirma que la única vía para resolver esta grave amenaza es I+D.

http://www.who.int/es/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

La epidemia que matará a más gente que el cáncer.

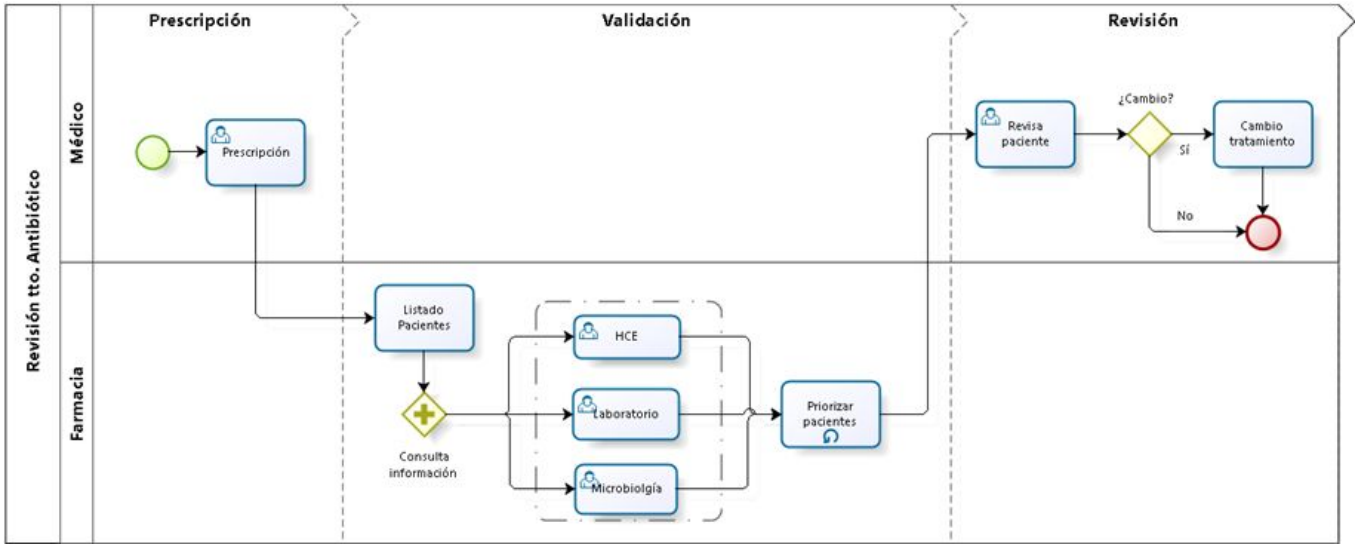
https://elpais.com/elpais/2017/09/21/planeta_futuro/1506004048_715947.html

Definición de problema

Problema

Para ponernos en contexto, se indica a continuación el circuito de validación de tratamientos de antibióticos en un hospital que tenga un equipo farmacéutico especializado en la revisión de estos tratamientos. Decir que en gran parte de hospitales no hay un equipo específico para esta importante tarea, sino que el tratamiento de antibiótico se valida del mismo modo que el resto de tratamiento farmacológico: Validado o No Validado.

Partiendo que el hospital tenga un equipo PROA (programas de optimización de uso de antimicrobianos) se muestra el flujo de trabajo del tratamiento de antibióticos



1. **Médico** realiza la prescripción: tras hacer el médico una exploración al paciente y consultar en su historia clínica electrónica ciertas constantes vitales, al identificar que tiene una infección bacteriana le prescribe un tratamiento antibiótico. Generalmente se indica uno genérico (a partir de tipo de infección), este tratamiento está protocolizado, tanto la dosis como la duración máxima del tratamiento. Mientras

tanto se solicita un estudio de microbiología para poder identificar el tipo de bacteria que ha originado la infección.

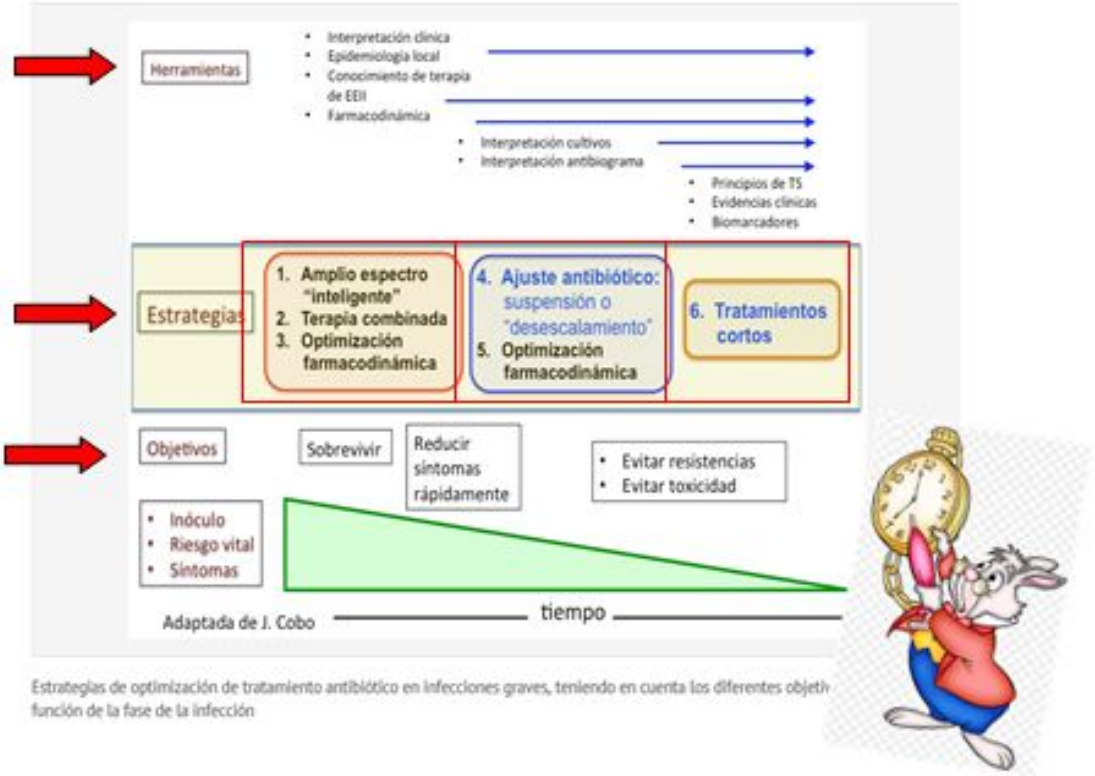
- **Problema**: Normalmente se comienza con un tratamiento antibiótico de amplio espectro (se refiere a un antibiótico que actúa contra una amplia gama de bacterias patógenas, tanto contra bacterias grampositivas como gramnegativas), al no tener aún los estudios de microbiología no se puede identificar el microorganismo en concreto, sólo el tipo de infección. Esto provoca que tomemos un exceso de este tipo de antibióticos y pueda provocar que con el tiempo seamos inmunes.

2. **Farmacia** obtiene un listado con los diferentes pacientes que tienen tratamientos farmacológicos con antibióticos. Tras hacer un primer filtro sobre los pacientes (generalmente los que identifica de mayor criticidad) entra en los diversos sistemas que tienen información clínica del paciente:
   - Historia Clínica Electrónica
     - Constantes vitales, alergias y factores de riesgo del paciente
     - Diagnósticos e intervenciones quirúrgicas
     - Antecedentes clínicos del paciente
     - Documentos médicos y notas de evolución del paciente
   - Laboratorio
     - Analíticas
     - Microbiología
   - Prescripciones farmacológicas
   - Pruebas complementarias: electrocardiograma, resonancia,..

   A partir de esta información se registra una recomendación a los pacientes analizados. De esta manera el equipo médico podrá priorizar la revisión de los tratamiento de antibióticos.

   - **Problema**: En la mayor parte de los hospitales los diversos sistemas de información no están integrados, por lo que el facultativo debe acceder a varios sistemas para poder consultar toda la información relevante del paciente, este tiempo perdido provoca que no se pueda revisar a todos los pacientes con tratamientos antibióticos (10% de pacientes ingresados).
3. **Médico revisa el tratamiento**: teniendo como punto de partida las priorización y recomendación realizada por el farmacéutico del centro hospitalaria el médico revisa el tratamiento. Para ello debe acceder de nuevo a los diferentes sistemas de información y poder medir la evolución del paciente con el actual tratamiento. Tras dicha evaluación el médico indicará el siguiente paso a seguir: continuar con el tratamiento, cambiarlo por otro o darlo por finalizado. En este punto, al tener los datos de microbiología (y por ello información de la bacteria causante de la infección) se puede cambiar el antibiótico de espectro reducido (sólo es eficaz contra familias específicas de bacterias), y por lo tanto, con menor impacto en el paciente.
   - **Problema:** El médico tiene poco tiempo para revisar la información clínica del paciente y hecha en falta un sistema integral para poder consultar toda la

información relevante para poder tomar una decisión correcta y de manera ágil.



Se indica a continuación un extracto de la entrevista que tuvimos con Maria Nuñez, farmacéutica en el hospital Virgen de las Nieves (Granada):

"Cuando el médico prescribe un fármaco afecta al paciente y lo valida el farmacéutico encargado del área (Unidosis). Sin embargo, al prescribir un antibiótico afectas a la ecología local y del propio individuo lo que se traduce en la generación de microorganismos resistentes para los que ya hoy en día se presentan ocasiones en los que no tenemos arsenal terapéutico capaz de tratarlos. Es decir, el antibiótico (ABT) que tomas para el resfriado puede hacer que en un futuro próximo, el ABT que necesites para una infección de herida no te sirva a ti o a tu vecino y pueda conllevar graves consecuencia (incluido el exitus). Por ello y por la complejidad de estos tratamientos han surgido los "PROA" programas de optimización de uso de antimicrobianos creados por un equipo multidisciplinar integrado por infectólogos, microbiólogos y farmacéuticos donde además de la clásica validación (dosis, vía de administración, alergias e interacciones fundamentales) se hacen programas de intervención específicos: seguimiento de cultivos, cambios de vía, duración de tratamientos… Se hace a día cero, a 48h y a días 5-7 de tratamiento. A estas intervenciones se conocen como "audits no impositivas" pues el objetivo es dar apoyo docente al médico prescriptor para ayudarle a tomar la mejor decisión y optimizar así el tratamiento. ***Uno de los grandes retos es que una pastilla para el corazón la prescribe sólo un cardiólogo, mientras que los ABT es cosa de todos y no siempre con la formación adecuada***"

Justo en esta última frase se identifica uno de los problema de la sobremedicación de antibióticos, es recetado por personal clínico no especializado, lo que provoca que se prescriba tratamientos antibióticos de manera 'estándar' sin tener en cuenta las particularidades de cada paciente, lo cual provoca, en gran parte de los casos, un exceso en la duración de los tratamientos antibióticos, teniendo las siguientes afecciones:

- Afecta al sistema inmune del paciente
- Efectos secundarios para el paciente
- Económicos para el centro hospitalario
- Impacto ecológico en el centro hospitalario.

Entrevistas a clientes

<u>Médico especialista en tratamiento de enfermedades infecciosas en Virgen Macarena (Sevilla)</u>

El circuito que hacemos a diario es ver paciente, conciliación de medicación, entrar en el programa de prescripciones y consultar/modificar el tratamiento farmacológico. Cada día se verifica o modifica. El aplicativo de prescripciones no está conectado con la Historia Clínica Electrónica (HCE), lo cual provoca: duplicidad, pérdida de tiempo y de información. Si bien un sistema de recomendación farmacológica sería muy interesante, el consejo sería centrarnos en enfermedades infecciosas.

En estos casos una misma patología puede tener diferentes tratamientos, algunos sensibles y otros resistentes. Deberían cruzarse los datos del programa de fármaco con el antibiograma y mostrar alertas, en caso de ser necesario.

Del mismo modo podría relacionarse con los datos de constantes vitales: frecuencia cardíaca, temperatura, tensión, arritmias, etc. Si estuviera unido al programa de tratamiento

podría tener alertas sobre constantes alternadas. Otra información relevante serían los factores de riesgo (antibióticos previos, diabetes, ingresos anteriores…)

Actualmente el tiempo de diagnosticar a un paciente es de 30’ el 1er día y 15’ los sucesivos, incluyendo visita al cliente.

¿Centralización de la información en una especie de Cuadro de Mandos Integral del Paciente o Visión 360, donde nos dé información y evolución incluyendo tratamiento de información no estructurada?

Farmacia con 2 médicos repasa 2 informes diarios de antibióticos restringidos y que llevan +7 días puestos en la medicación del paciente. ¿recomendación por el propio sistema de manera automática? ¿y si detectamos INFECCIÓN en los documentos no estructurados o en base a los medicamentos prescritos?

Los objetivos de la revisión de la prescripción son económicos, reducción del impacto ecológico, para evitar la disminuir resistencia de los gérmenes, y controlar los efectos secundarios en el paciente.

La medida de la efectividad de la prescripción se hace midiendo la mortalidad (sobre el 20%), y en qué día se produce. Si no hay mortalidad, se mide el número de días qué hay curación clínica (cuando desaparecen los síntomas de la enfermedad).

Otra fuente de datos importante puede ser la anatomía patológica (biopsias por ej.), que influye en la prescripción, aunque no siempre están disponibles o de fácil acceso.

Farmaceútica especialista en tratamiento de enfermedades infecciosas en Hospital de Virgen de las Nieves (Granada).

HOSPITAL UNIVERSITARIO VIRGEN DE LAS NIEVES
HOSPITALES DE GRANADA

La recomendación sería poner en foco en los tratamientos de antibióticos.

Existe una queja generalizada de la falta de integración en los sistemas. Debe entrar en varios para hacer su trabajo. En muchos casos la información no es clara y está en texto libre.

El programa de farmacia es diferente por hospital (APD software con tendencia a centralizar por el SAS, ellos siguen usando Farmatool). No hay previsto cambiar Farmatool a APD, pues APD genera quejas. También se utiliza para prescribir además de validar.

Hay que tener especial atención con fármacos de estrecho margen terapéutico. “Poco no hace nada. Mucho puede matar”.

Para determinados fármacos por seguridad o eficacia hace falta medir su nivel en sangre y estimar, basándonos en modelos poblacionales (bayesianos) la mejor dosis individualizada.

Para ello existen complejos y costosos programas que nadie usa (NONMEN creo que es la excepción), calculadoras on line (ej http://www.globalrph.com/vanco_single.htm) o lo que todo el mundo hace en españa: PKS de Abbott, un programa obsoleto en MSDOS nada friendly. Se que hay un grupo de la SEFH PKGEn que lleva años diciendo que iban a intentar sacar algo parecido versión moderna pero seguimos esperando.

Para poder **decidir** qué candidatos debe ver físicamente, se puede llevar 2 horas para revisar la documentación de 15 pacientes y poder priorizar según ésta. Es un trabajo de screening. El protocolo de revisión (tiempos) es claro y conocido. **¿Sistema de Alerta Temprana?**

- Foco 1. Optimización de tiempos.
- Foco 2. No sobre medicación. Economía tanto en costes y medio ambiente.
- Foco 3. Salvar vidas.

ZPD Compass. Para antibióticos. Aplicación con base de histórico anterior y proponer soluciones. Actualización en tiempo real con la bbdd de microbiología y los datos del propio paciente. Probabilidad de éxito con diferentes alternativas.

Interesante: Para realizar estudios, el SAS te puede facilitar bases de datos de pacientes con históricos. CRD cuaderno de recogida de datos. Se hace petición a un comité de ética y si es aprobada los datos se facilitan anonimizados.

Los datos a nivel ambulatorio están todos compartidos dentro de la red del SAS, pero a nivel hospitalario, aunque tengan al paciente dado de alta en su hospital no pueden consultar toda la historia generada en los otros hospitales (puede ser un tema de no compartición os implemente de permisos de acceso por el farmacéutico).

La evolución de los pacientes se suele medir mediante indicadores como la fiebre, valor de leucocitos, lo que diga el médico en la revisión del paciente (aunque se escribe poco y hay días que ni se ve al paciente).

Lleva más de 7 años en el área y al preguntarme qué haría la vida de los PROA más fácil lo tiene clarísimo: Una herramienta que sea capaz de capturar la info que necesito, en la forma que necesito, en tiempo real de los programas existentes y me de posibles soluciones al problema ajustadas a mi paciente, para ello:

1. Seleccione de la historia clínica una serie de parámetros preseleccionados (ojo no siempre completo)
2. Los cruce con los programas de farmacia y laboratorio
3. Se cruce con la guía de tratamiento empírico locales (que a su vez se actualice con informes anuales de sensibilidad)
4. Muestre el resultado más óptimo para mi paciente
5. Se creen repositorios uniformes de toda la info generada para estudios de intervención

Los sistema de información clínicos que hace uso diariamente para ver la evolución de un paciente son:

- Farmacia: farmatools o APD
- HCE: Diraya
- Laboratorio: MODULAB

La idea sería capturar de diraya, programas de farmacia y laboratorio pacientes candidatos y ayudarme a generar la mejor recomendación y generar repositorios de información. Este tema también está en auge por ser medicina personalizada, ahorrar costes, seguridad del paciente, etc, etc… ventaja, se basa en menos datos y más objetivos.

Farmacéutico con enfoque IT - Hospital de Puerto Real (Cádiz)

No existe flujo de información entre la aplicación de farmacia y datos de laboratorios, no solo para antibióticos, sino también para insuficiencia renal, etc.

Una parte clínica que se escapa en informática, es diagnóstica. Campo de texto libre donde el médico apunta el diagnóstico (juicio clínico). Esto lo apunta el médico a pie de cama del paciente y no es accesible para el farmacólogo, debe hacerlo mirando los papeles del médico previamente en la historia clínica del paciente (Diraya). Hasta que no se cierra el caso clínico no documentan los médicos por lo que deben ver las historias en papel de los médicos mientras están hospitalizados.
La definición de antibióticos en farmacia no es igual que en microbiología, hay que normalizarlos para que hablen igual.
En torno a 50 principios activos es el orden de magnitud que se maneja. Para cada diagnóstico hay que buscar el modelo adecuado que identifique el tratamiento más adecuado.

Concepto: según diagnóstico + microbiología antibiograma → Recomendación del siguiente paso más adecuado:

- cambio de antibiótico por uno más idóneo
- confirmación del antibiótico actual
- supresión del antibiótico

Mapa de resistencia → es propia de cada hospital, indica a que suele ser resistente el entorno de un hospital. Como punto de partida para construir algoritmo. Estos mapas de resistencia suelen estar disponibles en Internet o en cada Hospital.
El problema es el paso siguiente, tras ver ese mapa, tienes un paciente y debes evaluar según cada historia clínica para decidirlo. Esa reunión se realiza por las mañanas con el equipo completo PROA (farmacólogo, médico especialista, médico infeccioso, médico

internista, cirujano …) y deciden qué hacer. Esto habría que ver cómo podemos "automatizarlo" y que todos se sientan involucrados.
Para priorizar se evalúa los microorganismos del paciente en vez del tipo de infección, ya que hay algunos más "peligrosos" que otros por miedo a propagarse. En este caso, se prioriza a ese paciente, aunque no sea tan crítico como otro. Esto nos servirá para saber cómo tenemos que ponderar los antibióticos, bien por los microorganismos que ataquen o bien por tema económico.
La guía de cabecera para el PROA es el siguiente libro: **Guía de terapéutica antimicrobiana** de Josep Mensa (no está informatizado): Tipos de infecciones + primera elección + segunda infección. Este libro es como una guía de árbol de decisión, te explica qué hacer en primer paso, en segundo paso si el primero no fue bueno, … En base a datos de microbiología e historia clínica.

Solución Funcional

Fase 1: Visión 360º del paciente (Analítica de datos)

Se pretende dar acceso a toda la información de los pacientes desde un único Dashboard o cuadro de mando. El objetivo de este Cuadro de Mando es unificar e integrar toda la información del paciente, que está distribuida en diferentes bases de datos (silos de información) y en cuya recopilación manual se tarda varias horas al día, para que asista y ayude a priorizar de una forma rápida y ágil cuáles son las mejores decisiones a tomar. Permitirá lanzar alarmas según el tipo de infección y priorización de casos. A partir de ahora cuando hagamos referencia a éste lo llamaremos visión 360º del paciente.

La visión 360º del paciente construiría un primer producto que ayudaría a ahorrar una gran cantidad de trabajo diario derivada de la tarea de recopilación de información. La información que mostraremos procederá de los siguientes sistemas de información:

- Historia Clínica Electrónica:
    - Datos del paciente: edad, sexo, nacionalidad,...
    - Datos clínicos relevantes: alergias, factores de riesgo, antecedentes relacionados, diagnósticos, procedimientos realizados en el hospital e intervenciones quirúrgicas.
- Prescripción farmacológica electrónica:

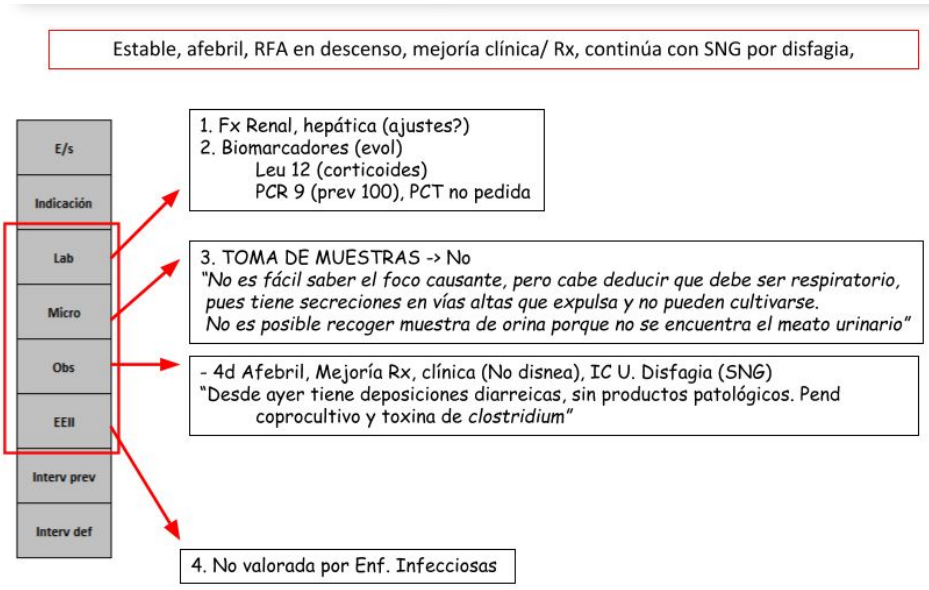
- Prescripciones del paciente, tanto las actuales como el histórico, incluyendo la dosis, principios activos, vías de administración y duración del tratamiento.
- Validación de los tratamientos farmacológicos por parte de la farmacia hospitalaria.
- Se obtendrá tanto la medicación prescrita en el centro hospitalaria como la que se haya administrado en su domicilio.

- Laboratorio:
  - Analíticas convencionales
  - Microbiología: con el resultado de los análisis y cultivos de los microorganismos estudiados.

Fase 2: Sistema de Recomendación

Una vez implementada la Fase 1 se realizaría el tratamiento de la información cruzando todos los datos recopilados de manera que el sistema pueda aprender de ellos y sea capaz de realizar recomendaciones de prescripciones óptimas (dosis, tratamiento, retirada, tiempo de continuidad…), reconociendo pacientes de características similares.

El sistema puntuará las recomendaciones de las prescripciones propuesta, dando a conocer al usuario la probabilidad de éxito de cada una de ellas, ayudando así al farmacéutico a la toma de decisiones y validación de tratamientos.



ncama	NHC	apelli	UGC	Fármaco	días	E/s	Indicación	Lab	Micro	Obs	EEII	Interv prev	Interv def
723-3	[redacted]	XXX	MIG	p/T 4/8	8	90M	Neum ¿asp? NOS	Leu 12 (c). PCR 3 (Prev 100). Cr OK	No	FQ prev. SNG	No		
530-3		XXX	MIG	A/C 2/8	11	85M	Neum necrot ¿empiema?	Leu N PCR 210 (d)	HC S pneum	2m prev: Pancre Ag Comp, quistes renales/hep con aislam K.Ox (AMC). NO TS (dieta trit). AF pend Drenaje + control Rx	Sí		
371-1		XXX	CARD	A/C 875/8	3	74H	? COM	Leu 7,5, PCR 4	No	Ing:Disnea. AF. Rx No condens, LOES, signos Edema ni derrame. En estudio por FA/anemia (Transf)	No		
527-3		XXX	MIG	CEFTAZ 1/8	7	82H	EPOC Reag bronquiectasias COM	Leu 9,5, pcr 162 (Prev 88)	Esputo: Ps R-cipro	Estable. Abt (AMC/FQ) + ingresos previos	No		
545-5		XXX	INF	Rimstar + Moxi	7	25M	TBC pulmonar	X	X	**Omeprazol**	Sí		
829-2		XXX	MIG	A/C 1/8	7	83M	Bronquitis	leu 14, PCR 2. **Cr2**	No	AF.	No		
524-1		XXX	MIG	CFTRX 2g/24 + metro IV	13	80M	Absceso lab	leu 6, PCR 23 Cr+Pr ok	Abc: KN (R a A/C y P/T) (Prev:KN, EC y S angin)	Previo 4d de P/T. Ca de pared Abd. Penrose drenanado. AF y pdte cirugía.	Sí		
847-2		XXX	DIG	CEFTAZ 2/8	13	56M	Colangitis aguda	Leu N , pcr 14	Bilis: Ps-MR	Prev P/T.AF.Drenaje biliar. Prog REPETIR CPRE en 2d	No		
728-3		XXX	MIG	A/C 1/8	8	87H	Colecistitis Ag litiasis COM	leu 14 (s) PCR 225 (d)	No	Pendiente CPRE	No		
246-3		XXX	TRAU	CIPRO 400/12	8	56H	Celulitis COM	leu N, pcr 76	Ex Mixta	**AF.**	Sí		
250-3		XXX	TRAU	**P/T 2/8**	7	70H	Pie diabético	Leu 13, PCR 115, **Cr 0,8**	No	MEG, amput	No		
356-2		XXX	CIR	CLOXA **2/4**	3	56H	BM SAMS	Leu N, pcr 30 (d). **IH, Cr ok**	HC neg (prev SAMS)	AF,Catéter retirado. CHILD C (NO septrim, clinda, line)	Sí		
528-3		XXX	MIG	Ceftriax 1/24	8	63M	PNF	leu N, pcr 76. Cr OK	HC, UC E Coli S-cipro, AMC	Sepsis grave. SU retirado. fTUs rep. (Visto por ERT). REG/febrícula	No		
540-1		XXX	INF	Septrim L-X_V	7	54H	Profilaxis (VIH)	CD4 <100	X	TAR - tto dom **OK**	Sí		

El sistema de Recomendación constituye el segundo producto, que se comercializará como una solución enfocada tanto a hospitales públicos como privados, y que contará con diversas release en las que se irá incrementando el alcance de la solución.

Retroalimentación del modelo / Feedback

Existen múltiples indicadores que nos permiten validar el correcto funcionamiento del modelo prescriptivos (tasa de consumo de antibióticos, tiempo de suministro de tratamiento, menor estancia en hospital de paciente con tratamiento antibiótico, menor contagio por ecosistema bacteriano del hospital, etc.)

Competencia

El clinical checking ofrecido por **Vademecum** se integra con la historia clínica del paciente y muestre un nivel de alerta con respecto al tratamiento.

Los valores que tiene en cuenta para consultar las interacciones son:

- Peso, estatura y edad
- Alergias, diagnósticos y factores de riesgo
- Valores de laboratorio y prescripciones activas

A partir de dichos valores nos indica por cada prescripción activa del paciente sus alertas. Priorizando por nivel de severidad.

http://es.vidal-dis.com/

Los médicos pueden buscar rápidamente posibles interacciones entre medicamentos o entre medicamentos y hierbas con el programa de interacciones de medicamentos de **UpToDate** (en asociación con Lexicomp®).

- Contenido basado en evidencia: más de 11.000 temas clínicos y 25 especialidades, con más de 9.500 recomendaciones valoradas para tratamientos
- Búsqueda de gráficos: permite la búsqueda y la visualización rápidas de más de 33.000 imágenes, tablas, ilustraciones, diagramas, gráficos, algoritmos y vídeos bien detallados.
- Calculadoras médicas: más de 175 herramientas de confianza que le ayudarán a diagnosticar y tratar pacientes sin tener que recurrir a fórmulas complejas.
- Bases de datos de medicamentos y de interacción de los mismos (Lexicomp®): información completa sobre cuestiones de seguridad, dosis, advertencias e interacciones (medicamento-medicamento y medicamento-planta medicinal)

https://www.uptodate.com/es/home/product

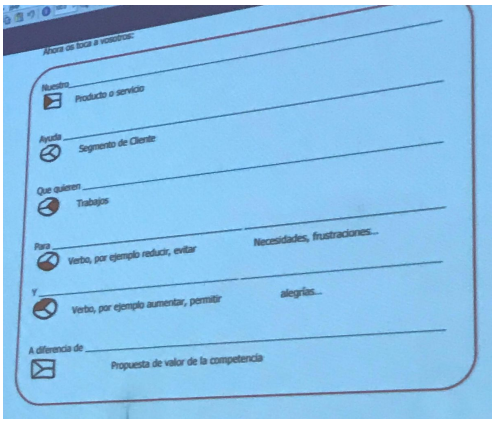
Consigue con técnicas big data mayor precisión que cualquier escala de riesgo clásica. Personaliza el riesgo de cada paciente para la toma de una decisión clínica. Genera un modelo predictivo en muy poco tiempo en base a lo recogido en las historias clínicas.

Entre sus aplicaciones más relevantes se encuentran:

- La evaluación de resultados en salud (muy útil para la gestión de los acuerdos de riesgo compartido en fármacos de alto impacto presupuestario)
- La gestión de la calidad (adherencia a las recomendaciones de práctica clínica)
- Descripción de poblaciones, conectando fuentes de información habitualmente desconectadas (impacto en atención especializada de decisiones tomadas en primaria y viceversa).

https://www.savanamed.com/savana-predict.php

¿En qué somos diferentes?

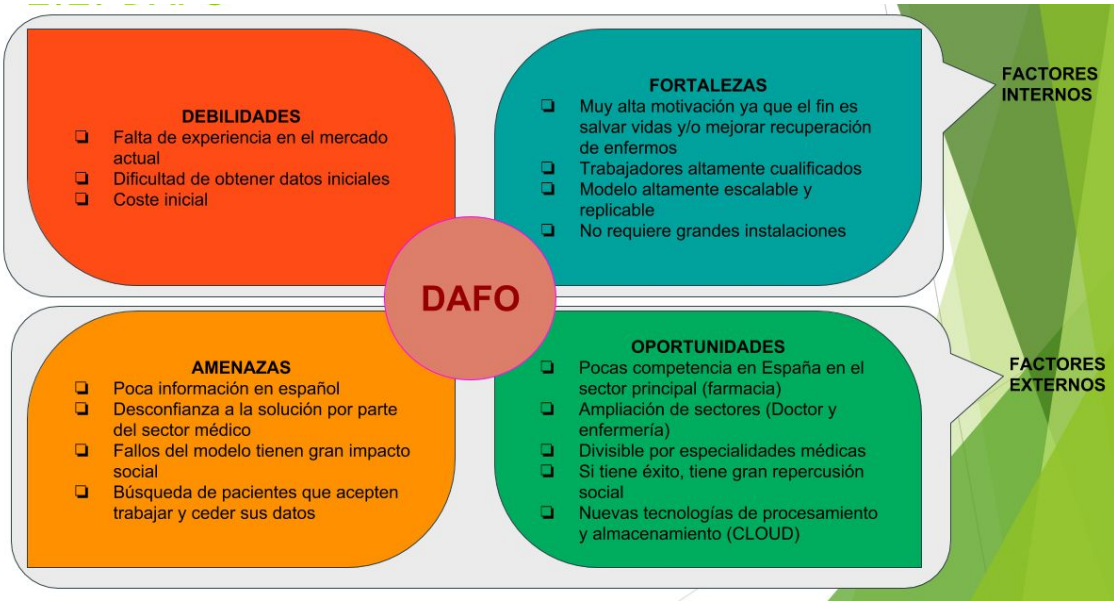


El Servicio sobre Prescripciones Antibióticas, mediante el análisis de modelos predictivos, permite al equipo médico hospitalarios adecuar los tratamientos de manera personalizada garantizando el mayor grado de eficacia de los fármacos con el menor tiempo de tratamiento, a diferencia de la extracción y cotejo de datos manual que actualmente existe en los hospitales para este tipo de valoración en las prescripciones antibióticas.

1. Especializados en tratamiento farmacológico con antibióticos
2. Tiene en cuenta, no sólo la información actual del paciente sino su histórico y el de otros pacientes con características similares.

DAFO

Se trata de un mapa a través del que se establecen las debilidades, amenazas, fortalezas y oportunidades de la organización. Un análisis interno y externo del entorno en el que se desarrolla la actividad para mejorar su rentabilidad, funcionamiento y posicionamiento en el mercado.



Modelo Normativo

En este proyecto nos afecta el marco normativo en cuanto a LOPD ya que parte de la solución pasa por guardar y analizar datos que podrían ser considerados sensibles. Teniendo en cuenta esto, debemos de prestar especial atención a la actualización de la LOPD que entrará en vigor el 25 de Mayo de 2018 y cuyo principal cambio a tener en cuenta es que tendremos que tener consentimiento explícito por parte de los clientes a la hora de guardar sus datos de carácter personal e informar del uso que le vamos a dar a esos datos.

Existen formas auditables, estandarizadas y controladas de gestionar datos sanitarios con total respeto a la privacidad de los pacientes. Por esta razón, en la actualidad ya hay proyectos europeos de Big Data, con gran alcance, en marcha. En estos, la solicitud del consentimiento informado a millones de pacientes de forma retrospectiva resulta imposible, por lo que deben idearse formas éticas para hacer avanzar la investigación y la gestión sanitarias, sin vulnerar derechos.

Los pasos a seguir son:

- **Anonimización y disociación**: Los algoritmos de anonimización son técnicamente sencillos de desarrollar y desproveer a los datos de su carácter personal. Los reglamentos de protección de datos exigen que la anonimización sea “razonable”; es

decir, no es exigible asegurar un nivel de anonimización absoluto, sabiendo que siempre existirá un riesgo despreciable de que puedan identificarse las peculiares características de un paciente por sus descripciones clínicas. Se cumple con los 2 puntos básicos en la anonimización:

  - Es un proceso irreversible.
  - No existe un soporte que permita relacionar el registro original con el anonimizado.

- **Agregación**: En el momento en que la información es agregada, deja de ser individualmente relevante y por tanto no pertenece a un paciente concreto. Lo que se reutiliza entonces no es la información (la cual pertenece al paciente), sino la propiedad intelectual (la cual pertenece a la institución). En realidad es el mismo procedimiento por el que se escriben y editan los libros de Medicina, pero llevados a un nivel de actualización y navegabilidad mayores.
- **Ruido**: Es recomendable introducir algo de ruido controlado en los datos, de forma que haga la base de datos inservible ante un eventual ataque malintencionado de un tercero.

Modelo de Negocio

Definición del reto

Mario es Farmacéutico y pertenece al equipo médico del Hospital donde trabaja. El, junto a otros 3 compañeros, otra farmacéutica y 2 médicos, forman el equipo PROA. Se dedican a revisar las historias clínicas con tratamiento aplicado.

¿Su misión?

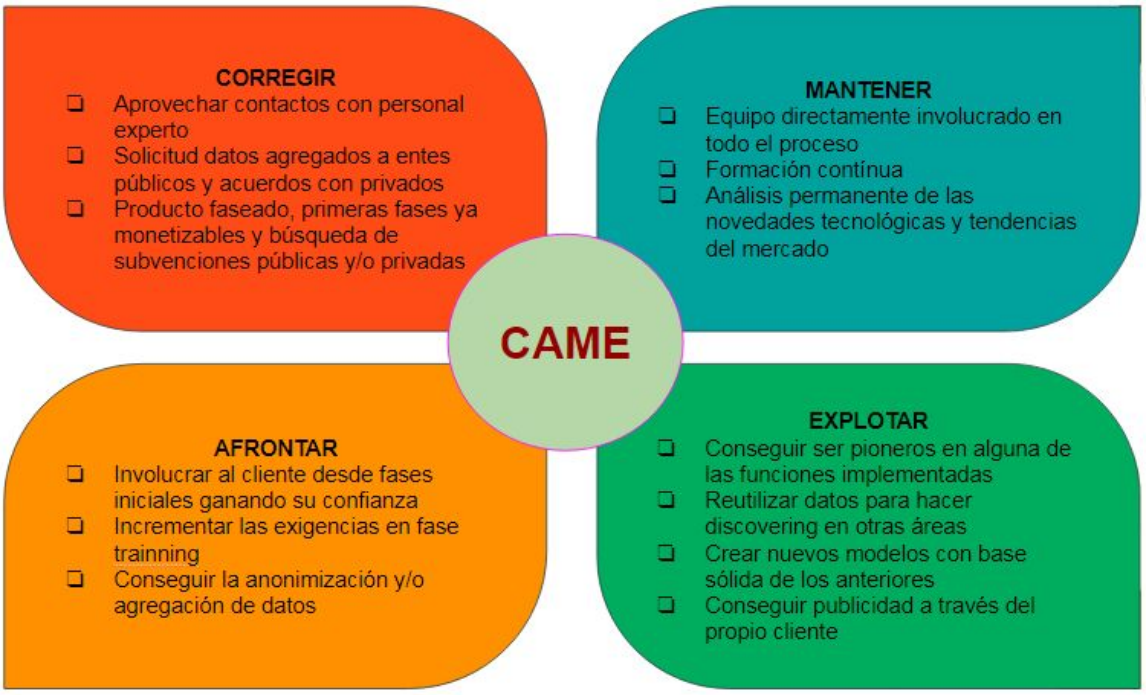
- Priorizar a quiénes deben visitar los médicos. (Seguridad del Paciente / Eficiencia & Eficacia)
- Revisar el correcto funcionamiento del tratamiento aplicado y tomar decisiones. (Seguridad del Paciente / Medicina personalizada / Ahorro de Costes)
- Luchar contra la medicación que crea enfermedades resistentes a los medicamentos. (Seguridad del Paciente / Medicina personalizada / Ahorro de costes futuros / Propagación de enfermedades)
- Economizar en el impacto medioambiental. (Sostenibilidad / Ahorro de Costes)
- Y lo más importante, "Salvar vidas".

Mapa de empatía

Conociendo el problema de nuestro amigo Mario, le realizamos su mapa de empatía comprendiendo a lo que diariamente se enfrente en su jornada laboral.

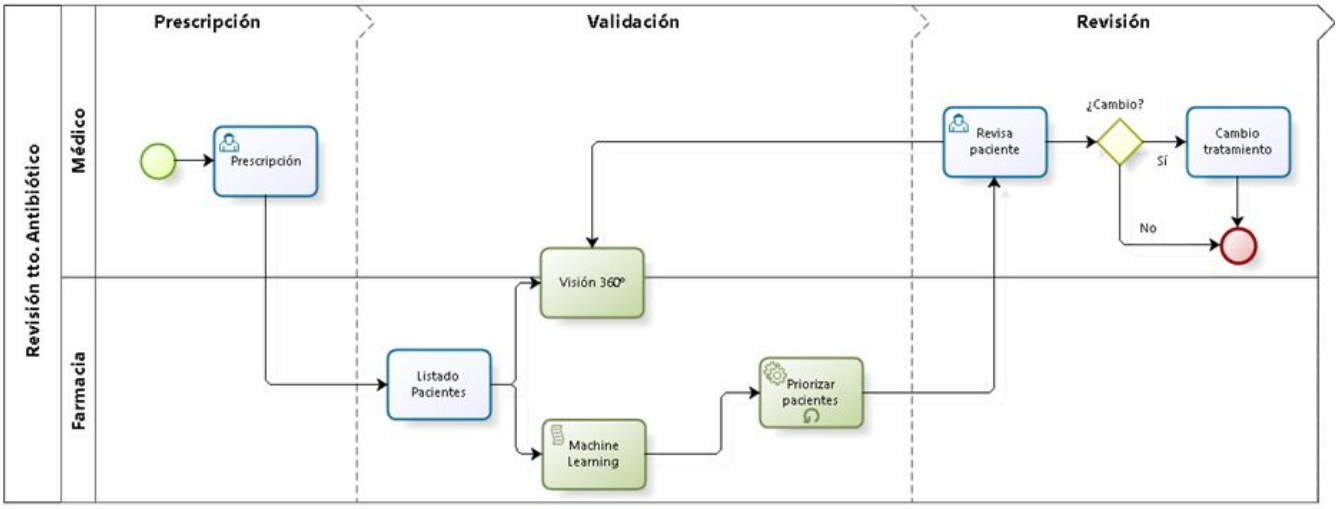
Análisis CAME

Tras valorar la situación inicial con el análisis DAFO, procedemos a realizar el contraanálisis CAME para obtener las acciones resultantes sobre las que debemos guiarnos.



Integración del Caso de Uso con los Procesos de Negocio

Anteriormente se mostró cual era el proceso que actualmente poseen nuestros clientes, pues tras definir las acciones que realizamos y la solución aportada, se muestra cómo la integramos en el proceso de negocio de nuestros clientes. En otras palabras, ¿cómo podemos ayudar a Mario?



En la figura anterior vemos cómo nuestra solución (en verde) se integra en la fase de *Validación - Evaluación* del tratamiento del paciente. El proyecto, como ya hemos comentado, es un producto software que posee varias versiones.

En la ***primera versión*** aportamos esa ***visión 360º del paciente*** que elimina todos los tiempos empleados en cuanto a recopilación de datos de distintas fuentes disponibles en los sistemas actuales del hospital (datos del paciente + diagnóstico + antibiograma + monitorización + etc). Esta fase la realizamos a modo de consultoría junto con el equipo PROA del hospital, generando un cuadro de mando embebido en nuestra aplicación con la herramienta *Tableau*.

Desde el primer momento comenzamos a monetizar el proyecto gracias a la integración de datos que le realizamos al hospital. Este cuadro de mando es con el que el equipo PROA comenzará a trabajar para:

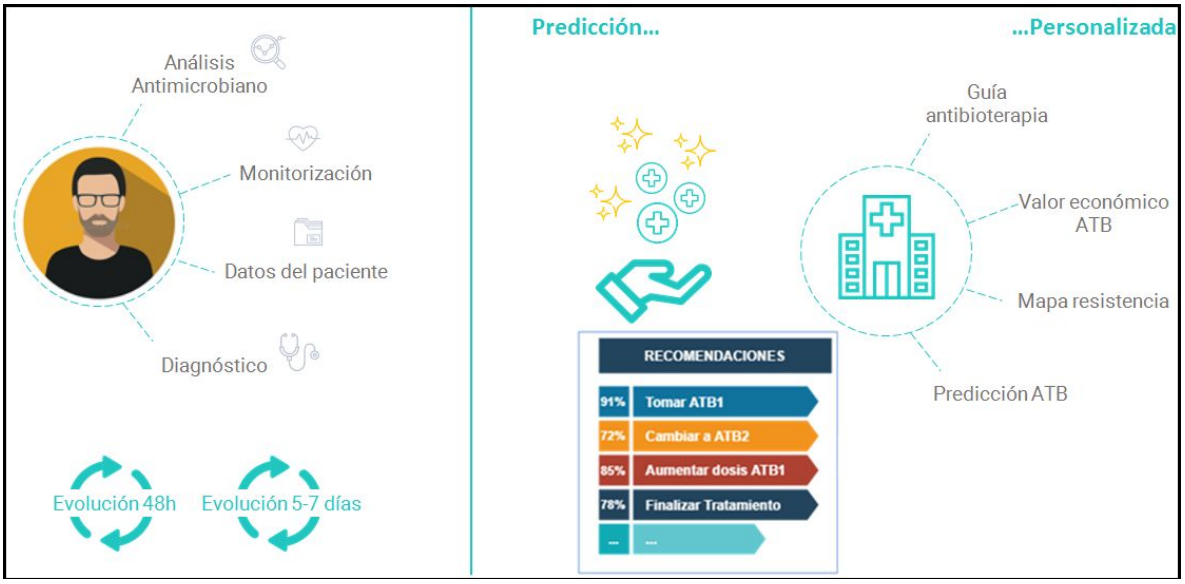
- priorizar la visita de los médicos a los pacientes.
- revisión de los tratamientos ATB actualmente en curso.

En la ***segunda release*** incluimos a la herramienta, el ***sistema recomendador***, realizado con el *proceso de machine learning*. Nuestra solución analizará la evolución del paciente (análisis antimicrobiano + monitorización + datos del paciente + diagnóstico + etc) dando una nueva recomendación en base de probabilidad que podrá ser:

- continuar con el tratamiento (el tratamiento propuesto es el mismo)
- modificación del tratamiento (propone nuevo tratamiento)
- finalización del tratamiento (el tratamiento propuesto ha sido efectivo)

La llamada al proceso de machine learning será descrita más adelante en el apartado de solución técnica, pero se avanza que será una llamada del tipo API REST en cloud.

Es en este momento dónde más valor añadido aportamos a nuestro cliente, personalizando el sistema recomendador al propio hospital mediante el peso asignado a cada variable de configuración al sistema.

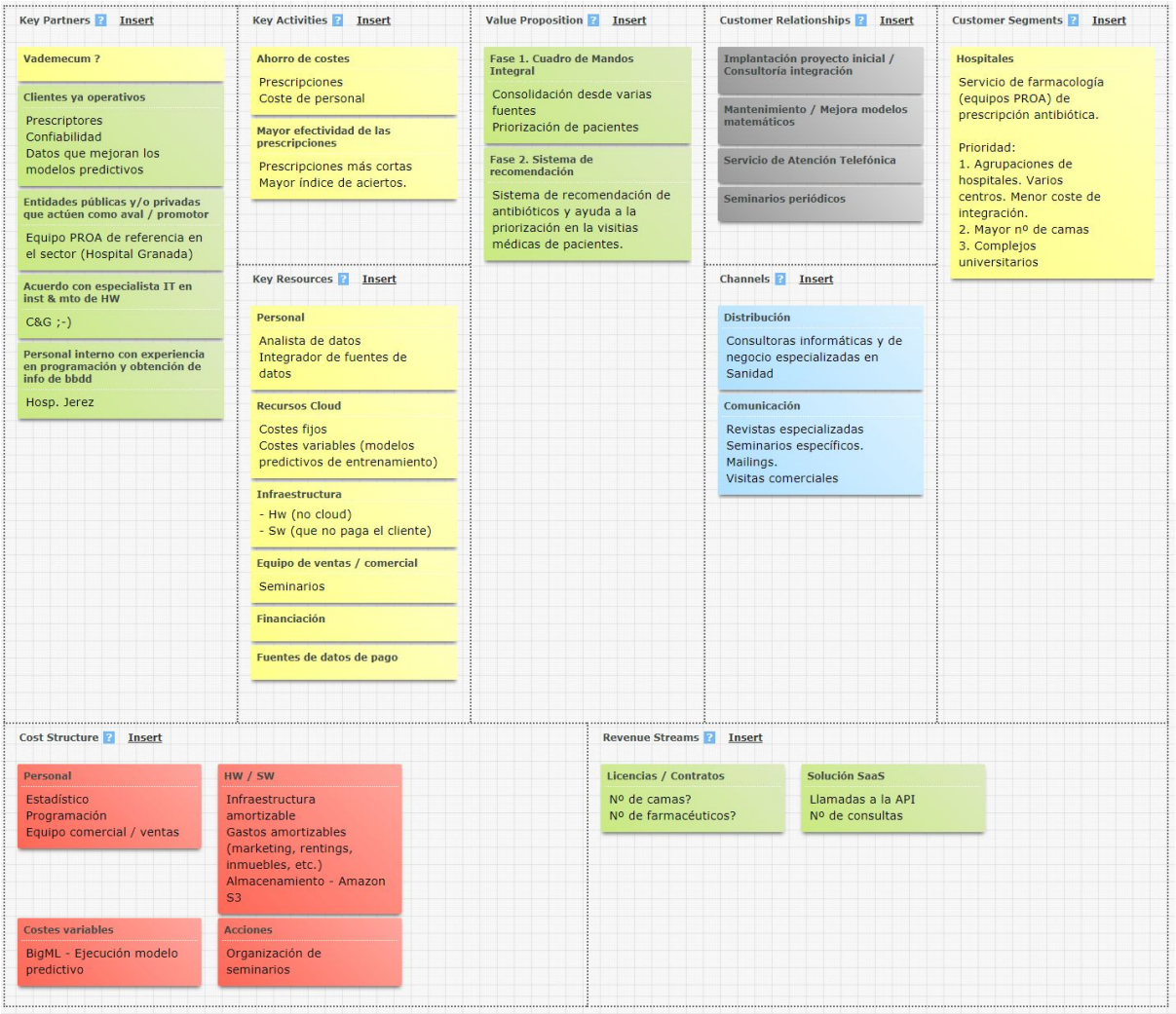


Dichas variables forman parte de las entradas de nuestro modelo por lo que más adelante en la solución técnica se explican su comportamiento.

Modelo Canvas

Una vez descrito el problema, Para realizar el modelo de negocio hemos aplicado la metodología canvas de manera gráfica mediante la herramienta de https://canvanizer.com. Esta metodología divide el análisis del modelo de negocio en 9 bloques repartidos desde dos puntos de vista; los aspectos externos de la empresa, mercado y entorno (parte derecha del canvas) y los aspectos internos de la propia empresa (parte izquierda del canvas).

Hemos comenzado con los bloques que componen la parte derecha del modelo y luego la parte izquierda. A continuación se muestra el resultado obtenido de la sesión de desarrollo del canvas y la descripción de cada bloque individual.



Definimos cada bloque a continuación en el orden en que lo hemos definido:

Segmentos de Clientes:

Equipos del Programa de Optimización de uso de Antimicrobianos (Equipos PROA) de:

- Cadenas de Hospitales privados: Quirón, VIAMED, etc.
- Consejerías de Salud Pública Autonómicas: Andalucía SAS, Cataluña (ASPCAT), etc.

Propuesta de valor:

El Servicio sobre Prescripciones Antibióticas, mediante el análisis de modelos predictivos, permite al equipo médico hospitalarios adecuar los tratamientos de manera personalizada garantizando el mayor grado de eficacia de los fármacos con el menor tiempo de tratamiento, a diferencia de la extracción y cotejo de datos manual que actualmente existe en los hospitales para este tipo de valoración en las prescripciones antibióticas.

Canales:

- Distribución:
    - Consultoras informáticas y de negocio especializadas en Sanidad.
- Comunicación:
    - Revistas científicas especializadas.
    - Seminarios específicos.
    - Mailings.
    - Visitas comerciales.
    - Web corporativa.

Relaciones con clientes:

- Implantación del proyecto inicial / Consultoría de integración.
- Servicio de atención telefónica.
- Seminarios periódicos.
- Mantenimiento / Mejora de los modelos matemáticos.

Fuentes de ingresos:

- Licencias de uso.
- Contratos de servicios.
- Llamadas a la API.
- Subvenciones públicas autonómicas.
- Subvenciones públicas europeas.

Recursos clave:

- Personal:
    - Consultor PROA.
    - Analista para la integración en hospitales.
    - Analistas técnicos y programadores.
- Infraestructura software y hardware.
- Recursos cloud, descritos en la solución tecnológica.
- Oficina de trabajo.
- Fuentes de datos privadas.
- Financiación.
- Formación.

Actividades clave:

- Incremento de la efectividad en las prescripciones aporta:
    - Prescripciones más cortas.
    - Mayor índice de aciertos.
- Ahorro de costes:
    - Personal.
    - Consumo de antibióticos.

Socios clave:

- Clientes ya operativos:
    - Prescriptores.
    - Confiabilidad.
    - Datos que retroalimentan el modelo para su mejora.
- Clientes consultores:
    - Equipo PROA de referencia en el sector. Hospital de Granada.
    - Farmacéutico con experiencia en integración de datos, Hospital de Jerez.
- Vademecum, fuente de datos privada de interacciones de medicamentos.
- Proveedores IT (instalación y mantenimiento HW).

Estructura de costes:

- Recursos Humanos:
    - Estadístico.
    - Técnicos desarrolladores de software.
    - Equipo comercial / ventas.
- Infraestructura (Hw y Sw).
- Recursos cloud.
- Costes variables:
    - BigML, ejecución del modelo predictivo.
- Acciones, organización de seminarios.

Análisis de riesgos de la empresa.

1) El cliente no esté dispuesto a pagar por nuestra solución.

   Tenemos previsto pivotar a un otros ámbitos diferentes al infeccioso, en el que se necesiten los mismos inputs del paciente (historia clínica, constantes vitales, test genético, etc), como podrían ser enfermedades víricas, oncología, etc.

2) No sepamos diferenciarnos de la competencia

   Puede darse el caso de que la competencia comience a ofrecer las mismas soluciones que las nuestras, lo que nos quite cuota de mercado y clientes potenciales. Para minimizar este riesgo potenciaremos exclusividad en clientes clave, de forma que la competencia no pueda ofrecer el mismo servicio a un cliente ya captado por nuestra empresa

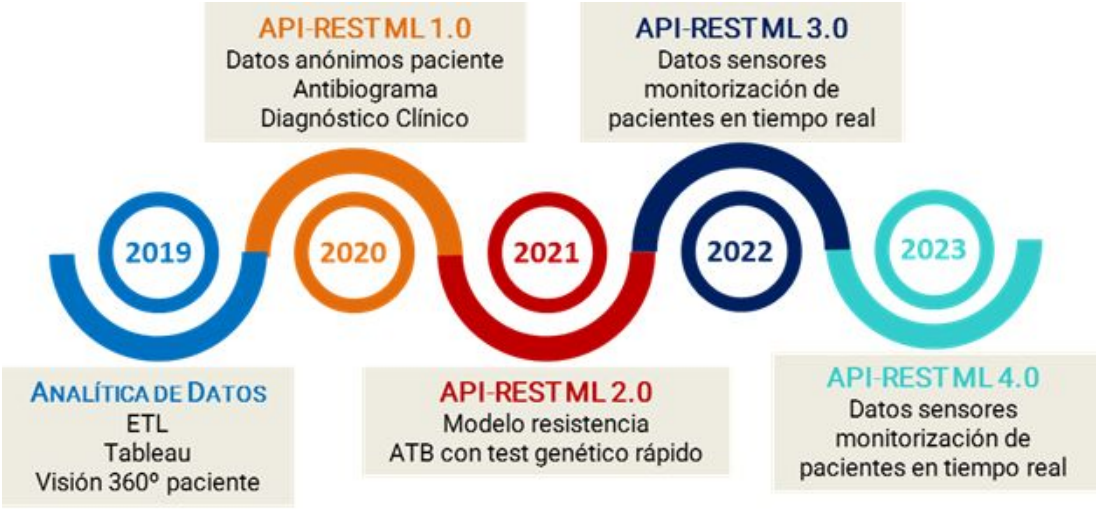
3) Lleguemos tarde al mercado.

   Nos hemos marcado un objetivo ambicioso para evitar este riesgo: conseguir el primer recomendador en el mes 18 del proyecto.

4) Largos plazos de la administración pública en la contratación..

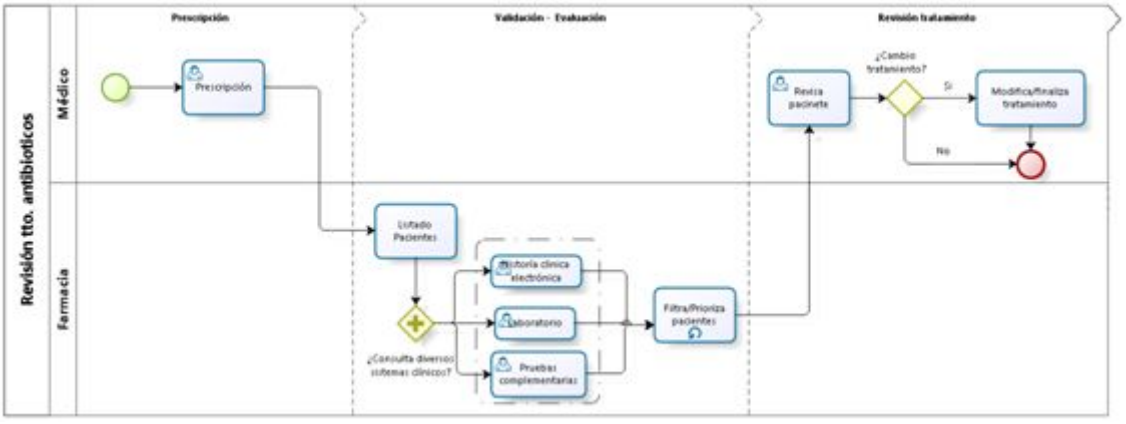
Aunque los Hospitales de las redes públicas son interesantes por sus grandes bancos de datos, los largos plazos que manejan para la toma de decisiones de contratación nos inclinan a trabajar con los hospitales de redes privadas.

Solución técnica

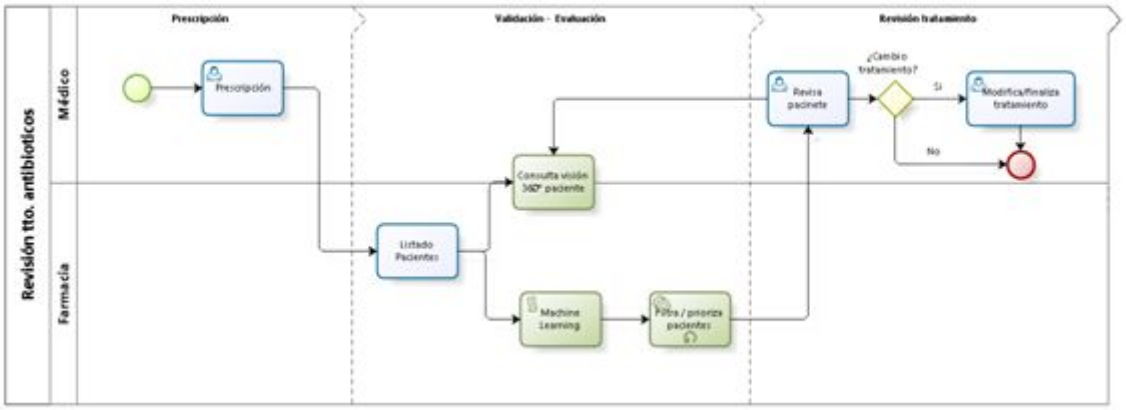


Como hemos comentado, nuestro proyecto se divide en dos productos o módulos bien diferenciadas, el primer producto se basa en la analítica de datos y la necesitaremos para obtener datos reales con los que poder desarrollar el segundo producto del proyecto, modelos predictivos. Para una correcta comprensión de donde se integran nuestras herramientas se vuelven a mostrar, a modo resumen, los diagramas de integración de los casos de uso con los procesos de negocio.

Situación Actual



Nuevo proceso de Negocio



El primer producto consiste en la creación de un Cuadro de Mando que recoja todos los parámetros relevantes del paciente hospitalizado de una manera visual y sencilla para el usuario.

Una segunda fase mucho más extensa en tiempo, consiste en la creación de un recomendador de antibióticos consultable a través de una API REST en cloud. El timeline para los modelos tienen sentido entendiendo los procesos de negocio dentro del Hospital.

El doctor prescribe un ATB de amplio espectro en base a sus conocimientos, datos del paciente y su diagnóstico, pero sin saber exactamente de qué microorganismo se trata hasta que llega el antibiograma (unos dos días).

Una vez llega el antibiograma el equipo PROA: según diagnóstico + microbiología antibiograma + datos del paciente + monitorización + mapa resistencia ATB dentro del hospital + peligrosidad del microorganismo + valor económico del ATB, realiza la recomendación del siguiente paso más adecuado:

- cambio de antibiótico por uno más idóneo
- confirmación del antibiótico actual
- supresión del antibiótico

Nuestra API REST Machine Learning intenta automatizar y ayudar en este proceso. Cada fase de desarrollo de la API está consensuada con el grado de importancia dado por los especialistas, de ahí los lanzamientos de las releases 1.0, 2.0, 3.0 y 4.0.

Release 1.0

Recomendador de antibióticos (ATB) de las prescripciones propuesta, dando a conocer al usuario la probabilidad de éxito de cada una de ellas, ayudando a la toma de decisiones y validación de tratamientos.

Entrada de datos:

- Conjunto de datos del paciente
- Antibiograma
- Diagnóstico clínico del doctor, campo de texto libre no estructurado donde el médico apunta el diagnóstico (juicio clínico)
- Mapa resistencia hospital
- Peligrosidad microorganismos
- Valor económico ATB

Salida datos:

- Propuesta ATB
- Propuesta dosis
- Propuesta duración
- Propuesta vía

El sistema puntuará las recomendaciones de las prescripciones propuestas, dando a conocer al usuario la probabilidad de éxito de cada una de ellas, ayudando así al farmacéutico a la toma de decisiones y validación de tratamientos.

Release 2.0

Creación de un modelo de resistencia antimicrobiana (AMR) basado en genes a través de un test rápido. Añadiendo a la API rest este servicio, se evita la prescripción de antibiótico de amplio espectro para los primeros días, donde se está a la espera de los resultados del cultivo. Para aquellos hospitales que tengan la instrumentación, este servicio viene a sustituir el primer modelo del recomendador donde no se envían datos de cultivo ni antibiograma.

Partimos del trabajo en conjunto realizado por varios departamentos de la universidad de Québec, Canadá centrado en modelar el AMR. El método consiste en tomar del microorganismo una serie de **genes diana** como identificador inequívoca de él. El modelo de machine learning toma como partida la base de datos del Centro de Integración de Recursos de Pathosystems (PATRIC) con más de 100 mil secuencias de genomas microbianos y datos de los **fenotipos de resistencia** AMR para los 56 principales ATB. Se consigue dar una predicción de resistencia al doctor que realiza el primer diagnóstico.

Release 3.0

Consistiría en añadir a los modelos, la posibilidad de que el hospital pueda enviar datos de monitorización del paciente en tiempo real. Somos conscientes de que cada vez son más

los dispositivos IoT disponibles en el mercado en el mundo de la salud, por lo tanto debemos de ser capaces a medio plazo de poder integrarlos en la plataforma.

Release 4.0

Lanzamiento del recomendador prescriptivo a un nivel mucho más complejo y fiable, consecuencia del gran volumen de datos históricos recogidos y mejoras de integración de información de libros de referencia, papers, publicaciones científicas, mayor número de colaboraciones con hospitales, gobiernos y contratación de médicos infectólogos que den validez y mejoren los modelos.

Visión 360º del paciente (Analítica de datos)

Se pretende dar acceso a toda la información de los pacientes desde un único Dashboard o cuadro de mando. El objetivo de este Cuadro de Mando es unificar e integrar toda la información del paciente, que está distribuida en diferentes bases de datos (silos de información) y en cuya recopilación manual se tarda varias horas al día, para que asista y ayude a priorizar de una forma rápida y ágil cuáles son las mejores decisiones a tomar. Permitirá lanzar alarmas según el tipo de infección y priorización de casos. A partir de ahora cuando hagamos referencia a éste lo llamaremos visión 360º del paciente.

La visión 360º del paciente constituiría un primer MVP que ayudaría a ahorrar una gran cantidad de trabajo diario derivada de la tarea de recopilación de información.

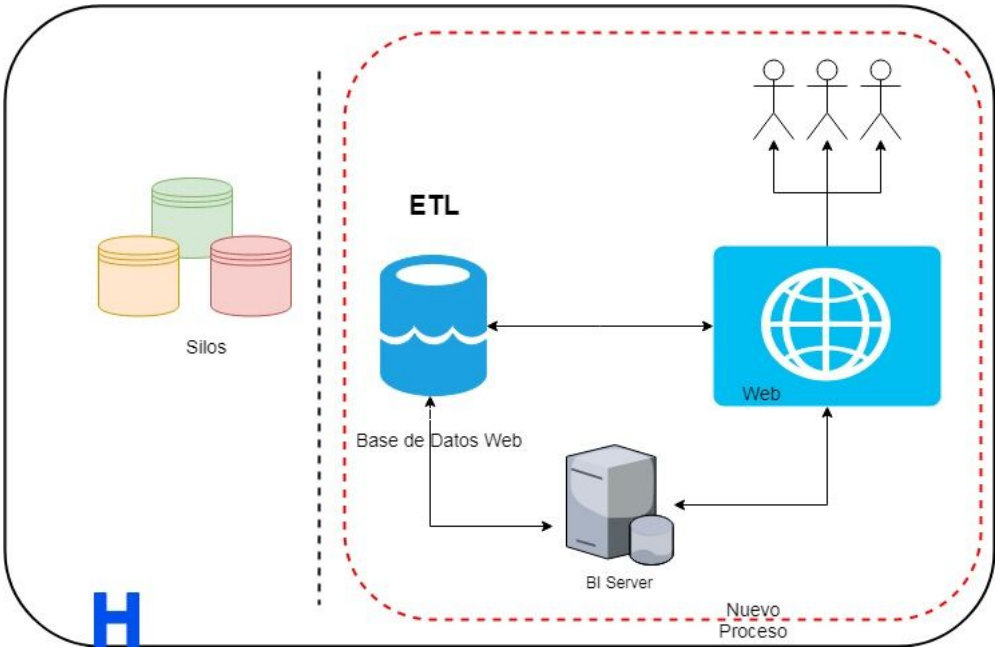
La solución ideada consiste en una aplicación web instalada en un servidor interno del hospital (intranet). Para ofrecer una visión 360 grados se hace uso de la herramienta de business intelligence&analytics TABLEAU server. Se hará uso de las librerías que permiten embeber los dashboard personalizados por tipo de usuario y será capaz de autenticar al usuario con el servidor de dashboard/report de TABLEAU. Si existe un nuevo dashboard publicado en TABLEAU Server, la aplicación lo detectará y permitirá su acceso, no siendo necesaria actualización ninguna. El principal motivo de hacer una aplicación web, está condicionado a que en la siguiente fase se integrarán las llamadas API REST de nuestros modelos de Machine Learning en AWS.

python™

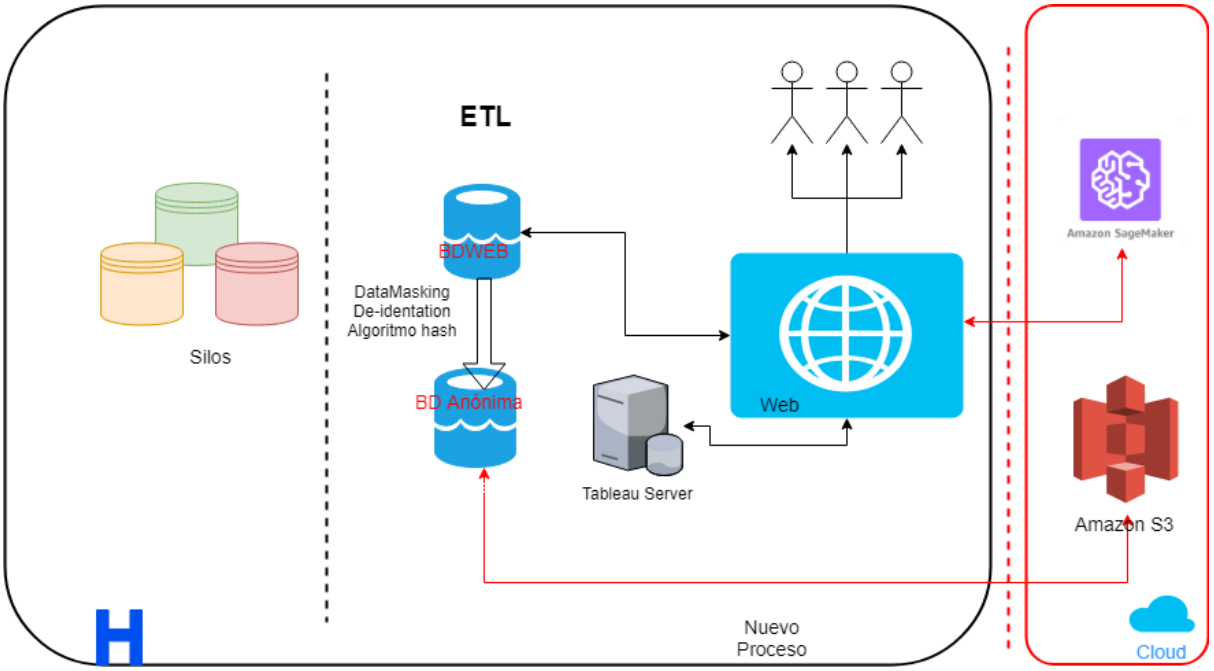
django

La aplicación web estará construida en DJANGO, está más que demostrado lo rápido, sencillo y robusto que es montar una web con este Framework, además Tableau posee una API REST en python necesaria para el control de usuarios y dashboards.

A continuación se muestra esquema arquitectura módulo 1 de alto nivel.



Hasta ahora no hemos hablado de dos elementos presentes en la arquitectura, el ETL y la base de datos, piedra angular del módulo 1, cuyo coste se estima que será entre el 50% y el 70% del trabajo y gasto del cuadro de mandos, es por ello que se dedicará un apartado completo a él. Por otro lado, dentro del ETL se definirán Jobs de extracción y anonimización de los datos necesarios para crear y entrenar nuestro modelos, así que vamos a presentar cómo se integra la arquitectura de alto nivel del producto 2 dentro del módulo 1.



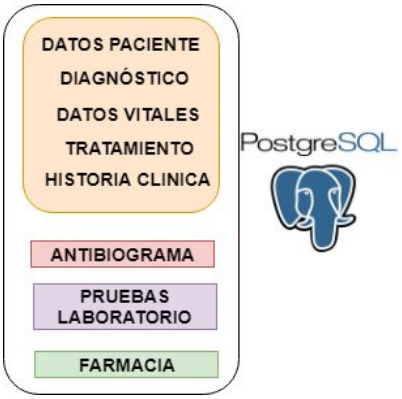
Principalmente se agregan jobs de anonimización y se crea una BD que aloja los datos a enviar a Amazon S3 que alimentará nuestros modelos.

ETL: Talend

La información necesaria por los equipos de validación de farmacia y PROA se encuentra distribuida en diferentes sistemas no integrados, llamados silos de información. En el caso de hospitales de pertenecientes al SAS (Sistema Andaluz Sanitario) tenemos:

- Farmacia: Domnium Farmatools. Antibióticos prescritos y contraindicaciones.
- Omnium: Antibiograma.
- Pruebas Laboratorio
- Diraya: Historia Clínica pacientes (Oracle).Tipos de infección, tratamiento recibido y éxito/no éxito del caso

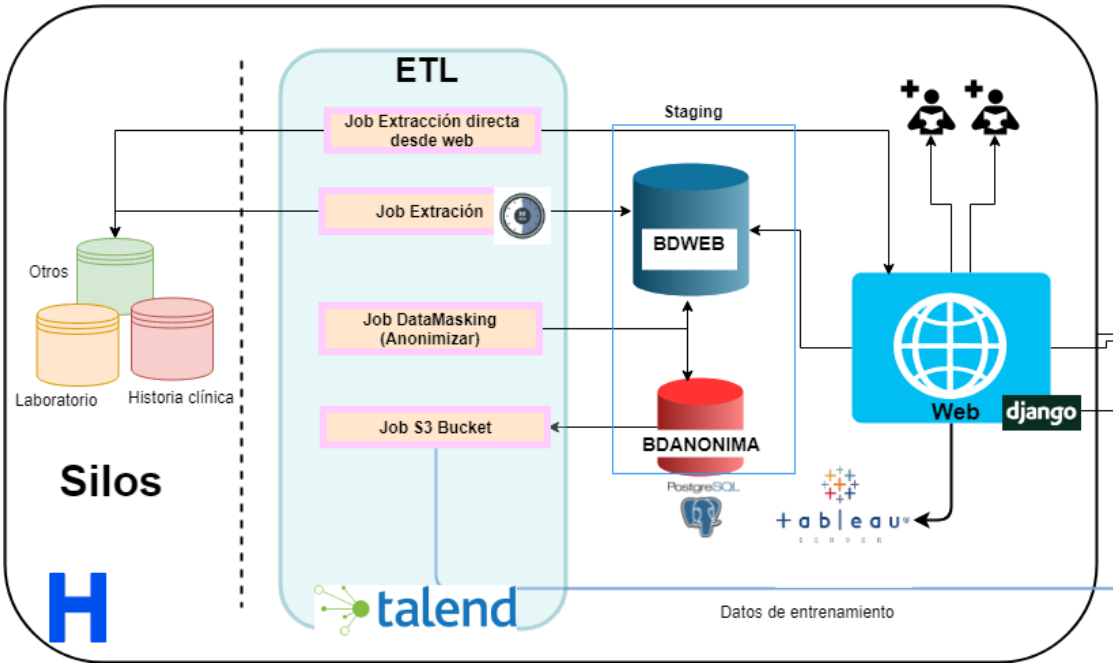
Cada aplicación tiene su base de datos independiente, por lo que será necesario usar una herramienta ETL que extraiga los datos, los armonice, categorize, filtre, limpie y aloje en una base de datos independiente a las de producción, llamada BD de Staging. Se decide usar PostgreSQL ya que es gratuita y tiene buen rendimiento. La llamaremos BDWEB y está compuesta por los datos de la extracción más los datos que van validando y añadiendo los usuarios de la aplicación web. Es estructurada y contiene los datos de pacientes del último año.



El enfoque ha sido utilizar una plataforma para la integración de datos (denominada, herramienta de extracción, transformación y carga, o ETL) para pasar los datos a un repositorio central. Operaciones de filtrado, agregación, combinación, limpieza y validación.

Se decide hacer uso del software de integración de datos empresariales TALEND. Aunque su curva de aprendizaje es mayor frente a herramientas como Pentaho, el producto Talend tiene más recorrido que otros. Los recursos invertidos en su desarrollo superan a los de cualquier otro software parecido y se está complementando con otras herramientas para crear una verdadera suite de integración de datos, cubriendo todas nuestras necesidades.

Se trata del software referencia del sector, con las herramientas ofrecidas en su versión open source suele cubrirse todas las necesidades aunque recomendamos usar la versión de pago, dejándolo a elección.



Con la herramienta de diseño construimos Jobs, utilizando el set de componentes disponibles.

- Jobs de extracción de datos BD microbiología, laboratorio, historia clínica y farmacia. Compromiso con los empresas de los distintos software involucrados, implementando actualizaciones del ETL si hubiera cambio en el modelo de base de datos usado. Expansión a otros software utilizados

- Jobs de transformación y carga de los datos para ser usados por la aplicación web y herramienta de BI. Uso Base de Datos PostgreSQL llamada BDWEB.

- Job disparador de inicio de extracción a demanda del usuario, lanzada desde la aplicación web.

- Jobs de transformación y carga de los datos para ser usados en los trainings de Machine Learning. Los datos se almacenan en Base de datos PostgreSQL llamada BDANONIMA. Existen rutinas de DataMasking que permiten realizar el enmascaramiento de los datos. Encriptar cadenas usando AES 128. Calculates MD5 hash value from String, etc. https://help.talend.com/reader/ztNTsvuA6OUYwnhO3e_YIQ/9HXDpsKTHsKBq8BhFu~0cw

- Job conexión AWS S3 bucket datos anonimizados.

Cada 30 minutos serán lanzados los jobs de extracción de datos de las distintas BD del hospital, con el objetivo de que la web muestre una visión 360º con información muy actualizada al equipo PROA y validador de Farmacia, sin comprometer rendimientos de bases de datos de producción.

Aún así, existe la posibilidad de llamar desde la app web al proceso disparador de extracción de datos para un paciente en concreto, teniendo el dashboard actualizado en **tiempo real**.

Los datos anonimizados, tendrán partes encriptadas cuyas semillas son almacenadas en base de datos. Se extraen todos los datos anonimizados de pacientes que hayan sido tratados y validados y finalizados por el equipo PROA y se guardan en BDANONIMA. Semanalmente se lanzará una job llamada tS3Put que permite conectar directamente con bucket S3 de Amazon y enviar todos los pacientes finalizados como alimento de nuestros modelos. Los puntos clave en el tratamiento de estos datos:

- Establecer el nivel de anonimización o seudo anonimización. Algoritmo hash usado.

- Definir qué datos identificativos no pueden ser extraídos fuera de la red hospitalaria. Todos los directos son eliminados y los indirectos cifrados.

- Impacto de la nueva GDPR, posibilidad de reutilizar los datos de pacientes para varios proyectos sin la nueva firma del paciente.

Seguimos las recomendaciones de la Agencia Española de Protección de datos para anonimizar datos y debido a la importancia y criticidad de esta materia en nuestro proyecto, le dedicaremos un punto completo a su desarrollo.

Anonimización de datos

La finalidad del proceso de anonimización es eliminar o reducir al mínimo los riesgos de reidentificación de los datos anonimizados manteniendo la veracidad de los resultados del tratamiento de los mismos, es decir, además de evitar la identificación de las personas, los datos anonimizados deben garantizar que cualquier operación o tratamiento que pueda ser realizado con posterioridad a la anonimización no conlleva una distorsión de los datos reales. Un análisis masivo de los datos o macrodatos que puedan derivar de los datos anonimizados no debería diferir del análisis que pudiera obtenerse si hubiera sido realizado con datos no anonimizados.

En el proceso de anonimización se deberá producir la ruptura de la cadena de identificación de las personas. Esta cadena se compone de microdatos o datos de identificación directa y de datos de identificación indirecta.

Siguiendo las recomendaciones de la Agencia Española de Protección de datos y todas las limitaciones y sanciones de la nueva ley GDPR seguimos cada una de las fases y cumplimos con cada principio establecido en:

https://datos.gob.es/sites/default/files/doc/file/orientaciones_y_garantias_anonimizacion_0.pdf.

En un proceso de anonimización es aconsejable definir un protocolo de actuación.

Definir un equipo de trabajo, segregando funciones atendiendo a perfiles o roles. Responsable de fichero, responsable GDPR, equipo riesgos, etc y garantizando la independencia de las funciones de cada uno. Inicialmente será realizado una evaluación de riesgos completa (EIPD) de reidentificación de los datos, vital para no excedernos en recursos o caer en errores graves.

Preanonimización de los microdatos es la parte inicial del proceso de anonimización, en el que se determinarán las posibles variables de identificación (**directas e indirectas**) a tener en cuenta en el diseño de las herramientas de anonimización. Durante el proceso de **preanonimización** se tendrá en cuenta:

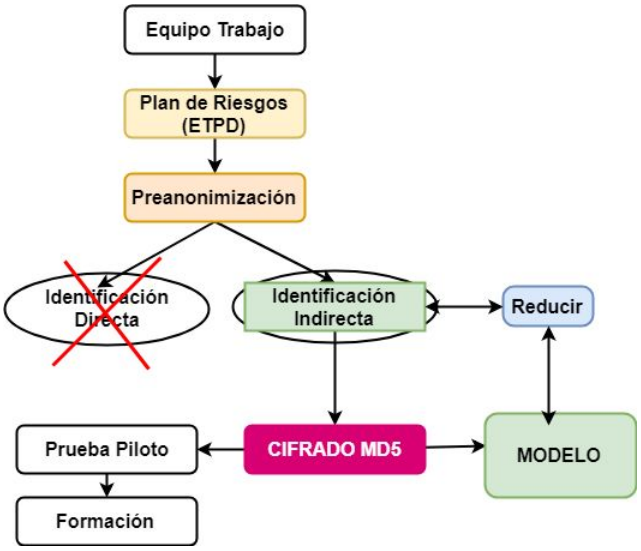
- La determinación de variables: datos personales, identificadores directos e indirectos, datos especialmente protegidos y otros datos con carácter confidencial. La clasificación y sensibilidad de las variables por categorías: de identificación directa, de identificación geográfica, de carácter especialmente protegido, numéricas, temporales, metadatos, etc.

- Variables de identificación que no puedan ser anonimizadas y que sea preciso eliminar del proceso de anonimización.

- Variables anonimizadas que sean imprescindibles para la finalidad a la que se van a destinar los datos anonimizados.

Eliminación/reducción al mínimo de variables no necesaria para nuestros modelos que permitan la identificación de las personas, restringiendo el acceso a la información confidencial al equipo de trabajo implicado en el proceso y optimizando el coste computacional de las operaciones con datos anonimizados. Volveremos a este punto tras entrenar los modelos y cuando este me descarte todas aquellas variables independientes que no tienen peso en la predicción de la variable dependiente.

Selección técnica de anonimización, **función hash md5** realizada por Job del ETL. Realización de una **prueba Piloto** y finalmente **formación al personal**.

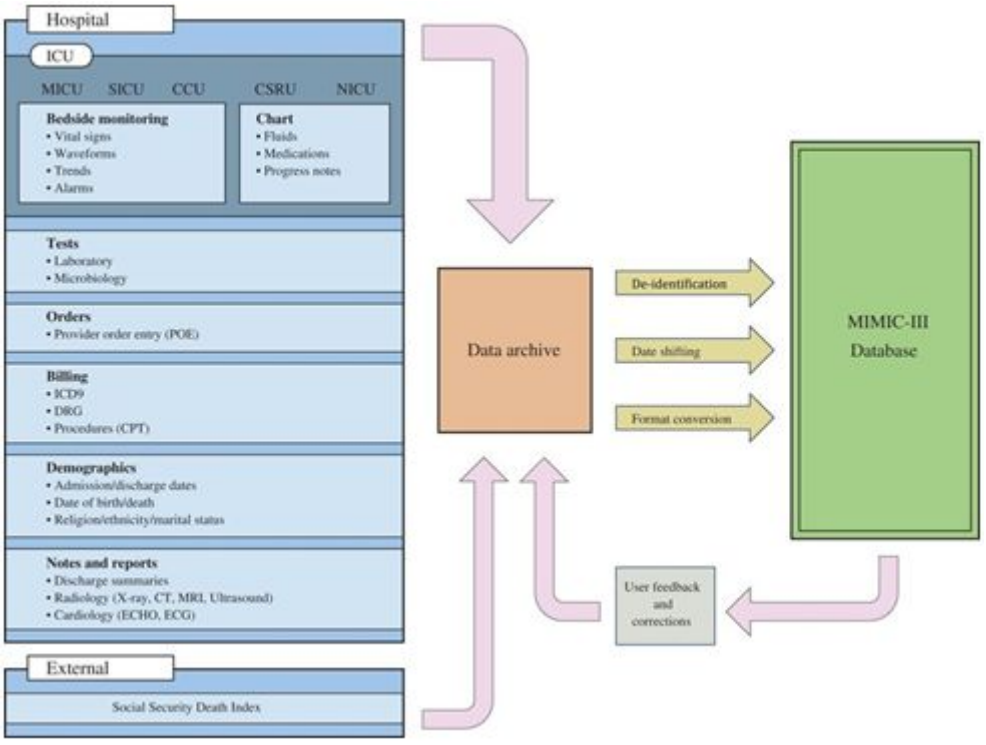


Datos reales para pruebas

Aún no hemos podido acceder a datos reales de pacientes en hospitales en España, hemos contactado con el departamento de informática del hospital Macarena, pero han rehusado a colaborar con nosotros debido a que ya están trabajando con otra empresa en un proyecto parecido y se lo impiden los contratos de confidencialidad. Al menos nos confirma que existe la posibilidad de acceso al dato.

Por fortuna existe una base de datos real llamada MIMIC-II, con registros de más 53000 pacientes en cuidados intensivos registrados a lo largo de una década. Ha sido internacionalmente usada por investigadores y universidades, por lo tanto está muy bien documentada. Existen números papers con estudios de oncología, diabetes, antibióticos, etc. En el siguiente enlace se describe al completo como usarla y los pasos para tener acceso a ella.

https://www.nature.com/articles/sdata201635#f1



Tableau

Se opta por crear los dashboards con una herramienta de BI en vez de crear una aplicación web a medida usando librerías de gráficos tales como D3.js, bokeh, etc. Como casi todo el trabajo de exportación de la aplicación a una nueva red hospitalaria se lo llevará el ETL, se hará un trabajo concienzudo dando lugar a que la base de datos de STAGING (DBWEB) esté lo más normalizada posible entre distintos hospitales, siendo por tanto cien por cien exportables los dashboard creados.

- Tableau permite embeber los dashboard en web.
- Se hará uso de tecnologías como SQL Direct o LiveConnection, las cuales permiten hacer consultas directamente sobre nuestra base de datos BDWEB.
- Se instala on premise en la red hospitalaria.
- Con la licencias Explorer, el personal del centro hospitalario puede modificar y/o crear nuevos dashboards según sus necesidades, ya que ellos son los grandes conocedores del negocio. La aplicación web permitirá acceder a esos nuevos dashboard alojados en Tableau server.

- Permite varios métodos de autenticación de usuarios y gran variedad de conectores con fuentes de datos.
- Tiene herrameintas de analítcia de datos avanzadas, muy útil para estadísticas, evoluciones, detección de alarmas, etc.

Gracias a la base de datos MIMIC-II podemos crear los primeros dashboards como presentación y venta de la herramienta.

Para más información:

https://www.tableau.com/es-es/products/techspecs#server

https://onlinehelp.tableau.com/current/pro/desktop/en-us/embed.html

API Rest AWS modelos Machine Learning

Una de las razones de haber elegido hacer funcionar los modelos en cloud y no on premise dentro del hospital es que será la Producto 2 y expansión de la API REST a un mayor número de clientes, lo que hará disparar los ingresos frente a un crecimiento lineal de los costes, además de los posibles altos costes en hardware on premise que serían necesarios. El lenguaje de programación usado será Python ya que varios componentes del grupo lo dominan y se decide hacer uso de la herramienta para construir modelos AWS Sagemaker en vez de montar un servidor elástico en AWS (ElasticBeans) ya que es más complejo de gestionar, mantener y debiéndose de crear el webservice en Flask para la API, funciones de tratamiento y proceso de datos de entrada y salida, instalación de librerías, etc. Sagemaker además te permite poner en producción modelos sostenidos sobre una arquitectura serverless altamente escalable en muy poco tiempo a coste razonable, siendo números los casos de éxito.

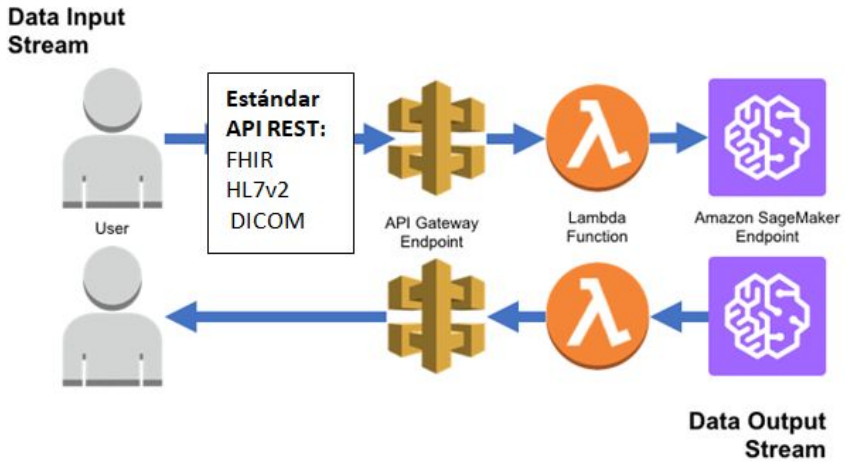
Amazon SageMaker, ofrece una poderosa plataforma donde construir, desarrollar y entrenar modelos de machine learning dentro de AWS. Además combinando esta poderosa herramienta junto con las capacidades de arquitectura sin servidor ofrecidas por Amazon *Simple Storage Service (S3), Amazon API Gateway, y AWS Lambda*, es posible transformar un AWS Sagemaker endpoint en un servicio web que acepta datos de entrada y devuelve las inferencias resultantes al usuario. El flujo de tareas de las funciones lambda será gestionado por el servicio Step Functions.

En resumen, se ha decidido usar una ***arquitectura Serverless*** para nuestros modelos API REST, servida por Amazon a través de servicios AWS Lambda quien llamará al endpoint de Sagemaker y API creada y publicada con AWS API Gateway.

En API Gateway solo nos preocupamos de publicar los métodos GET, POST y PUT, y Amazon se encarga de atender a la peticiones y además de escalar ante grandes volúmenes de peticiones. Solo se deberá pagar por las llamadas que reciba la API y por la cantidad de datos de salida transferidos.

Según Amazon Web Services: *Amazon API Gateway se integra a la perfección con AWS Lambda para crear API sin servidores. Amazon API Gateway permite crear API REST para que aplicaciones móviles y web puedan llamar a los servicios de AWS disponibles públicamente mediante el código ejecutado en AWS Lambda. AWS Lambda ejecuta el código en una infraestructura informática de alta disponibilidad sin preocuparnos de la administración de los recursos informáticos.*

En nuestro caso, las funciones Lambda en primer lugar tratan los datos de entrada, filtran, formatean y deciden a qué modelo llamar de los endpoint Sagemaker publicados, y los datos pueden ir de vuelta a que se ejecute otro modelo o enviarlos a API Gateway de vuelta al usuario. Este flujo de trabajo es gestionado por AWS Step Functions. Si se tuviera que desencriptar algún dato o modificarse también se encargaría.



Un ejemplo sería: recibo llamada a la API con estudio de Antibiograma, la función lambda hace un filtro de los 56 ATB disponibles ya que sabe cuales tienen sentido usar, añade algunos datos particulares para distribución de pesos en el algoritmo y realiza una llamada a Sagemaker, el cual ejecuta para cada ATB. El conjunto de resultados los envía de nuevo a Sagemaker, en este caso al modelo recomendador junto otros datos identificativos del hospital que hace la llamada y finalmente devuelve la recomendación.

Endpoints a ser creados en la **release 1.0**.

- Modelo supervisado clasificador solo con datos paciente y Diagnóstico doctor.
- Modelo supervisado clasificador ATB elegido, 3 variables dependientes (Dosis, duración, via). Puede que sea dividido en tres modelos.
- Modelo supervisado clasificador Recomendador

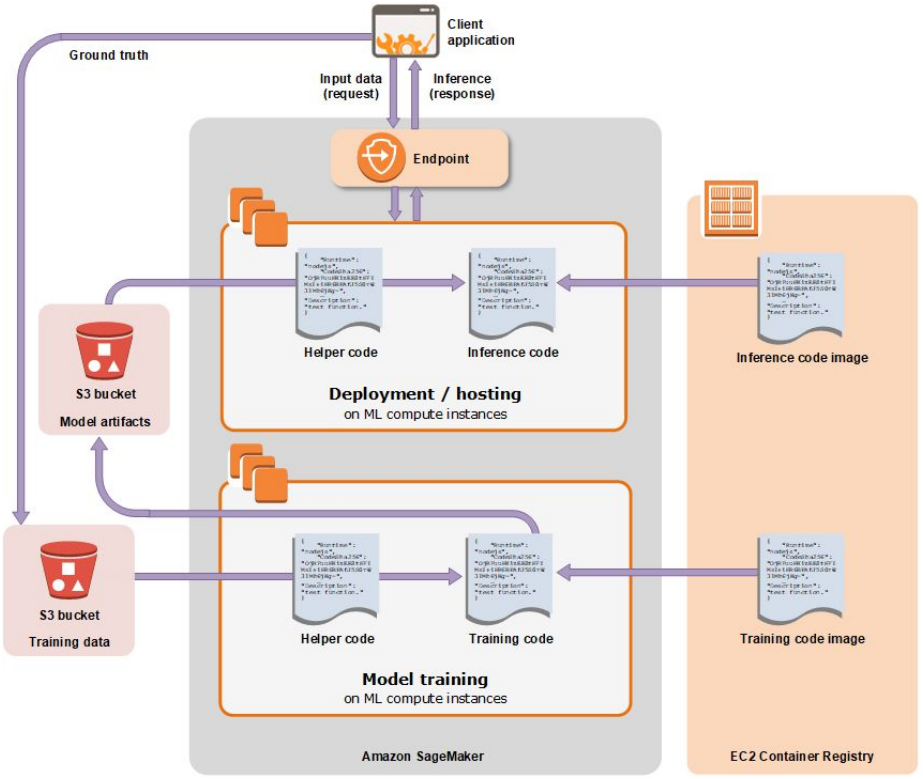
Estarán configurados en alta disponibilidad, al menos dos instancias del endpoint en zonas geográficas distintas, siempre dentro de EU por motivos de GDPR (Irlanda y Luxemburgo)

Configurar en Sagemaker un endpoint es muy sencillo, los pasos a seguir son:

1. Crear S3 Bucket
2. Crear instancia Notebook Sagemaker
3. Entrenar y desarrollar un modelo:
    - Crear Jupyter Notebook, Descargar, explorar y transformar los datos.
    - Entrenar el modelo
    - Desarrollo del modelo y Validación del modelo
    - Automatic Model tunning
4. Publicación de las Instancias Jupyter Notebooks.
5. Creación de los endpoints a ser llamados por Lambda

Sagemaker brinda al científico de datos todas la herramientas necesarias para crear sus modelos, despreocupandose de tener que configurar un servidor de entrenamiento con Jupyter Notebook y demás librerias python ML.

Se muestra la infraestructura de entrenamiento y desarrollo de modelos.



Protocolo estándar medicina API Rest

Es esencial que una aplicación que quiera ser vendida a nivel global utilice un estándar ya utilizado por la comunidad, en nuestro caso son varios los estándares más usados en salud cuando utilizamos servicios web REST. No podemos caer en la tentación de crear

aplicaciones a medida y que luego se adapten a nosotros. Nos comprometemos a que la API REST sea totalmente compatible con los estándares más usados en medicina, facilitando la integración de la API REST a aquellos hospitales que solo quieran hacer usos de la FASE 2 del proyecto.

- FHIR (Fast Healthcare Interoperability Resources) Es el estándar emergente para la asistencia sanitaria en interoperabilidad. FHIR especifica un modelo de datos robusto y extensible usando REST. Semántica para interactuar con los recursos. FHIR se beneficia de los avances técnicos en el desarrollo web para ofrecer una tracción significativa de la industria, incluidos los principales sistemas de registros de salud electrónicos y apoyo entre gobiernos.

- HL7v2 es una modalidad de comunicación esencial para cualquier aplicación que busque integrarse con sistemas clínicos heredados. La API HL7v2 implementa una interfaz REST para ingerir, enviar, buscar y recuperar mensajes HL7v2

- DICOM es el estándar establecido para almacenar y intercambiando imágenes médicas a través de una amplia gama de modalidades, incluyendo Radiología, cardiología, oftalmología y dermatología. DICOMweb es una API REST para organizar y consultar imagenes medicas.

Recomendaciones de instalación del conjunto de herramientas

Siguiendo criterios de buenas prácticas en seguridad informática recomendamos seguir el esquema de infraestructura propuesto. Sobre todo separar el servidor que aloja la aplicación web que tendrá conexión a internet con el servidor de base de datos.

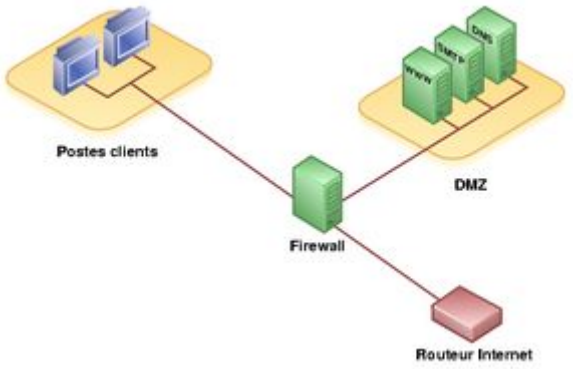
Infraestructura ON PREMISE compuesta por 3 servidores, compatibles en Linux / Windows.

Servidor DMZ 1 Apache7 + Django + WSGI

Servidor 2 base de datos para PostgreSQL

Servidor 3 Talend + Tableau Server

Un servidor DMZ es aquel que se encuentra alojado en una zona desmilitarizada o área perimetral (DMZ). Los servicios alojados en este servidor son los únicos que pueden establecer tráfico de datos entre la DMZ y la red interna, en este caso la conexión de datos entre un servidor web y una base de datos protegida situada en la red interna. Las conexiones que tendría el cortafuegos que habilitar serían con PostgreSQL, API python de Tableau y API Rest del Extractor TALEND, API REST con AWS API Gateway.

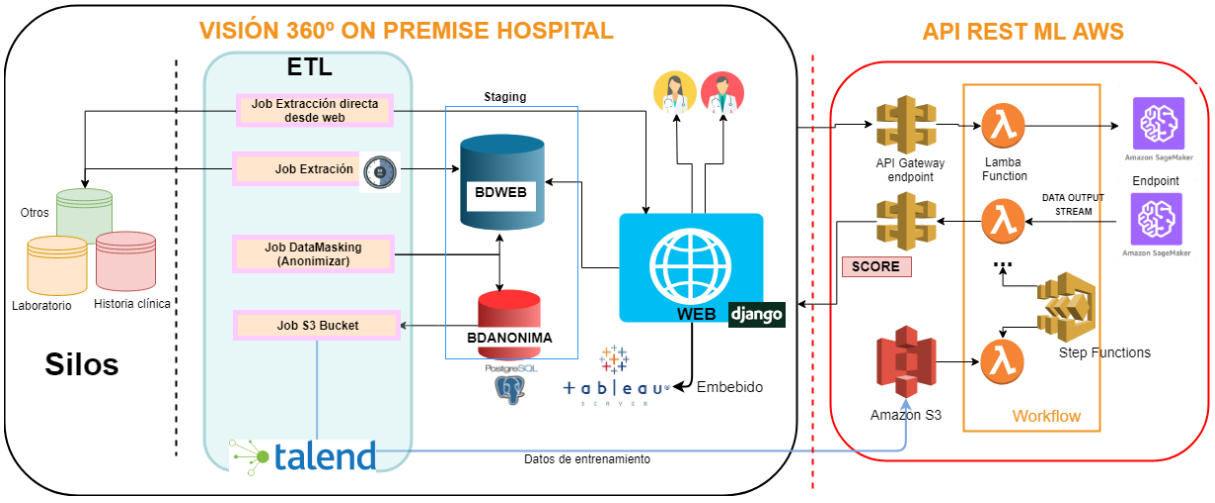


Los servidores 1 y 2 tienen requerimientos hardware en línea con el estándar actual, el servidor 3 si requiere de mayores prestaciones en RAM y CPU. Las recomendaciones serán las dadas por Talend y Tableau:

https://help.talend.com/reader/5SCAnQgH2QjsoRvKUKvtbw/DEfM1YKWrxM3qNFA6AZjmw

https://onlinehelp.tableau.com/current/server/es-es/server_hardware_min.htm

Diagrama arquitectura completa



Modelos

Tras el data discovery, data collection, data exploration, data validation y data cleaning del trabajo conjunto entre el científico de datos y personal sanitario altamente cualificado, llega el momento de planear los modelos, ejecutarlos, estudiar los resultados y probar diferentes algoritmos y entrenarlos.

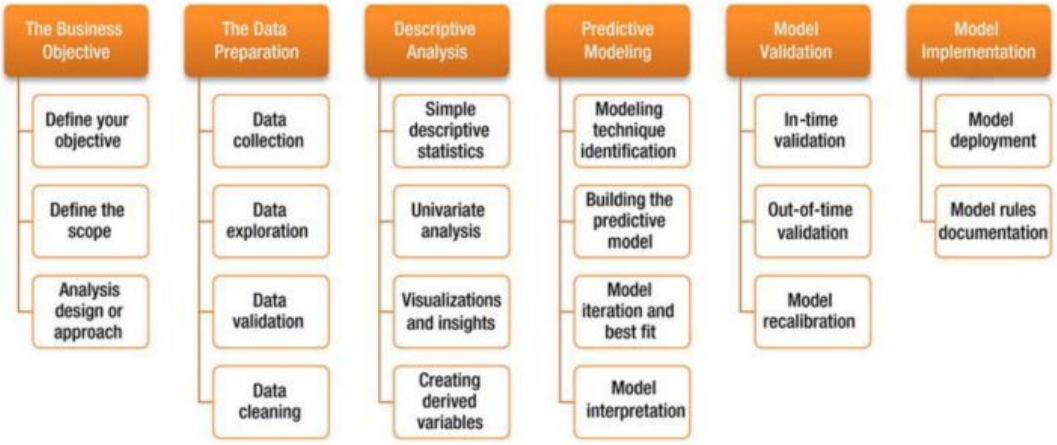
El planteamiento inicial de los modelos es crearlos simplificados utilizando para ello variables con capacidad predictiva contrastada. Se ajustan los pesos con la ayuda de los

expertos y a medida que evolucionan y entrenan se añaden y prueban nuevas variables y redistribuyen los pesos (calibración).

Las métricas de validación del algoritmo usado serán:

- Especificidad > 90%
- Tiempo de ejecución < 5sg
- Clasificación correcta > 90 instancias
- Clasificación incorrecta < 10 instancias
- Sensibilidad >90%
- Mean absolute error < 0.1 ideal
- Kappa statistic >0.79, >0.9 ideal
- Root mean squared error < 0.3, <1 ideal
- Relative absolute error <25%, <1 ideal
- Root relative squared <50%, 0 ideal

Ciclo de vida modelado



Fuentes de datos

Guia Antibioterapia conjunta del hospital virgen Macarena y virgen del Rocío de recomendaciones de tratamiento empírico en adultos. Se trata de una clasificación enfermedades infecciosas y recomendaciones del ATB a prescribir, dosis, duración y vía, alternativas y pasos a seguir si no funciona ATB previo. Se trata de información desestructurada pero de fácil conversión a estructurada. De aquí se obtiene gran cantidad de variables independientes para nuestros modelos.

http://www.hospitalesmacarenayrocio.es/activos/antibioterapia/index.php/temprico/12-infecciones-orales,-far%C3%ADngeas,-sinusales,-%C3%B3ticas-y-oftalmol%C3%B3gicas

Un trozo de la tabla es tal que así:

Síndrome	**Etiologías frecuentes**	**Tratamiento empírico**	**Comentarios**
Candidiasis oral	*Candida* spp	Oral: nistatina suspensión 100.000UI/ml 5 ml/6-8h, enjuagar y tragar 7 días. Alternativa: fluconazol 100 mg/día	Realizar cultivo para antibiograma si de repetición o no responde
Gingivoestomatitis herpética	Virus Herpes Simple	Aciclovir 200 mg 5 dosis al día, o bien 400mg 3 veces al día durante 5 días. Famciclovir 500mg/12h ó valaciclovir 1g/12h son **alternativas** más cómodas, pero **más caras**. Si no tolera vía oral, aciclovir 5 mg/kg/8h IV.	Valorar necesidad de tratamiento frecuentes, fiebre y dificultad para la ingesta. En casos graves tomar muestra para PCR y cultivo (medio de transporte de virus). Considerar el uso de analgesia (p.ej. lidocaína viscosa).

Para el síndrome y la etimología usaremos la codificación numérica ICD-9 recomendada por la ONU. Para el cuadro tratamiento se crean columnas en base de datos.

Oral: nistatina suspensión 100.000UI/ml 5 ml/6-8h, enjuagar y tragar 7 días. Alternativa: fluconazol 100 mg/día

- Vía: oral (0), sublingual (1), bucal (2), tópica (3), transdérmica (4), inhalación (5), instilación ocular (6), rectal (7), vaginal (8), subcutánea (9), intramuscular (10), endovenosa (11)
- ATB recomendado: nistina (numeración de 0-56 antibióticos con los que se trabaja en hospitales)
- Dosis: 100.000 UI/ml
- Duración: 7
- Intervalo: 6-8
- ATB alternativo: fluconazol (campo numérico)

El libro Guia de terapeutica antimicrobiana de Josep Mensa que no está informatizado hace exactamente lo mismo pero con un mayor grado de profundidad.

Mapa de resistencia propia de cada hospital, a qué suele ser resistente el entorno de un hospital. Estos mapas de resistencia suelen estar disponibles en Internet o en cada Hospital.

Lista de microorganismos MENSA priorizados.

Precio Antibióticos oficiales.

A la espera de poder tener datos de pacientes en España tenemos los de la base de datos real MIMIC-III ya mencionada anteriormente y usada en numerosos paper científicos. Contiene los siguientes datos:

- Datos paciente
- Datos Antibiograma
- Diagnóstico médico tras aplicar técnicas de text mining
- Datos pruebas laboratorio
- Datos farmacia, ATB, contraindicaciones y alergias

Clasificación de notas médicas y diagnóstico del doctor en lenguaje natural

Para ello se aplicaran técnicas de text-mining con las que etiquetar y estructurar la mayor cantidad posible de información relevante anotada. Lo primero a etiquetar es el tipo de enfermedad, para ello usaremos la lista estándar internacional ICD-9 de códigos alfanuméricos de seis caracteres para describir diagnósticos. Los códigos 001–139 corresponden a infecciones.

Nuestro modelo utilizará de base de un paper publicado por Amitabha Karmakar University of California Berkeley, CA amitkarmakar@berkeley.edu, titulado Classifying medical notes into standard disease codes, utilizando para ello las notas médicas de la base de datos MIMIC III. El algoritmo recomendado es una **red neuronal** con una capa de convolución.

https://arxiv.org/pdf/1802.00382.pdf

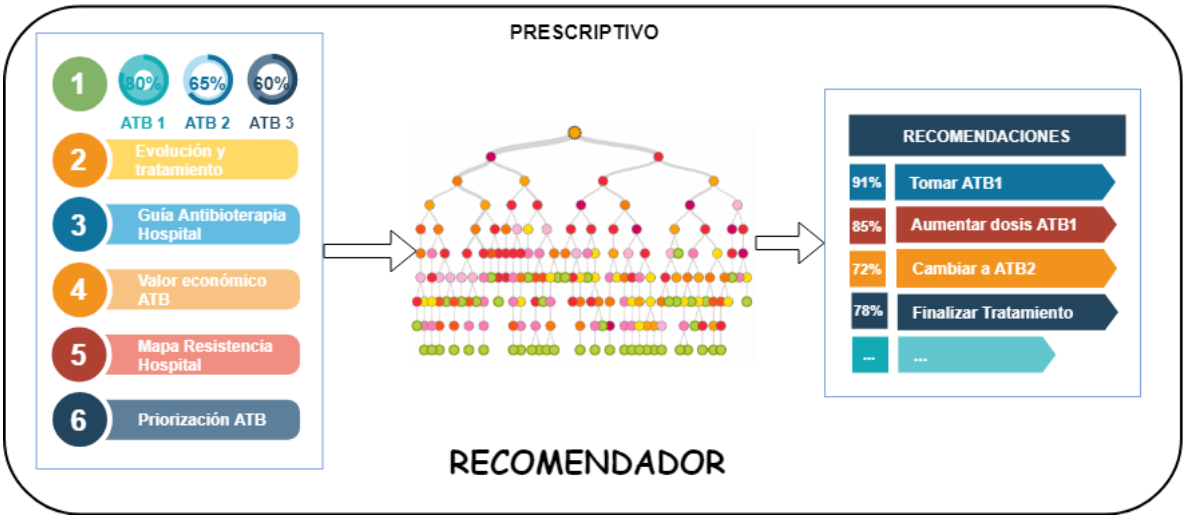
Recomendador de antibióticos (análisis prescriptivo)

Nuestro recomendador es el resultado de la ejecución de un modelo supervisado clasificador con algoritmo árboles de decisión. Los árboles de decisión son una de las opciones más utilizadas para el razonamiento y el aprendizaje basado en ejemplos, para tratar con el lenguaje natural y la incertidumbre.

Como entrada se tiene una lista de los mejores antibióticos junto a sus probabilidades de éxito que a su vez son la salida de la ejecución modelos predictivos basados en datos históricos, más toda la información relativa a recomendaciones de tratamiento de la Guia Antibioterapia conjunta del hospital virgen Macarena y virgen del Rocío, Mapa resitencia del hospital, valor económico de cada ATB, priorización ATB especificada y recomendaciones

de uso de la ONU. El resultado final es una una nueva recomendación en base de probabilidad que podrá ser:

- Continuar con el tratamiento (el tratamiento propuesto es el mismo)
- Modificación del tratamiento (propone nuevo tratamiento)
- Finalización del tratamiento (el tratamiento propuesto ha sido efectivo)

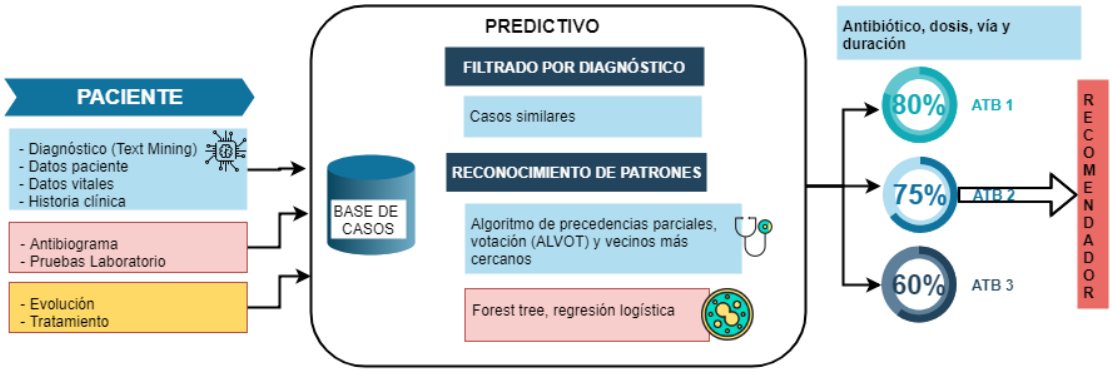


Se ofrece la posibilidad de envío a través de la llamada a la API de los datos valor económico ATB, mapa resistencia hospital y priorización específica ATB propuesto por un cliente en concreto. Si se quiere hacer uso de otra guia antibioterapia se debe de realizar un desarrollo previo con el cliente, que almacenaremos en la cloud. Así ofrecemos un recomendador adaptado a las necesidades del cliente.

Destacar que la complejidad en la construcción del recomendador se la lleva casi en su totalidad la creación y calibrado de los modelos predictivos que generan la lista de antibióticos junto a probabilidades de éxito para cada paciente. La calidad del dato es extremadamente crítica en nuestro sistema incluso más que la solidez de nuestros modelos, es la base de casos de pacientes validados por los mejore especialistas, la que hará que nuestro recomendador tenga éxito o no. El ajuste de peso y la elección de las mejores variables predictivas será un trabajo en conjunto con los mejores especialistas del sector y el científico\os de datos.

Mencionar que los hospitales trabajan con 56 antibióticos distintos, cada enfermedad puede ser tratada por un número de ATB definido y conocido, una función Lambda se encarga de decidir para qué ATB se hará predicción, obteniéndose finalmente la lista de los más certeros.

A continuación se realiza una descripción detallada de cada modelo de predicción entrada del recomendador, diferenciado por los datos que se tienen, finalmente como extra se describe parcialmente el modelo predictivo a usar en la release 2.0.



Modelo Predicción Antibiótico sin antibiograma

La llamada se hace antes de la prescripción del antibiótico de amplio espectro, ya que aún no se dispone de cultivo, antibiograma ni pruebas de laboratorio, las pruebas llegan en dos días. Si se tiene éxito, se podría prescindir de las pruebas (ahorro costes hospital), ahorro costes tiempo hospitalización y ahorro coste medioambiental al no seguir haciendo aún más resistentes a las bacterias con ATB no adecuados o con poco efecto sobre estos, sobremedicando o extendiendo la medicación en exceso.

Nuestro objetivo es predecir con la mayor probabilidad el mejor ATB a prescribir utilizando el mejor diagnóstico posible y datos de paciente, a través de la comparación de patrones de nuestro dataset validado por doctores. Utilizaremos Algoritmos de reconocimiento de patrones y nos hemos basado en un paper publicado por M. R. Ortiz-Posadas Maya-Beh~ y M. Lazo-Cortés.

http://host-article-assets.s3-website-us-east-1.amazonaws.com/ceb/5889fb885d01231a018b468c/fulltext.pdf

Menciona como las técnicas de reconocimiento de patrones, han resuelto algunos problemas de **clasificación supervisada en el área de la medicina.**

La metodología se sustenta en el uso del enfoque lógico-combinatorio de la teoría de reconocimiento de patrones, usada en la modelación de problemas relacionados con el diagnóstico médico diferencial. Se desarrolla una nueva metodología que permite pronosticar la rehabilitación del paciente y evaluar la tendencia en la evolución de la atención médica, mediante la comparación del resultado alcanzado en cada caso con el resultado pronosticado.

Los datos disponibles son:

- Datos paciente
- Monitorización (Fiebre, Presión arterial,Pulso...)
- Datos Antibiograma: No
- Diagnóstico médico tras aplicar técnicas de text mining
- Farmacia, contraindicaciones y alergias

Definición de las variables independientes del modelo junto dominios para cada antibiótico prescrito:

- Localización infección: Garganta, oido, sangre, orina, pulmones,...
- Dolor: Si o No
- Dosis: 1, 2 , 3, 4
- Vía: Píldora, inyección, ...
- Fiebre: No, moderada, alta, muy alta
- Días sin tratamiento: 1, 2 , 3 , 4
- Días con tratamiento: 1, 2 , 3, 4
- Cambio de ATB: Si o No
- Antibiótico anterior: 56 diferentes
- Edad: bebé, adolescente, adulto, anciano
- Alto riesgo: Si o No
- Antecedentes: Si o No
- Hábitos saludables: Ninguno, medio, alto.
- Muerte: Si o No
- Efectos secundario: Ninguno, alguno, medio, muchos
- Nivel: Leve, Grave, Muy Grave
- Otras muchas pertenecientes al campo de la medicina presentes en análisis de orina/sangre (leucocitos, plaquetas, …)

Una vez definidas las variables y sus dominios montamos las tablas de comparación para cada una de ellas, por ejemplo para la fiebre sería:

Fibre	No, moderada, alta, muy alta		no	moderada	alta	muy alta
		no	0	0.25	0.5	1
		moderada		0	0.25	0.75
		alta			0	0.5
		muy alta				0

Se utiliza un **algoritmo de precedencia parcial conocido como Conjuntos de Representantes** (Baskakova y Zhuravlev, 1981; Carrasco-Ochoa, 1994). Se utiliza un algoritmo basado en precedencia parcial porque, los expertos en muchos problemas trabajan sobre la base de analogías (totales o parciales), es decir, precedentes totales o parciales de sus experiencias y observaciones acumuladas. Las analogías parciales pueden interpretarse como tener un "punto de vista" particular en el que se ignoran ciertas variables; esto significa en términos matemáticos, hacer una proyección en el espacio de representación. En este sentido, los modelos que se proponen para la solución de los problemas de clasificación en las ciencias poco formalizadas se fundamentan en las precedencias permitiendo el análisis de la semejanza entre los objetos de estudio por partes

y posteriormente sobre la base de ellas, hacer un análisis global para tomar una decisión final.

Hablar un poco más de otros algoritmos supervisados clasificados de reconocimiento de patrones llamados por votación (ALVOT) o los del vecino más cercano.

La desventaja de este modelo es que necesito tener datos muy buenos y precisos, además de gran cantidad de ellos. También está la problemática de que por muchas variables independientes de datos de paciente que tenga, las reacciones a ATB de los cuerpos humanos son muy diferentes entre sí, factores como estado de ánimo, nivel de estrés, etc, afectan al sistema inmune.

Modelo predicción Antibiótico con antibiograma

En este caso tenemos perfectamente identificado el microorganismo y un estudio de sensibilidad de ATB, por lo que los pesos de las variables independientes de los datos del paciente tienen menor importancia y es mucho mayor el del estudio de sensibilidad. Problema de **clasificación supervisada.**

Los datos disponibles son todos los mencionados con anterioridad más el antibiograma (MIC) y otras pruebas de laboratorio.

Ejecución tras ejecución del modelo se detectarán qué variables independientes tienen mayor peso en la probabilidad de que mi variable dependiente o target sea más alta, dando como resultado una función matemática con pesos de variables independientes, siendo rechazadas muchas de ellas, ya que apenas aportarán valor al modelo predictivo, haciéndolo más sencillo y explotable. El algoritmo empleado será el que mejores métricas obtenga entre, forest tree, regresión logística, redes neuronales y redes bayesianas, aunque preferiblemente ante métricas parecidas nos quedaremos con los forest tree y regresión logística ya que son más fáciles de entender y usar.

Hemos encontrado analogía entre la construcción de un modelo de scoring para el otorgamiento de crédito de una entidad financiera y la construcción de un modelo de scoring de un antibiótico sobre un microorganismo conocido, salvaguardando las distancia de las variables independientes de un sector y otro pensamos que el camino a seguir es parecido y afortunadamente aunque para antibióticos apenas existen paper publicados si lo existen para riesgo en crédito. Adjuntamos algunos consultados en el apartado paper.

Algoritmos a ser probados:

- Las Redes Neuronales
    - Multilayer Perceptron
    - RBFNetwork
- Las Redes Bayesianas
    - BayesNet
    - Bayes simple estimator
    - BMA bayes

    - Naive-Bayes
    - BayesNet Kernel
    - Naive-Bayes Discretización Supervisada
- Árboles de decisión
    - J48
    - Decision Table
- Regresión Logística
    - MultiClassClassifier

Las métricas de validación del algoritmo usado serán:

- Especificidad > 90%
- Tiempo de ejecución < 10sg
- Clasificación correcta > 90 instancias
- Clasificación incorrecta < 10 instancias
- Sensibilidad >90%
- Mean absolute error < 0.1 ideal
- Kappa statistic >0.79, >0.9 ideal
- Root mean squared error < 0.3, <1 ideal
- Relative absolute error <25%, <1 ideal
- Root relative squared <50%, 0 ideal

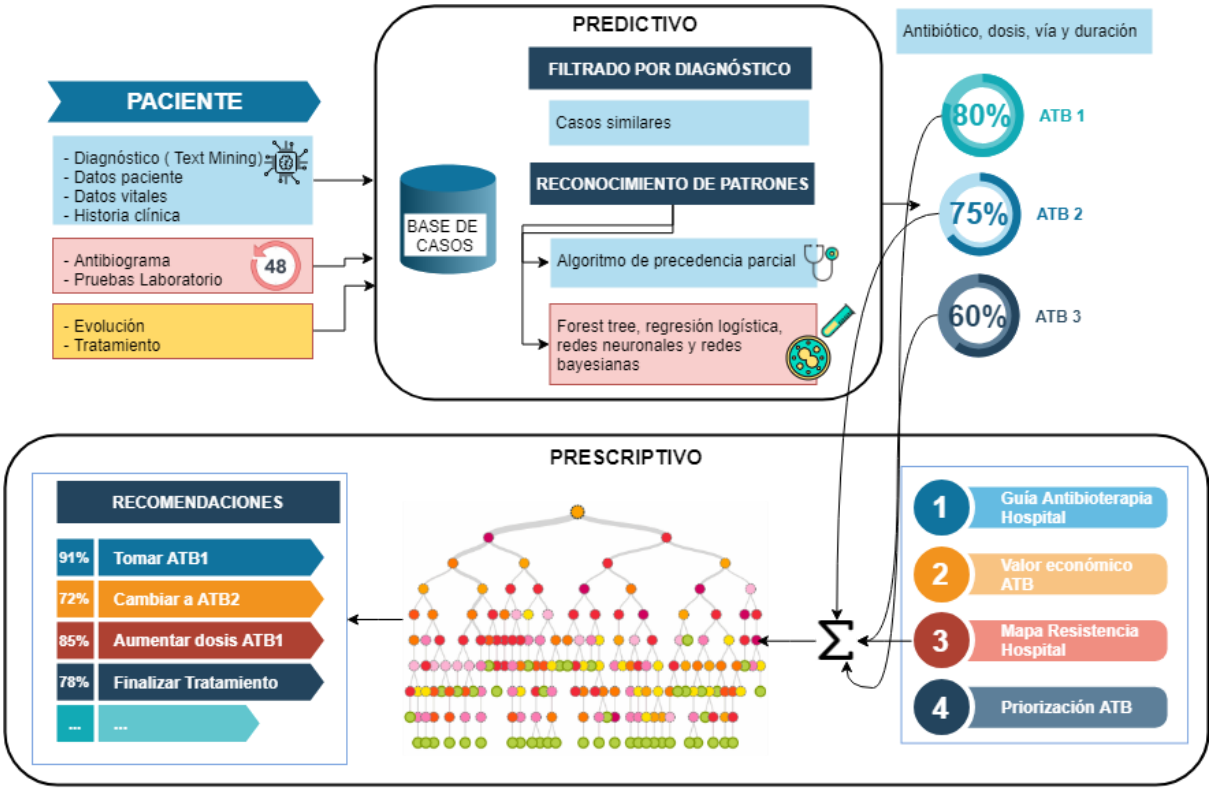
Muchas de las variables independientes del modelo las haremos categóricas, positivo o negativo , cero o uno, estableciendo para cada una un umbral para cuando es 0 o 1, y así simplificar los modelos.

- Sensible ATB X1: 1-0 (Si o No)
- Sensible ATB X2: 1-0 (Si o No)
- Sensible ATB X3: 1-0 (Si o No)
- Dolor: Si o No
- Peso paciente: 3-200
- Edad: 0-99
- Dósis: 100, 200, 300, 4000
- Días Tratamiento: 1, 2, 3 , 4, 5 ...
- Fiebre: 1-0
- Cambio de ATB: Si o No
- Alto riesgo: Si o No
- Antecedentes: Si o No
- Muerte: Si o No
- Nivel: Leve, Grave
- Otras muchas pertenecientes al campo de la medicina presentes en análisis de orina/sangre (leucocitos, plaquetas, …)

Para cada variable dependiente ejecutaré un modelo previamente entrenado y no necesariamente con la mismas independientes válidas(ATB adecuado, dosis, vía y

duración). Por ejemplo para la dosis, la variable númerica peso tendrá mucha importancia, o para el caso de la vía lo será la edad.

Finalmente el diagrama con todo completo quedaría tal que así:



Modelo Predicción antibiótico tras test rápido molecular

Ya lo mencionamos en un apartado anterior, se parte del trabajo en conjunto realizado por varios departamentos de la universidad de Québec, Canadá centrado en modelar el AMR. Donde fueron obtenidos modelos altamente precisos para 12 especies de microorganismos y 56 antibióticos. Se presenta la KOVER AMR platform de acceso libre y open source. (https://aldro61.github.io/kover-amr-platform/).

Ya existen el mercado instrumentación que realiza tales tests moleculares tales como PCR o polymerase chain, a un precio cada vez más económico, son rápidos y eficaces.

El objetivo es predecir el mejor ATB sabiendo de este una serie de genes diana y comparando con el histórico de los fenotipos de resistencia antimicrobiana (AMR) de una gran base de datos de genomas y fenotipos AMR de microorganismos (PATRIC), nosotros mejoraríamos la precisión de los modelos añadiendo todo el conocimiento de datos de paciente ya adquirido en esta etapa.

Se trata de un problema de supervisión clasificación y lo genial de todo ello es que el proyecto es open source, teniendo acceso a los modelos en python.

Papers consultados

Algoritmo reconocimiento de patrones:

http://host-article-assets.s3-website-us-east-1.amazonaws.com/ceb/5889fb885d01231a018b468c/fulltext.pdf

Clasificadores supervisados para el análisis predictivo de muerte:

http://www.informatica2013.sld.cu/index.php/informaticasalud/2013/paper/viewFile/292/212

Combinación de clasificadores supervisados:

http://www.cenatav.co.cu/doc/RTecnicos/RT%20SerieAzul_048web.pdf

Construcción de un modelo de scoring para el otorgamiento de crédito en una entidad financiera:

http://www.scielo.org.co/pdf/pece/n16/n16a10.pdf

Modelo Scoring para el otorgamiento de crédito de las pymes:

https://core.ac.uk/download/pdf/159381488.pdf

El paper de partida release 2.0 donde explican el tratamiento de los datos y el algoritmo empleado es:

https://arxiv.org/pdf/1612.01030.pdf

Junto a este otro:

https://arxiv.org/pdf/1406.1219.pdf

Planes operativos

Plan de implantación y recursos necesarios.

Sistemas on-premise.

Analizando la estructura de costes, se observa que el mayor coste se produce durante los seis primeros meses de trabajo con el hospital, correspondientes a la fase de la implantación on-premise de nuestra solución en sus instalaciones. Estos costes son los asociados al análisis de las bases de datos existentes y al empaquetado de nuestra solución para la generación de los dashboards de la Visión 360º. Se estima que para esta fase de implantación serán necesarias, aproximadamente, unas 1350 horas de personal analista.

Teniendo en cuenta la inercia en el funcionamiento de la administración pública (posibles retrasos en la contratación, burocracia, pruebas y validación por parte de Dirección y otros posibles imprevistos), se prevé que el tiempo de implantación en hospitales públicos se mueva dentro de una horquilla temporal de 6 meses (caso ideal) hasta un año natural (caso más desfavorable).

Por otra parte, es muy importante tener en cuenta que los hospitales suelen estar integrados en distintas Redes, según Comunidades Autónomas u otros aspectos (por ejemplo, Red del SAS en Andalucía, Red del Servicio Madrileño de Salud, Red de Hospitales Jesuitas, etc). Todos los hospitales (públicos o privados) pertenecientes a una misma Red utilizan la misma estructura de datos para sus historias clínicas, mismos procedimientos de diagnóstico, etc. Esto nos lleva a la conclusión de que el trabajo de implantación realizado en el primer hospital de una Red ("hospital piloto de la Red") será fácilmente replicable en los todos los restantes hospitales pertenecientes a la misma Red.

Aún así, se tiene en cuenta que los detalles en cada centro implicarán un cierto trabajo de personalización, para lo que tendremos disponible una plantilla de programadores que darán servicio y soporte tanto a la expansión a otros hospitales de una Red ya conocida, como a las labores de mantenimiento habituales en estos sistemas on-premise en los años posteriores a la implantación.

A modo de resumen, para la labores de desarrollo y mantenimiento en los sistemas de los hospitales cliente, estimamos unas necesidades de 1 analista durante un año (1.800 horas anuales según convenio), por cada hospital piloto en fase de implantación. además, por cada Red nueva que contratemos el primer año, se estiman 600 horas hombre anuales de programadores (0,33 personas en cómputo anual) por Hospital implantado, sea piloto o réplica.

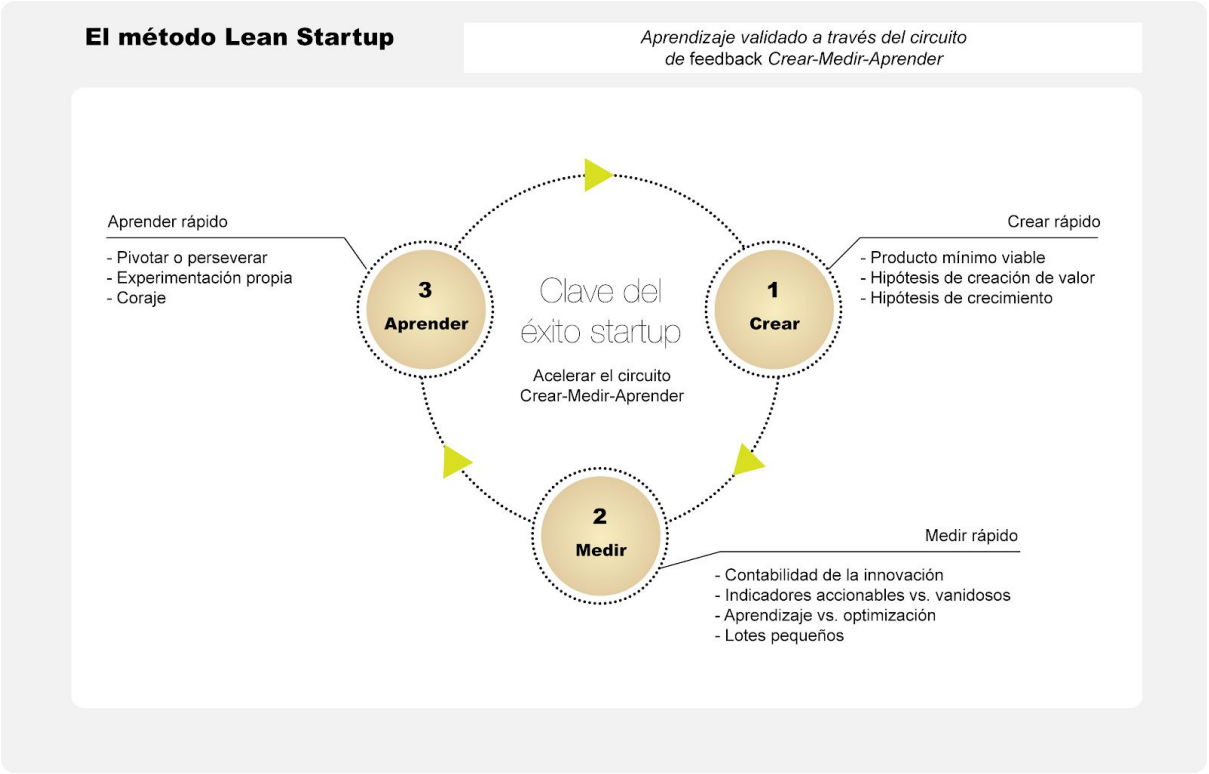
Etapa de descubrimiento y entrenamiento de modelos.

Para acortar los plazos y facilitar la evaluación de los resultados, dispondremos de un Data Scientist con experiencia hospitalaria desde el primer momento en que se comiencen los trabajos con el primer hospital piloto.

Para el desarrollo de la infraestructura sobre AWS contaremos con la experiencia y el conocimiento de nuestros CIO y CDO, que como el resto del equipo fundador de la empresa, percibirá un salario equivalente a una dedicación del 50%, desde el primer momento del proyecto y considerando que el otro 50% de salario provendrá vía beneficios futuros e incremento de valor de la empresa.

Los primeros modelos del Sistema Recomendador estarán disponibles a los 12 meses desde el comienzo de los trabajos en el hospital piloto de la Red. Esto debe ser así para que los trabajos de desarrollo, entrenamiento de los modelos y la realización de ajustes de las llamadas a la API-REST puedan iniciarse una vez transcurrido este periodo y de esta manera se pueda tener en producción el Sistema Recomendador a partir del mes 18.

En otras palabras, en el mes 18 se debe tener la primera versión de Sistema Recomendador (API-REST ML 1.0) validada y en producción, para así completar el ciclo de acuerdo con la metodología Lean Startup, y desarrollar las posteriores versiones, 2.0, 3.0 y 4.0 de nuestro Sistema Recomendador .



Métricas.

En el paso 2 debemos analizar las métricas que estemos obteniendo en nuestra empresa, en el cliente (hospitales), y en el usuario, y aprender. A priori, y sólo como ejemplo, citamos las siguientes, típicos en un modelo de suscripción como el nuestro, que deberemos ir afinando según avancemos en el conocimiento de la solución y del cliente:

IMMUNOWAR (nuestra startup)
• Ingresos totales. • Evolución del número de clientes implantados. • Cash burn rate (CBR). Es el dinero que “quemamos” mensualmente, nos indica cuántos meses faltan para quedarnos sin fondos. • Ingresos recurrentes mensuales (MRR - Monthly Recurring Revenue): Son los ingresos recurrentes obtenidos de los clientes instalados. • Uso (indicador clave de churn futuro): Número mensual de llamadas a la API. Nos ayuda además a prever si un usuario va a darse de baja. • Ingresos medios por cliente (ARPU - Average Revenue per User/Customer): Se trata de los ingresos por hospital, considerando los ingresos recurrentes mensuales más la facturación por el número de llamadas a la API. • Coste de servicio (segmentado por tipo): Costes imputables a cada cliente en concreto.
Hospitales (clientes)
• Productividad PROA: % de tratamientos con antibióticos revisados respecto del total aplicable. • Precisión del recomendador: cuantos de los cambios propuestos fueron adecuados para el paciente (mejoraron su historia). • Disminución en días del tratamiento necesario. • Estimación del número de vidas salvadas. (# y %). • Ahorro en costes (menores dosis, menos tiempo).

En cada ciclo de crecimiento deberemos decidir si seguir avanzando perfeccionando nuestro recomendador, o bien, si es necesario pivotar nuestra empresa de algún modo.

Plan de expansión

Como hemos comentado, el plan de expansión tiene dos ejes:

Eje 1: implementación de nuestra solución en el primer hospital de una Red (SAS, QUIRÓN, Jesuitas, etc) al que denominaremos hospital piloto, para posteriormente replicar la solución a los hospitales restantes de esa Red con un esfuerzo mucho menor. Esto es así debido a que, como se ha detallado anteriormente, los hospitales pertenecientes a una misma Red utilizan la misma estructura de datos para sus historias clínicas, mismos procedimientos de diagnóstico, etc.

Eje 2: introducir nuestra solución en nuevas Redes hospitalarias, gracias a la experiencia y la solvencia de los resultados obtenidos anteriormente. En la siguiente tabla mostramos el mercado existente sólo en España:

Comunidad Autónoma	Hospitales		Camas	
	Público	Privado	Público	Privado
Cataluña	67	147	15.299	19.239
Madrid	34	47	14.473	7.005
Andalucía	46	62	15.089	5.889
Canarias	15	23	4.966	2.680
Comunidad Valenciana	36	25	10.899	2.446
País Vasco	17	24	5.613	2.347
Galicia	14	24	7.444	2.341
Castilla y León	15	21	7.213	2.262
Murcia	11	16	3.275	1.575
Islas Baleares	11	13	2.584	1.332
Asturias	9	11	2.764	998
Aragón	20	9	4.367	926
Navarra	4	7	1.405	883
Cantabria	3	4	1.419	658
Castilla – La Mancha	19	9	5.247	472
Extremadura	11	8	3.587	401
La Rioja	5	2	911	137
Ceuta	1	0	252	0
Melilla	1	0	168	0
Total	**339**	**452**	**106.975**	**51.591**

Se muestra a continuación un cronograma del plan de expansión de nuestra solución en un hospital que comienza en el primer año de vida de la empresa (2019):

Si se desea implantar la solución en un nuevo cliente, por ejemplo, a principios del año 3 del proyecto (enero 2021), la solución a implantar se correspondería al paquete Visión 360º más la versión Sistema Recomendador más actualizada que se disponga (véase en el apartado *Definición de problema - Solución* los detalles y actualizaciones que contendrá cada versión del Sistema Recomendador) en el momento de la implantación, que en este ejemplo se correspondería con la versión 2.0.

Para el desarrollo del business case, consideraremos que nos introducimos en las Redes de hospitales paulatinamente y de forma escalonada, para optimizar el uso de nuestros recursos (analistas y programadores, CIO y CDO). Por tanto, el esquema de expansión que se ha considerado para modelar el business case es el siguiente:

Clientes	Años	2019	2020	2021	2022	2023
Total Clientes		1	5	12	20	32
Total clientes - piloto		1	2	1	1	1
Total clientes - réplicas		0	3	11	19	31

Donde los clientes se corresponden con el número de hospitales, que pueden ser tanto públicos como privados.

Plan de precios

El plan de precios estará en línea con los esfuerzos de implantación y con nuestros costes, pero sobre todo con el valor aportado al cliente. Con esto queremos decir que en el primer año, cuando más costes de personal se tendrán, solo podremos facturar al cliente el servicio de la Visión 360º, por lo que trabajaremos con un margen bastante ajustado.

Precios base	Implant.	Precio Implantación	Cuota mantenimiento /mes (€)	Precio llamada a la API /mes (€)
Año 1: Visión 360	Año 1	25.000,0 €	550,0 €	- €
Año 2: V360 + Recomendador	Año 2	8.000,0 €	1.200,0 €	1,8 €

Una vez el Sistema Recomendador se encuentre en producción podremos trabajar con márgenes mayores, lo que generará una curva de ingresos que se irá despegando de los costes de producción.

Para calcular los ingresos por cliente, estimaremos que:

- En al año 1 facturaremos la implantación, más 6 cuotas mensuales de la Visión 360º funcionando.

Año 1: Visión 360	Descuento comercial	Precio Implantación, mes 6	Precio mantenimiento, mes	Estimación de llamadas/mes	Ingresos año 1	Mix tamaños	Ingresos medios año 1
Hosp. Pequeño <> 200 camas	20%	20.000 €	440 €	500	22.640,0 €	20%	
Hosp. mediano <> 500 camas	10%	22.500 €	495 €	1600	25.470,0 €	30%	**26.319,0 €**
Hosp. grande <> 1000 camas	0%	25.000 €	550 €	3000	28.300,0 €	50%	

- En el año 2 facturaremos 6 cuotas mensuales de la visión 360º, la implantación del recomendador, más otras 6 cuotas mensuales del recomendador. Supondremos un número de llamadas a la API (nuestro recomendador ya estarà funcionando y ajustado), durante seis meses, al precio de 1,8 €/llamada. Estimamos el número de llamadas al mes en función del tamaño del Hospital).

Año 2: V360 + Recomendador	Descuento comercial	Precio Implantación, mes 6	Precio mantenimiento, mes	Estimación de llamadas/mes	Ingresos año 2	Mix tamaños	Ingresos medios año 2
Hosp. Pequeño <> 200 camas	20%	6.400 €	960 €	500	20.200,0 €	20%	
Hosp. mediano <> 500 camas	10%	7.200 €	1.080 €	1600	33.930,0 €	30%	**39.669,0 €**
Hosp. grande <> 1000 camas	0%	8.000 €	1.200 €	3000	50.900,0 €	50%	

- En el año 3 y siguientes en cada hospital facturaremos 12 cuotas mensuales del servicio completo más el número estimado de llamadas mensuales durante 12 meses.

Es a partir del año 3 cuando podemos reflejar en el precio de nuestro servicio todo el valor que estamos aportando con el recomendador. como vemos en el siguiente cuadro resumen:

Ingresos estimados	Descuento comercial	Mix tamaños	Ingresos año 1	Ingresos medios año 1	Ingresos año 2	Ingresos medios año 2	Ingresos año 3 y siguientes	Ingresos medios año 3
Hosp. Pequeño <> 200 camas	*20%*	20%	22.640 €		20.200 €		22.320 €	
Hosp. mediano <> 500 camas	*10%*	30%	25.470 €	**26.319,0 €**	33.930 €	**39.669,0 €**	47.520 €	**58.320,0 €**
Hosp. grande <> 1000 camas	*0%*	50%	28.300 €		50.900 €		79.200 €	

Hemos considerado que en el mix tendremos un mayor número de clientes grandes, lo que parece coherente con el servicio que prestaremos.

Plan de marketing

El plan de marketing estará centrado principalmente en la difusión y alimentación de información en ámbito nacional, especialmente durante los primeros años del proyecto, ya que todos los hospitales cliente con los que se estima trabajar en los 5 primeros años son españoles. No obstante, es vital mantenerse actualizado con los últimos avances que se

realicen a nivel internacional, especialmente en Europa, dadas las grandes oportunidades de difusión de información y colaboración que se ofrecen desde la Comisión Europea.

El CMO (Chief Marketing Officer) será el principal responsable del desarrollo de los planes de marketing de la empresa. Las herramientas para la difusión de información, canales de venta y marketing a utilizar pueden resumirse en:

- **Visitas a clientes:** visitas personales por parte del CMO a responsables de Redes Hospitalarias, equipos médicos, equipos PROA, etc., para dar a conocer nuestras soluciones e informar de los grandes beneficios derivados de su implantación en el hospital.
- **Página web**: landing page en la que exponer información sobre nuestra empresa, equipo humano, soluciones que aportamos, actividad desarrollada (papers, artículos publicados, congresos…), clientes con los que colaboramos y sus testimonios y datos para el contacto. Serviría además para reforzar nuestro logo y marca *INMUNOWAR*. Se muestra a continuación un ejemplo de landing page:

- **Revistas y publicaciones científicas:** el uso de estos recursos está ampliamente extendido y es muy influyente en la comunidad científica. Estos canales de información nos permitirían compartir nuestro conocimiento y experiencia en el sector de la Medicina y dar a conocer la empresa, las soluciones innovadoras que aportamos y nuestras colaboraciones con hospitales y clientes.

  Se muestran a continuación algunos ejemplos de las revistas científicas más influyentes de la actualidad, a nivel nacional e internacional:

- Enfermedades Infecciosas y Microbiología Clínica

Esta revista forma parte de Elsevier (http://www.elsevier.es/es), "*empresa de análisis de información global que asiste a instituciones y profesionales en el progreso de la ciencia, cuidados avanzados en materia de salud, así como mejorar la ejecución de los mismos para el beneficio de la humanidad*"

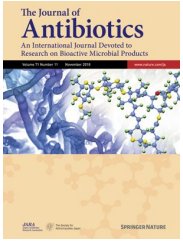
"La revista Enfermedades Infecciosas y Microbiología Clínica es el órgano de expresión de la Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica (SEIMC) (Spanish Society of Infectious Diseases and Clinical Microbiology), tiene como objetivo dar respuesta a los retos que actualmente plantea todo lo relacionado con la patología infecciosa, desde el punto de vista clínico, microbiológico y de salud pública.

La revista está incluida en Science Citation Index Expanded, Medline/Pubmed y SCOPUS".

Factor de Impacto 2017**:** 1,707

- The Journal of Antibiotics

Perteneciente a la famosa revista internacional Nature (https://www.nature.com/), que publica las mejores investigaciones revisadas por expertos en todos los campos de la ciencia y la tecnología y que proporciona noticias e interpretaciones rápidas, autorizadas, perspicaces y reveladoras de las tendencias actuales y futuras que afectan a la ciencia, a los científicos y al público en general.

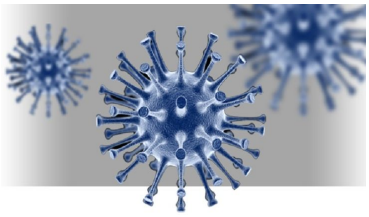


"*La investigación sobre antibióticos y otros productos microbianos es fundamental para la batalla mundial contra el creciente problema de la resistencia a los antibióticos. Además del modelo tradicional de investigación y desarrollo para encontrar medicamentos de reemplazo para abordar este problema, existen otras nuevas estrategias importantes para abordar la resistencia que incluyen la investigación de productos microbianos bioactivos.*

The Journal of Antibiotics forma parte de esta batalla contra la resistencia al alentar y publicar artículos de investigación originales que informan hallazgos de gran importancia biológica en el campo multidisciplinario de los antimicrobianos. Alentamos específicamente los artículos que representan avances importantes en el campo de la investigación antimicrobiana, incluidos los antibióticos descubiertos recientemente, los productos microbianos bioactivos y las nuevas estrategias para combatir la resistencia a los antibióticos. The Journal of Antibiotics proporciona una plataforma dinámica para comunicar la investigación sobre productos microbianos y enfoques antimicrobianos a través de la inclusión de revisiones detalladas y actualizadas, comunicaciones breves y cartas al editor".

2017 Impact Factor: 2.033 (2017 Journal Citation Reports® Science Edition (Clarivate Analytics, 2018))

- Journal of Immunology and Infectious Diseases (JIID)



The Journal of Immunology and Infectious Diseases (JIID, por sus siglas en inglés) es una revista multidisciplinaria de acceso abierto, arbitrada, que fomenta la investigación precisa y contribuye de manera significativa al avance del conocimiento para la aplicación inmunológica en el tratamiento de diversas enfermedades infecciosas. Su objetivo es publicar la fuente de información más completa y confiable sobre los descubrimientos y desarrollos actuales en forma de investigación, revisión, informes de casos y artículos breves de comunicación. Todos los artículos son revisados por pares y publicados bajo la guía de los miembros de nuestro Consejo Editorial".

- International Journal of Environmental Research and Public Health

"The International Journal of Environmental Research and Public Health (ISSN 1660-4601; CODEN: IJERGQ; ISSN 1661-7827 para edición impresa) es una revista interdisciplinaria de acceso abierto y arbitrada que se publica mensualmente online a través de MDPI Open Access (https://www.mdpi.com/).

Abarca las Ciencias e Ingeniería Ambientales, Salud Pública, Salud Ambiental, Higiene Ocupacional, Investigación Económica y Salud Global, etc. La Sociedad Internacional de Médicos para el Medio Ambiente (ISDE) y la Sociedad Italiana de Medicina Ambiental (SIMA) están afiliadas a IJERPH y sus asociados.

Acceso abierto gratuito para lectores, con cargos de procesamiento de artículos (APC) pagados por los autores o sus instituciones".

Impact Factor: 2.145 (2017); Factor de impacto de 5 años: 2.608 (2017)

- **Seminarios y Congresos científicos:** asistencia a eventos del sector para darnos a conocer y estar al tanto de los últimos avances y novedades (vigilancia tecnológica). Al igual que en el caso de las revistas científicas, existen multitud de Congresos, Conferencias y Seminarios relacionados con el sector de los antibióticos. Se exponen a continuación algunos de los más relevantes:

  - Congreso Nacional de la Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica

    https://www.seimc2019.org/

  - World Congress on Antibiotics

    https://www.meetingsint.com/conferences/antibiotics

"El objetivo del World Congress on Antibiotics es reunir a científicos, investigadores, maestros, delegados de negocios, estudiantes y asociados de investigación para compartir su experiencia y conocimiento, y las investigaciones en las que están trabajando. El tema principal de la conferencia es mejorar la salud mundial con antibióticos.

El World Congress on Antibiotics incorpora conferencias plenarias, conferencias magistrales de personalidades famosas de todo el mundo, además de presentaciones de pósteres, sesiones de jóvenes investigadores, simposios, talleres y exposiciones. Se enfoca en una amplia gama de temas que incluyen los siguientes: terapia con antibióticos, resistencia y prevención de antibióticos, farmacología de antibióticos, usos médicos, antibióticos para enfermedades emergentes y reemergentes, antibióticos en oncología, interacción y efectos secundarios, el enfoque de la próxima generación de antibióticos, la resistencia antimicrobiana genética, la alternativa a los antibióticos".

Estudio económico financiero

Plan de Ingresos

Es fundamental seguir la expansión que hemos diseñado, para optimizar nuestros recursos y el flujo de ingresos, que quedaría de este modo:

Clientes	Años	2019	2020	2021	2022	2023
Total Clientes		**1**	**5**	**12**	**20**	**32**
Total clientes - piloto		1	2	1	1	1
Total clientes - réplicas		0	3	11	19	31

Ingresos	Años	2019	2020	2021	2022	2023
Red hospitalaria A - piloto		12.090 €	26.625 €	43.020 €	43.020 €	43.020 €
Red hospitalaria A - réplica			12.090 €	26.625 €	43.020 €	43.020 €
Red hospitalaria A - réplica			12.090 €	26.625 €	43.020 €	43.020 €
Red hospitalaria A - réplica				12.090 €	26.625 €	43.020 €
Red hospitalaria A - réplica				12.090 €	26.625 €	43.020 €
Red hospitalaria A - réplica					12.090 €	26.625 €
Red hospitalaria A - réplica					12.090 €	26.625 €
Red hospitalaria A - réplica						12.090 €
Red hospitalaria A - réplica						12.090 €
Red hospitalaria A - réplica						
Red hospitalaria B - piloto			12.090 €	26.625 €	43.020 €	43.020 €
Red hospitalaria B - réplica				12.090 €	26.625 €	43.020 €
Red hospitalaria B - réplica				12.090 €	26.625 €	43.020 €
Red hospitalaria B - réplica					12.090 €	26.625 €
Red hospitalaria B - réplica					12.090 €	26.625 €
Red hospitalaria B - réplica						12.090 €
Red hospitalaria B - réplica						12.090 €
Red hospitalaria B - réplica						
Red hospitalaria C - piloto			12.090 €	26.625 €	43.020 €	43.020 €
Red hospitalaria C - réplica				12.090 €	26.625 €	43.020 €
Red hospitalaria C - réplica				12.090 €	26.625 €	43.020 €
Red hospitalaria C - réplica					12.090 €	26.625 €
Red hospitalaria C - réplica						12.090 €
Red hospitalaria C - réplica						12.090 €
Red hospitalaria D - piloto				12.090 €	26.625 €	43.020 €
Red hospitalaria D - réplica					12.090 €	26.625 €
Red hospitalaria D - réplica					12.090 €	26.625 €
Red hospitalaria D - réplica						12.090 €
Red hospitalaria D - réplica						12.090 €
Red hospitalaria E - piloto					12.090 €	26.625 €
Red hospitalaria E - réplica						12.090 €
Red hospitalaria E - réplica						12.090 €
Red hospitalaria E - réplica						12.090 €
Red hospitalaria F - piloto						12.090 €
Total ingresos - pilotos		12.090 €	50.805 €	108.360 €	155.685 €	172.080 €
Total ingresos - réplicas		- €	24.180 €	125.790 €	342.510 €	702.240 €
TOTAL Ingresos		**12.090 €**	**74.985 €**	**234.150 €**	**498.195 €**	**874.320 €**

Plan de Costes

Ls inversiones que estimamos necesarias son muy pequeñas, mobiliario para la oficina y un equipo de trabajo para nuestro Data Scientist:

Inversiones	
Maquinas / HW	5.000,00 €
Mobiliario	1.500,00 €
TOTAL INVERSIONES	6.500,00 €
Amortización	1.300,00 €
Años	5,00 €

Los gastos de comercialización y generales (un modesto alquiler de la oficina):

Gastos de comercialización y otros		20.000 €	20.400 €	20.808 €	21.224 €	21.649 €
Gastos generales		3.600 €	3.672 €	3.745 €	3.820 €	3.897 €

Para soportar este crecimiento en clientes, tendremos unos recursos humanos proyectados:

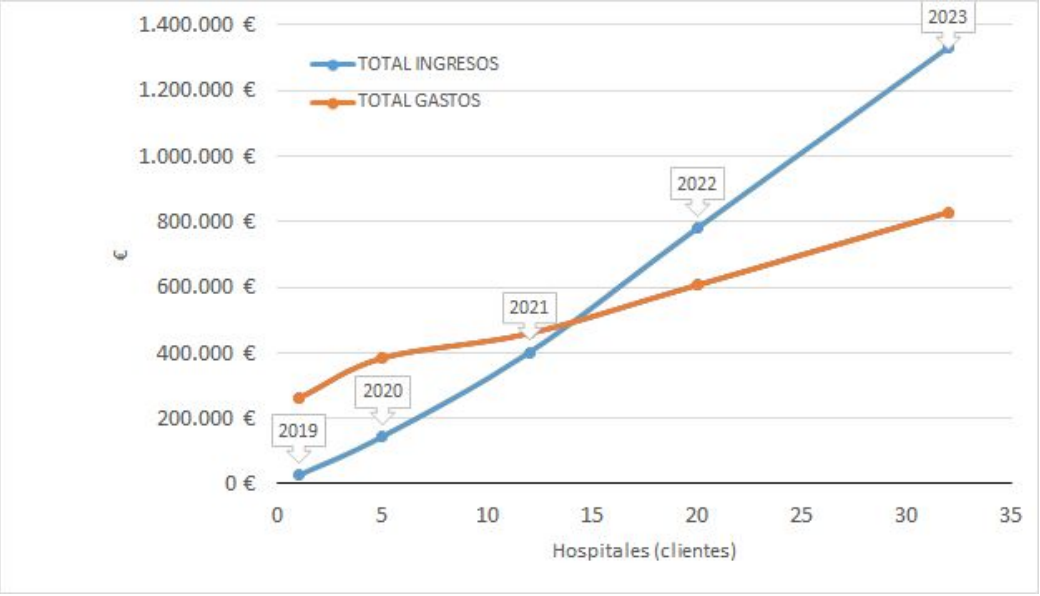
Recursos Humanos	Años	2019	2020	2021	2022	2023
CEO		1	1	1	1	1
CDO		1	1	1	1	1
CTO		1	1	1	1	1
CMO		1	1	1	1	1
COO		1	1	1	1	1
CFO		1	1	1	1	1
DataScientist		1	1	1	1	1
Analistas (1 por piloto)	1	1,00	2,00	1,00	1,00	1,00
Programadores (0,33 /Hospital)	0,33	0,33	1,65	3,96	6,60	10,56

Y unos costes de software:

Costes de SW	Años	2019	2020	2021	2022	2023
Total Costes SW		**9.677 €**	**29.774 €**	**64.673 €**	**106.047 €**	**168.463 €**
OTRO SW, MANTENIMIENTO		1.824 €	5.581 €	12.335 €	20.324 €	32.577 €
API para consulta		2.376 €	11.881 €	28.514 €	47.524 €	76.039 €
SAGEMAKER IRLANDA		5.477 €	12.312 €	23.823 €	38.199 €	59.847 €

Punto muerto

Con esto obtenemos el punto muerto poco después de iniciado el año 2021:

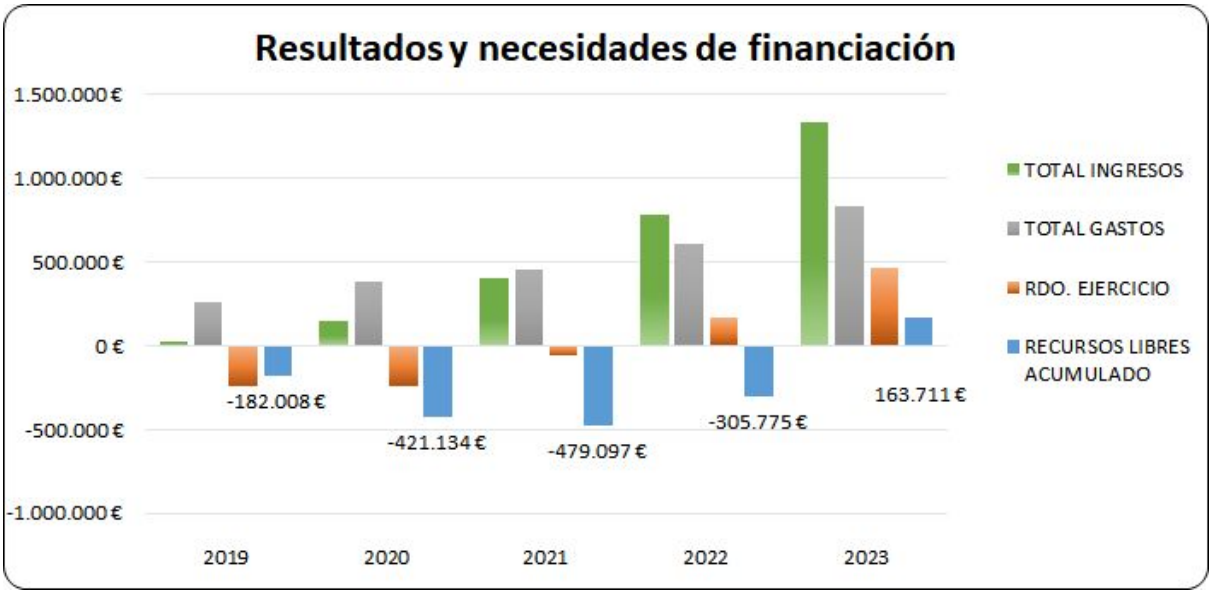


Cuenta de resultados.

Nuestra cuenta de resultados queda:

CUENTA DE RESULTADOS - IMMUNOWAR	2019	2020	2021	2022	2023
TOTAL INGRESOS	26.319,00 €	144.945,00 €	401.229,00 €	779.835,00 €	1.333.020,00 €
Total ingresos - pilotos	26.319,00 €	92.307,00 €	163.977,00 €	214.629,00 €	233.280,00 €
Total ingresos - réplicas	0,00 €	52.638,00 €	237.252,00 €	565.206,00 €	1.099.740,00 €
TOTAL GASTOS	261.826,88 €	384.071,42 €	459.192,19 €	606.512,98 €	828.962,80 €
Aprovisionamientos (SW)	9.676,88 €	29.774,42 €	64.672,51 €	106.047,29 €	168.462,70 €
Gastos de personal	228.550,00 €	330.225,00 €	369.966,24 €	475.421,18 €	634.954,71 €
Otros gastos de explotación	23.600,00 €	24.072,00 €	24.553,44 €	25.044,51 €	25.545,40 €
EBITDA	-235.507,88 €	-239.126,42 €	-57.963,19 €	173.322,02 €	504.057,20 €
Amortización Inmovilizado	-1.300,00 €	-1.300,00 €	-1.300,00 €	-1.300,00 €	-1.300,00 €
Imputación de subvenciones de	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
Provisiones, deterioros	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
Resultado por enajenación de	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
RDO. EXPLOTACIÓN	**-236.807,88 €**	**-240.426,42 €**	**-59.263,19 €**	**172.022,02 €**	**502.757,20 €**
Resultado Financiero	0,00 €	0,00 €	0,00 €	0,00 €	0,00 €
RDO. ANTES DE IMPUESTOS	**-236.807,88 €**	**-240.426,42 €**	**-59.263,19 €**	**172.022,02 €**	**502.757,20 €**
Impuesto de sociedades (compensando)	0,00 €	0,00 €	0,00 €	0,00 €	-34.570,43 €
RDO. EJERCICIO	**-236.807,88 €**	**-240.426,42 €**	**-59.263,19 €**	**172.022,02 €**	**468.186,77 €**
RECURSOS LIBRES EN EL PROYECTO	-182.007,88 €	-239.126,42 €	-57.963,19 €	173.322,02 €	469.486,77 €
RECURSOS LIBRES ACUMULADO	-182.007,88 €	-421.134,30 €	-479.097,50 €	-305.775,48 €	163.711,29 €

y de un modo gráfico:



Subvenciones

Existen diferentes programas de subvención pública. Las más relevantes son las que ofrece la Comisión Europea. Se muestran a continuación las convocatorias más interesantes de aplicación a nuestro proyecto:

<u>SME Instrument:</u>

https://ec.europa.eu/programmes/horizon2020/en/h2020-section/sme-instrument

"Las pequeñas y medianas empresas que tienen su sede en la UE o están establecidas en un país asociado a Horizonte 2020 ahora pueden obtener financiación y apoyo de la UE para proyectos innovadores de innovación con un potencial de creación de mercado bajo el nuevo instrumento de PYME, que se implementa como parte de El Consejo Europeo de Innovación (EIC) piloto. El instrumento para pymes impulsará el rápido crecimiento de la empresa y la innovación creadora de mercado gracias a la financiación por etapas y la aceleración de los servicios de aceleración de negocios".

"El SME Instrument ofrece a las pequeñas y medianas empresas lo siguiente:

- *Subvenciones de innovación empresarial para fines de evaluación de viabilidad (fase I opcional): 50.000 EUR (suma global) por proyecto (70% del costo total del proyecto);*
- *Subvenciones de innovación empresarial para fines de desarrollo y demostración de la innovación (posible fase II): un monto en el rango indicativo de EUR 500,000 y 2,5 millones (70% del costo total del proyecto como regla general);*
- *Capacitación empresarial gratuita (opcional) para respaldar y mejorar la capacidad de innovación de la empresa y ayudar a alinear el proyecto con las necesidades comerciales estratégicas;*
- *Acceso a una amplia gama de otros servicios de aceleración de negocios y acceso facilitado a financiamiento de riesgo, para facilitar la explotación comercial de la innovación* ".

Working Programme 2018-2020, Health, demographic change and wellbeing

http://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-health_en.pdf#page=23&zoom=100,0,660

Programa colaborativo en el que presentar proyectos a las diferentes convocatorias que casen con la temática y objetivos para obtener una subvención por parte de la Unión Europea. En general, el proyecto a presentar debe participar un consorcio de empresas (PYMES, gran empresa), organismos público, universidades y/o Centros de Investigación de diferentes países de Europa.

Algunas convocatorias de interés en el marco de H2020 Health, demographic change and wellbeing a las que podría aplicar nuestro proyecto:

- **SC1-BHC-13-2019: Mining big data for early detection of infectious disease threats driven by climate change and other factors**

"Reto específico: una serie de factores es responsable de la (re) aparición de amenazas de enfermedades infecciosas, incluida la resistencia a los antimicrobianos, la alteración de la epidemiología y la propagación de la enfermedad en un entorno global cambiante. Estos incluyen factores como el cambio climático y los impactos ambientales asociados, el

crecimiento de la población, la urbanización no planificada y la alta movilidad, así como la cría de animales o las prácticas agrícolas intensivas.

Al mismo tiempo, las herramientas para el diagnóstico y la vigilancia de enfermedades infecciosas están evolucionando rápidamente, lo que permite un diagnóstico cada vez más preciso en un tiempo cada vez más corto. El uso de la secuenciación de próxima generación combinada con datos de vigilancia, registros de salud y datos sociales de fuentes informales / no tradicionales (por ejemplo, redes sociales) promete mejorar la salud individual y de la población.

Las tecnologías avanzadas de IT actuales ofrecen la oportunidad de integrar estos grandes conjuntos de datos y podrían permitir el tratamiento rápido y personalizado de los pacientes infectados y reforzar la detección, el seguimiento y el control de los brotes de enfermedades infecciosas.

Alcance: Se espera que las propuestas desarrollen:
1. la tecnología para permitir la puesta en común, el acceso, el análisis y el intercambio de datos relevantes, incluida la secuenciación de la próxima generación;
2. las innovadoras metodologías de bioinformática y modelado que permiten la creación de modelos y mapeo de riesgos; y
3. las herramientas analíticas para la alerta temprana, la evaluación de riesgos y el monitoreo de (re) emerger amenazas de enfermedades infecciosas.

La Comisión considera que las propuestas que soliciten una contribución de la UE de entre ***12 y 15 millones de euros*** *permitirían abordar este desafío específico de manera adecuada".*

- **SC1-HCO-08-2018: Creation of a European wide sustainable clinical research network for infectious diseases**

"Reto específico: las enfermedades infecciosas representan una grave amenaza para la salud mundial. Las epidemias emergentes, las pandemias y el aumento de los niveles de resistencia a los antimicrobianos requieren una respuesta sólida y coordinada para proteger a los ciudadanos en Europa, como se indica en el Plan de Acción Europeo de Salud Única contra la AMR.

Existe la necesidad de establecer una red de investigación clínica en toda Europa que tenga la capacidad y la capacidad de inscribir directamente a pacientes con enfermedades infecciosas, para aumentar la eficacia de las pruebas y el desarrollo de nuevas estrategias y terapias diagnósticas, preventivas y / o terapéuticas. Esto debería permitir generar pruebas rigurosas para mejorar el diagnóstico, la prevención y el tratamiento de infecciones y para responder mejor a las amenazas de enfermedades infecciosas, y contribuir al objetivo del G7 con respecto a la necesidad de establecer una red mundial de estudios clínicos sobre resistencia a los medicamentos que proporcione acceso a una gran Infraestructura de investigación clínica para el diseño, la coordinación y la realización de ensayos clínicos y estudios en cooperación con las redes e infraestructuras de expertos globales existentes.

*La Comisión considera que una propuesta que solicite una contribución de la UE de entre **2 y 3 millones de euros** permitiría abordar este desafío específico de manera adecuada ".*

- **SC1-HCO-12-2018: Innovation in healthcare - a CSA towards using pre-commercial procurement and public procurement of innovative solutions in healthcare systems**

"Reto específico: las soluciones innovadoras para el cuidado de la salud tienen el potencial de mejorar el cuidado del paciente en el ámbito del cuidado de la salud en Europa. Los principios de atención integrada permiten que la atención de los pacientes esté mejor coordinada y sea planificada conjuntamente por profesionales de la salud y asistencia social a través de servicios preventivos y curativos de relevancia vertical y horizontal.

Para responder a la cambiante organización de la atención y apoyar la transición de los servicios hospitalarios hacia un modelo de atención integral centrado en el paciente, se alienta a los proveedores de atención médica a unir sus fuerzas y crear una demanda de tales innovaciones a través de la contratación pública conjunta, sirviendo al triple objetivo de la atención médica: mejor atención y experiencia, mejores resultados de atención y atención más eficiente.

La implementación de diagnósticos oportunos y correctos para enfermedades infecciosas que acelerarán la identificación de patógenos causantes de enfermedades infecciosas causantes, la resistencia y la susceptibilidad a los medicamentos es crucial para adaptar el tratamiento antimicrobiano a fin de garantizar el uso apropiado de medicamentos antimicrobianos y reducir las prescripciones innecesarias.

Alcance: El objetivo de esta acción de coordinación y apoyo (CSA) es crear un consorcio a nivel europeo de proveedores de atención médica y proveedores públicos en el sector de la salud y la asistencia social que definan conjuntamente las necesidades de compras no cubiertas para implementar soluciones innovadoras en la atención médica.

*La Comisión considera que una propuesta que solicite una contribución de la UE entre **1,5 y 2 millones de euros** permitiría abordar este desafío específico de manera adecuada".*

Valoración de la empresa y necesidades de fondos.

Utilizaremos como método de valoración el de la dilución. Dilución es el "control" que se pierde a nivel accionarial de la empresa cuando entran nuevos socios.

Necesidades de financiación e hitos de control.						
Hito	Año 0	2019	2020	2021	2022	2023
Fondos propios	60.000 €					
Necesidades de financiación	182.008 €	239.126 €	57.963 €			
Recursos libres (caja)					173.322 €	469.487 €
Fecha	ene.-19	ene.-20	ene.-21	ene.-22	dic.-22	dic.-23
Valoración de la empresa (EV)	800.000 €	2.100.000 €				7.500.000 €
EV / EBITDA						14,88
Aportación inversionistas ronda A	**200.000 €**					
Valoración post-money	1.000.000 €					
Dilución	20,0%					
Aportación inversionistas ronda B		**400.000 €**		***x 6,30***		
Valoración post-money	***x 1,76***	2.500.000 €				
Dilución		16,0%		***x 3,00***		
Valor para los inversionistas ronda A		352.800 €				1.260.000 €
Valor para los inversionistas ronda B						1.200.000 €

Los 6 socios fundadores aportaremos 60.000 € en enero de 2019, y solicitamos una aportación de 200.000 €, que valoraremos como un 20% del capital, valorando la empresa en 800.000 € (pre-money).

En enero de 2020 debemos tener el primer cliente en producción con la visión 360 y haber obtenido 26.319 € de ingresos. En esas fechas necesitaremos una nueva aportación esta vez de 400.000 €, que valoraremos como un 1% del capital, valorando la empresa en 2.100.000 € (pre-money).

Con estas previsiones, podemos valorar nuestra empresa en el año 5 en 7,5 MM €, 14.88 veces el EBITDA, ratio común para una empresa aún en fuerte expansión como IMMUNOWAR. Los inversionistas de la ronda A pueden obtener entonces un multiplicador 6,30 por su inversión, y los inversionistas B un multiplicador 3,00, si optasen por salir del capital en ese momento.

Por tanto, los socios fundadores, que inician la sociedad con una aportación inicial de 60.000 €, se habrán diluido hasta el 100% x 80% x 84%, esto es, hasta el 67,2%.

Los balances resultantes pueden verse en el anexo de cálculo.

Líneas de Negocio Futuras y Nuevos Mercados

Líneas de negocio futuras

Lanzamiento de un servicio de consultas al Sistema Recomendador sin necesidad de implantación on-premise en el hospital, directamente mediante llamadas a la API-REST. Esto permitiría llevar nuestra solución fácilmente y de forma inmediata a nuevos hospitales.

Este servicio sería monetizado mediante un paquete de suscripción mensual y cobro por llamadas a un precio más elevado que los ofrecidos a hospitales en los que estemos

implantados on.premise, para compensar que no se extraerían datos de este tipo de clientes.

Esta posible línea de negocio futura permitiría un rápido aumento de los ingresos y ofrecer una prueba real de los servicios que ofrece nuestro Sistema Recomendador. Sin embargo, es importante destacar que este servicio se lanzaría una vez se tuviera un modelo de Sistema Recomendador muy robusto y bien desarrollado, así como una base de datos suficientemente grande y representativa, de forma que se asegure que la prescripción que se hace es correcta y fiable, lo que podría darse sobre el año 3 o 4 de proyecto.

Nuevos mercados

Considerando el actual grado de desarrollo de las herramientas informáticas los hospitales (especialmente en hospitales públicos), se puede afirmar que la digitalización es una necesidad en este sector. Por tanto, nuestra solución puede ser ampliada y aplicada a otros ámbitos hospitalarios para la recomendación y validación de prescripciones diferentes a enfermedades infecciosas.

Conclusiones

Queremos remarcar como conclusiones las siguientes ideas:

- Es un producto reclamado por los PROAs consultados.
- Es un producto viable técnicamente.
- Es un modelo de negocio escalable rápidamente.