



Universitat Oberta
de Catalunya

Máster universitario de Ciencia de Datos

Prueba de Evaluación Continua – PEC1

**Fundamentos de la Ciencia de Datos – Bloque 1 ¿Ciencia
en los Datos?**

Autor:

Mario Ubierna San Mamés

Índice de Contenido

Índice de Contenido	3
Índice de ilustraciones	4
1. Enunciado	5
1.1. Pregunta 1	5
1.1.1. ¿Qué obstáculos y retos comporta?	5
1.1.2. ¿Qué limitaciones tiene?	6
1.1.3. ¿Qué beneficios representa?	7
1.1.4. ¿Qué impacto ha tenido durante la pandemia?	8
1.2. Pregunta 2	9
1.2.1. ¿Qué conocimientos previos y formación deben tener?	9
1.2.2. ¿Qué habilidades deben poseer?	11
1.2.3. ¿Qué buenas prácticas deben tener en cuenta?	12
1.3. Pregunta 3	13
1.3.1. Describe el ciclo de vida de los datos en <i>Covid data save lives</i>	13
1.3.2. ¿Cómo es la pirámide D-I-K-W asociada al proyecto?	14
1.4. Pregunta 4	16
1.4.1. ¿Podemos asegurar que es una organización orientada al dato?	16
1.4.2. ¿En qué fase del modelo de madurez se encuentra?	17
2. Bibliografía	19

Índice de ilustraciones

Ilustración 1 - Pirámide DIKW [11].....	15
---	----

1. Enunciado

1.1. Pregunta 1

Argumenta cómo la filosofía del Open Data impacta en los avances del sector de la salud:

1.1.1. ¿Qué obstáculos y retos comporta?

Lo primero de todo cabe mencionar lo que se entiende por *Open Data*, los datos abiertos u *open data* son aquellos datos externos, es decir, que no están bajo el dominio de nuestra organización, los cuales pueden ser utilizados con motivos profesionales o personales sin ánimo de lucro, ya que son datos sin derechos de autor, patentes o similares.

La práctica del *open data* va adquiriendo cada vez más importancia ya que, gracias a la liberación de los datos por parte de empresas privadas u organizaciones públicas, se consigue obtener mucha información útil, la cual de no ser así tendríamos dificultades para poder realizar determinados estudios.

A día de hoy, el sector de la salud sigue siendo uno de los sectores más atrasados respecto a las tecnologías usadas, y con ello me estoy refiriendo a la explotación de los datos. Por lo tanto, poco a poco estas empresas del ámbito sanitario se van digitalizando con el fin de mejorar sus servicios.

Como resultado de dicha digitalización estas empresas se encuentran con una serie de obstáculos [1], los cuales se definen a continuación:

- **Barreras institucionales, burocráticas y culturales:** cuando hablamos de *open data* estamos hablando de algo que hasta hace poco en este sector no existía y es la transparencia en los datos, es decir, que la información no sea privada. Este cambio de percepción de lo privado a lo público no es fácil, supone un gran cambio cultural y como bien se sabe, realizar estos tipos de cambios dentro de instituciones, organizaciones ya sean del Estado o en este caso en el sector de la salud requieren tiempo.

- **Barrera en la protección de datos:** este punto junto con el anterior son las mayores barreras que existe dentro del *open data*, pero dentro del sector de la salud ésta tiene una gran importancia, ya que estamos tratando con datos sensibles, es decir, la liberación sin consentimiento de dicha información puede perjudicar a ciertas personas. Todo esto nos conlleva a que hay una gran desconfianza por parte de los pacientes en este caso, ya que no saben si se va a tratar con delicadeza su información personal.
- **Barreras relacionadas con estándares técnicos e interoperabilidad:** otro obstáculo que presenta los proyectos de *open data* es que no se garantice la efectividad del uso de los datos, es decir, que se pueda distribuir dichos datos de una forma rápida y sencilla entre diferentes organizaciones e individuos.

Aunque hay una serie de barreras u obstáculos a la hora de abrir los datos, también podemos observar las mejoras que pretenden dichos proyectos sanitarios gracias a sus retos [1]:

- **Reforzar la prevención y la salud pública:** tiene como objetivo realizar una medicina que no busque el solucionar problemas, en este caso enfermedades, sino que su mayor reto sea el anticiparse a dichos problemas. Por otro lado, también se quiere conseguir el retraso de enfermedades mejorando así la calidad de vida de las personas.
- **Una sanidad continuada y de calidad:** se pretende que se pueda hacer uso de este tipo de sanidad durante un periodo largo de tiempo y que éste sea de la mayor calidad posible, es por ello que, para cumplir este objetivo se debe tener un gran volumen de datos de las personas.
- **Mejorar la coordinación socio-sanitaria:** es decir, conseguir la información entre diferentes organizaciones, departamentos o individuos se realice de la forma más efectiva posible, para así poder garantizar y mejorar la calidad en la sanidad.

1.1.2. ¿Qué limitaciones tiene?

Respecto a las limitaciones que puede presentar el *open data* dentro del sector de la sanidad considero que no hay limitaciones técnicas en sí, es decir, las herramientas/tecnologías que tenemos hoy en día nos permitirían hacer uso de una forma fácil de dichos datos, tanto a la hora de manipularlos como compartirlos.

Por otro lado, sí que veo un problema respecto al cómo poder almacenar, manipular y publicar dichos datos dentro de un marco legislativo. Con esto me refiero a que el principal inconveniente es el hecho de que estamos “atados” de pies y manos respecto

a la privacidad de dicha información, no quiero decir que sea un problema en sí ya que siempre hay que garantizar la privacidad de los datos y sobre todo si éstos son sensibles, lo cual sucede en el ámbito sanitario. Sin embargo, tal y como he podido leer en uno de los recursos proporcionados [1], las leyes vigentes que hay a día de hoy en nuestro país, España, estás un poco anticuadas ya que buscan mantener la privacidad desde un punto de vista tradicional, es decir, asegurarse que la información es privada entre el paciente y el conjunto de médicos.

Como bien he mencionado se tiene esta limitación debido a que cuando surgieron dichas leyes no se tuvieron en cuenta el gran impacto que podrían tener las nuevas tecnologías en el futuro y las ventajas que esto conllevaba.

Por otro lado, fuera de nuestro país pero dentro de Europa se han tomado medidas para permitir un mayor uso de la datos en el ámbito de la sanidad, gracias a esto han “obligado” a los países que conforman la Unión Europea a tomar medidas y si no las han tomado todavía tomarlas en un futuro próximo, ya que desde Europa la idea está muy clara, y ésta es la de mejorar en línea generales la calidad del sistema sanitario y para ello se precisan de gran cantidad de datos.

1.1.3. ¿Qué beneficios representa?

Respecto a los beneficios [1] que presenta el *open data* en el sector de la sanidad son los siguientes:

- **Mejora en la toma de decisiones:** como es lógico, al tener una mayor cantidad de datos, ya que empresas privadas u organizaciones públicas liberan esta información al público, se pueden realizar estudios con una mayor profundidad haciendo uso de las tecnologías como *machine learning*, *deep learning*... De esta forma conseguimos predecir lo que puede suceder, lo cual nos da una gran ventaja.
- **Transparencia de gestión:** al pasar la información de un ámbito privado a uno público conseguimos que no haya secretismo, lo cual permite que haya una gestión más clara y transparente de lo que se hace con los datos que se recopilan de los pacientes.
- **Aumenta la eficiencia del sistema:** al poder manejar más información conseguimos por ejemplo que los tratamientos que son aplicados a los pacientes sean más efectivos, ya que tenemos un mayor conocimiento sobre cómo actúan dichos tratamientos y sobre qué tipo de personas tienen un mayor rendimiento.
- **Aumenta el número de pacientes informados:** siguiendo con la misma tónica que en los puntos anterior, al tener más datos puedes ser más preciso

a la hora de diagnosticar a un paciente, por lo que esto junto con el uso de *wearables* podemos mantener informada a una persona las 24 horas del día de una forma más precisa.

- **Reducción de costes:** al mejorar todos los puntos anteriores respecto a la sanidad más tradicional, conseguimos reducir los costes, siendo esto algo muy ventajoso.
- **Nuevos estudios:** una de las ventajas de compartir los datos a coste cero es que conseguimos que nuevos grupos de investigación puedan obtener datos para o bien, mejorar sus estudios en el campo, o descubrir por ejemplo nuevas enfermedades.

1.1.4. ¿Qué impacto ha tenido durante la pandemia?

Aunque el año pasado no fue el mejor para la gran mayoría de habitantes de este planeta, hay que reconocer que en los peores momentos es cuando salen las mejores soluciones, y este caso no iba a ser una excepción.

Si hay algo que debemos tener claro es que, la única forma de avanzar como sociedad y hacer frente a los problemas que han venido o están por llegar es que hay que invertir en la ciencia, es la única forma de solventar estos imprevistos.

El *open data* ha sido y es una gran inversión en el campo de la ciencia y sobre todo en esta pandemia. Tal y como han anunciado la empresa *HM Hospitales* [2] la única forma de acabar con el virus es haciendo uso de la tecnologías de la información y del *big data*.

Siguiendo con el ejemplo de *HM Hospitales* capturaron y a día de hoy siguen capturando información sobre los pacientes durante todo el proceso del *COVID*, permitiendo así tener una visión más detallada de cómo se comporta el *SARS-CoV-2*.

Una vez que habían recopilado esa información decidieron que lo mejor era publicarlo a la comunidad científica, gracias al compartir estos datos permiten a otras instituciones realizar estudios para conocer mejor cuáles son las causas y los efectos del *SARS-CoV-2*.

En resumidas palabras, el principal objetivo de *HM Hospitales* fue el de compartir los datos para así poder realizar modelos predictivos sobre cómo evolucionaba el virus y modelos sobre cómo afectaba a la población. Este no es más que un ejemplo, pero nos permite hacernos una idea de lo que se podría conseguir si tanto las empresas privadas como las públicas compartieran toda la información que tienen tanto en el ámbito de la sanidad como en otros ámbitos, ya que no hay mayor impacto que el compartir el conocimiento.

1.2. Pregunta 2

Describe los posibles perfiles involucrados en la transformación digital de HM Hospitales.

1.2.1. ¿Qué conocimientos previos y formación deben tener?

Lo primero de todo cabe mencionar que, es difícil saber a ciencia cierta todos los perfiles involucrados en la transformación digital de *HM Hospitales* ya que para ello tendríamos que estar dentro de la organización. Sin embargo, gracias a los recursos [3]–[5] que se nos ha proporcionado respecto a *HM Hospitales* podemos saber cuáles son sus objetivos (hacer uso del *cloud computing*, del internet de la cosas, de la inteligencia artificial y *machine learning*, de una infraestructura y servicios de almacenamientos de datos, de un óptimo mantenimiento y calidad de los mismos y por último de la seguridad), y con ello hacernos una idea de los perfiles necesarios.

Cuando hablamos de los roles involucrados en la transformación digital tenemos que hacer una división entre: perfiles de negocio y perfiles tecnológicos [6].

Los perfiles de negocio son aquellos en los que se encuentran los *stakeholders*, entendiendo por ellos como las personas u organizaciones que están relacionadas con las decisiones y actividades de una empresa [7]. Dentro de éstos podemos encontrarnos:

- **Patrocinador del proyecto:** es el encargado de apoyar el proyecto, ya sea tanto en dinero/soporte como en tiempo para que el proyecto se pueda realizar. El principal conocimiento que hay que tener en este perfil es el de saber invertir y buscar las colaboraciones necesarias para llevar a cabo el proyecto. Con este rol podemos identificar tanto a *HM Hospitales* como a *Telefónica*, ya que ambos tratan de sustentar el proyecto los primeros proporcionando el dinero y los segundos dando todo el apoyo necesario para poder cumplir con los objetivos.
- **Gerente del proyecto:** tiene como objetivo asegurarse de que el proyecto se ejecuta de forma correcta en la parte del negocio. Debe tener los conocimientos necesarios para poder monitorizar y evaluar si el proyecto va por el buen camino en cuanto a presupuesto, tiempo y objetivos cumplidos. Tanto en la documentación proporcionada como en la que he buscado por internet, no se sabe quién es el encargado en este perfil, quiero suponer que habrá un representante de *HM Hospitales* y *Telefónica*, pero lo que sí sabemos es que este rol es muy necesario.

- **Analistas/usuarios de negocio:** son los encargados de proporcionar los requisitos necesarios y hacer uso de diferentes herramientas para poder obtener conclusiones de los datos obtenidos gracias a la digitalización. Para poder desempeñar este rol tienen que saber el alcance del proyecto, es decir, los objetivos que se desean cumplir y hacer uso de herramientas para extraer información.

Por otro lado, los perfiles tecnológicos son los encargados de realizar la digitalización de la empresa a partir de los conocimientos técnicos que tienen sobre las tecnologías. Como bien he mencionado antes hay una serie de objetivos técnicos que se tienen que cumplir y para ellos hay determinados perfiles.

Respecto al responsable de la parte técnica encontramos el siguiente rol:

- **Chief information officer [8]:** es el director de las tecnologías y se encarga de que todos los procesos tecnológicos de la empresa se cumplan. Para ello debe tener grandes conocimientos a nivel técnico sobre las diferentes tecnologías, y por otro lado una gran visión general sobre cómo solventar diferentes problemas haciendo uso de las mismas. Este rol está representado por Alberto Estirado [4].

Por otro lado, para cumplir el objetivo del *cloud computing* necesitamos el siguiente rol:

- **Arquitectos de *cloud computing*:** este perfil es relativamente nuevo y su objetivo es el de diseñar y desplegar toda la arquitectura haciendo uso de tecnologías de *cloud computing*. Para ello deben tener conocimientos sobre las tecnologías que van a usar ya sean las de *Google*, *Amazon*...

Continuando con el cumplimiento de objetivos se encuentra el internet de las cosas, la inteligencia artificial y el *machine learning*. Para ello encontramos los siguientes perfiles:

- **Data Scientist:** es el encargado de identificar el problema que tiene que resolver y cómo lo tiene que hacer, haciendo uso de un conjunto de datos. Debe tener conocimientos de matemáticas (estadística), programación, *machine learning*, lógica de negocio y capacidad para comunicar la información.
- **Analista de BI:** responsable de las necesidades del negocio respecto al *Business Intelligence*. El BI está más focalizado en analizar los datos desde el presente al pasado para saber qué ha pasado según determinadas métricas, mientras que el *data scientist* se centra en los datos del pasado y presente

para predecir el futuro. Por lo tanto, el conocimiento entre ambos perfiles es similar.

Respecto a los objetivos de infraestructura y servicios de almacenamiento de datos, y su correspondiente mantenimiento y calidad encontramos los siguientes perfiles:

- **Arquitecto de datos:** es el encargado de diseñar la evolución del sistema para que siga funcionando según lo definido. Debe saber sobre las diferentes tecnologías sobre el *big data*.
- **Arquitecto de BI:** es un perfil similar al anterior rol, pero con el enfoque en los sistemas de *business intelligence*.
- **Analista de calidad de datos:** se encarga de garantizar la calidad de los datos que se usan y se almacenan. Debe tener un amplio conocimiento sobre los datos en general.

Finalmente, para cumplir con el objetivo de seguridad necesitamos el siguiente rol:

- **Gerente de seguridad:** es el responsable de que la seguridad en la plataforma junto con la arquitectura de la misma se cumpla según lo definido. Su mayor conocimiento está enfocado en la seguridad y en el uso de las tecnologías relacionadas con la misma.

1.2.2. ¿Qué habilidades deben poseer?

Cuando hablamos de una organización que está realizando un proceso para digitalizarse, debemos de tener en cuenta que se tiene que realizar un cambio cultural, es decir, pasamos de hacer uso de la información a ser analíticos con el uso de esa información.

Este cambio supone nuevas habilidades [6] que los perfiles comentados anteriormente deben adquirir:

- Tener un pensamiento crítico y analítico, es decir, que sean capaces de identificar un problema, seleccionar el mejor método para resolver dicho problema, y ser crítico con los resultados obtenidos.
- Basarse en los datos para poder tomar mejores decisiones.
- Ser capaces de usar las diferentes tecnologías, ya sean de *machine learning*, inteligencia artificial, bases de datos analíticas... Con el objetivo de poder resolver un problema a partir de los datos.

- Tienen que ser capaces de comunicar bien, ya que de nada sirve tener una cultura analítica dentro de la empresa si después no se aplica de forma correcta entre ellos.
- Tener la capacidad de adaptarse rápidamente al cambio, vivimos en un mercado muy competitivo y la única forma de mantenerse arriba es si se consigue hacer un buen uso de los datos durante un largo periodo de tiempo, ya que éstos pueden servir hoy pero no mañana.

1.2.3. ¿Qué buenas prácticas deben tener en cuenta?

Cuando hablamos de buenas prácticas que debe seguir una organización orientada a los datos, tenemos que hacer una clasificación entre los roles de negocio y tecnológicos [6].

Dentro de los roles de negocio las prácticas más características son las siguientes:

- Deben tener en cuenta que los perfiles tecnológicos necesitan bastante tiempo para conocer en profundidad el problema, y así poder desarrollar la mejor solución posible.
- Deben desarrollar un pensamiento crítico y analítico para poder discernir cuándo es mejor un modelo u otro.
- Hay que tener en cuenta todo el tiempo los objetivos, el presupuesto y tiempo que se tiene para cumplirlos.
- Es importante remarcar a lo largo de los empleados de la organización la importancia del proyecto, tanto su valor como beneficios.
- Finalmente, tiene que haber una gran comunicación dentro de la empresa por lo que, si hay algo que requiere más atención, o no se entiende es necesario dedicarle más tiempo y cambiar las metodologías para así comprenderlo mejor.

Por otro lado, los roles tecnológicos deben de seguir las siguientes prácticas:

- Deben desarrollar un conocimiento más enfocado en el negocio, para así poder comprender mejor las decisiones que se toman desde ese campo. Además, es necesario adquirir el vocabulario de negocio para comunicarse de una mejor forma con ellos.
- Al igual que sucedía en los roles de negocio, es necesario una buena comunicación entre los empleados del ámbito técnico, para así evitar

problemas en la toma de decisiones sobre qué modelo seguir. Por otro lado, si los roles de negocio no son capaces de entender las soluciones que se proponen, habrá que buscar otras formas para poder explicarlas.

- Para ver cómo evoluciona el proyecto es necesario estimar de una forma correcta el tiempo y coste de las tareas relacionadas con el mismo.
- Hay que tener un *modus operandi* basado en la organización, es decir, tener un ecosistema ordenado y estructurado.
- Finalmente, hay que ver el proyecto no solo como un proyecto, sino que también es una manera de mejorar la madurez digital de la organización.

1.3. Pregunta 3

1.3.1. Describe el ciclo de vida de los datos en *Covid data save lives*.

Cuando hablamos del ciclo de vida de los datos nos estamos refiriendo a cómo se gestionan éstos desde que son datos hasta que obtenemos información valiosa con ellos. Este ciclo de vida tiene seis etapas diferenciadas, las cuales vamos a ver a continuación:

En primer lugar, encontramos la fase de captura, en ella se busca el conseguir recolectar el mayor número de datos, ya sea con la creación de nuevos datos o con la extracción de datos ya existentes, en *Covid data save lives* los datos son generados y capturados por *HM Hospitales*. Se recopila información sobre el proceso de tratamiento del *COVID-19* [9], esto incluye datos del ingreso, altas sanitarias, medicación, tratamientos, constantes vitales...

Una vez que ya tenemos capturados los datos, el siguiente paso es almacenarlos, ya que de lo contrario estaríamos perdiendo información. En esta fase tenemos que almacenar los datos teniendo en cuenta el uso posterior que vamos a hacer con ellos. A la hora de guardarlos se puede realizar con ficheros simples o haciendo uso de base de datos. Sinceramente, en este caso creo que estos datos se guardan en una base de datos, ya sea relacional o no, pero posteriormente los comparten a la comunidad científica haciendo uso de tablas [9], por lo que entiendo que será un fichero *Excel* o similar.

Posteriormente, al ya tener los datos almacenados, podemos hacer uso de ellos para realizar tareas analíticas. Antes de nada, tenemos que preprocesar esa información, ya que de lo contrario no nos estaríamos garantizando la calidad de los datos. Dentro de

esta fase podemos realizar muchas tareas como fusión de datos, seleccionar solo aquellos campos que nos interese analizar, transformar datos de un tipo a otro, quitar valores nulos, crear nuevos campos a partir de los que ya tenemos... Como bien he mencionado, todas estas tareas se realizan en esta fase para asegurarnos de que cualquier persona pueda después hacer uso de estos datos sin tener en cuenta factores ya mencionados.

Siguiendo con el ciclo de vida, ya podríamos realizar un análisis de los datos, de tal forma que éstos dejen de ser datos para convertirse en información. Cabe destacar que esta fase la podemos enforar de maneras diferentes, ya que dependiendo de si somos *HM Hospitales* o un miembro de una organización científica buscamos objetivos distintos, el primero trata más de realizar un análisis desde el punto vista del *business intelligence* (leyendo diferentes artículos sobre ellos tratan de ver cómo estos datos capturados afectan a medidas internas de la organización, para así mejorar la calidad del servicio, es decir, se fijan en el pasado para mejorar el presente), mientras que la comunidad científica busca crear modelos predictivos de evolución y modelos epidemiológicos, en otras palabras, basarse en el pasado y presente para mejorar el futuro. El objetivo de crear estos modelos no es más que el entender cómo son los datos y qué les representa, de tal forma que nos permita responder preguntas haciendo uso de éstos.

Una vez realizado el análisis pasamos de los datos a la información, pero en la siguiente fase se busca convertir esa información en conocimiento, para ello se hace uso de representaciones visuales, estoy hablando de la fase de visualización. La manera más fácil de poder extraer conclusiones es haciendo uso de representaciones visuales (gráficos estáticos, dinámicos, interactivos...), gracias a ellos podemos tener una visión más general del problema. Dentro de este caso práctico, seguramente tanto *HM Hospitales* como cualquier científico haga uso de gráficos para poder entender y explicar mejor la realidad de la cuestión a responder.

Finalmente, una vez hemos adquirido el conocimiento suficiente sobre el problema y hemos respondido a las preguntas que nos habíamos planteado, llega la fase de publicar tanto el *dataset* como los resultados obtenidos. El principal objetivo de esta fase es compartir el conocimiento, para que así se pueda reutilizar en un futuro o que otros puedan usar esos mismos datos para resolver las mismas o diferentes preguntas. *HM Hospitales* en este caso lo hace muy bien, ya que los datos que adquiere sobre los pacientes durante el *COVID-19* los ha publicado a la comunidad científica con el fin de ponerle fin a este problema.

1.3.2. ¿Cómo es la pirámide D-I-K-W asociada al proyecto?

La pirámide DIKW (datos, información, conocimiento y sabiduría) o pirámide de conocimiento [10] la podemos definir como la forma en la que la información se define

a partir de los datos, el conocimiento a su vez a partir de la información y la sabiduría es el juicio óptimo sobre el conocimiento.

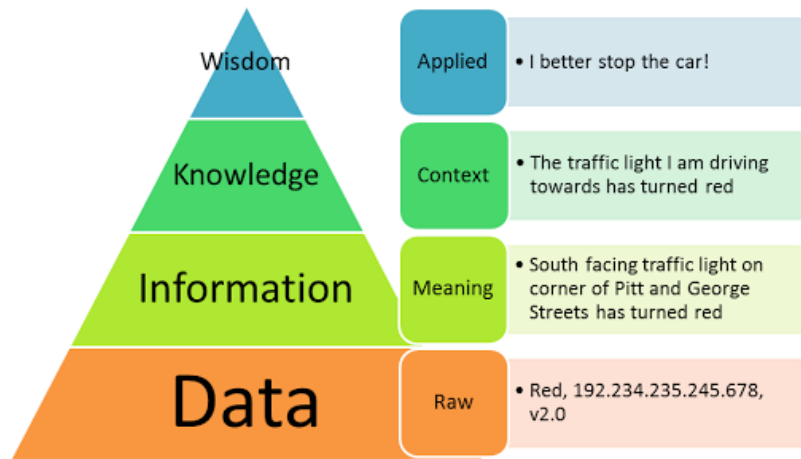


Ilustración 1 - Pirámide DIKW [11]

Como podemos observar en la anterior ilustración la pirámide se divide en cuatro bloques:

- **Datos:** podemos definir los datos como aquellos valores que sin un contexto carecen de sentido. En nuestro caso, los datos podrían ser por ejemplo: la altura, el peso, la frecuencia cardiaca... Son datos que tienen un valor pero que sin un sentido no aportan nada.
- **Información:** es el siguiente paso a los datos, es decir, dar a los datos un sentido. En el caso que nos corresponde sería el indicar que los datos que estamos recogiendo son relacionados con los pacientes de *COVID-19* en los hospitales de *HM Hospitales* a lo largo de España en el año 2020 y 2021, es decir, por ejemplo que el 80% de los hombres que han sufrido dicha enfermedad han estado en la unidad de cuidados intensivos, mientras que solamente el 40% de las mujeres han necesitado de dichos cuidados.
- **Conocimiento:** la diferencia entre el conocimiento y la información es que éste trata de dar un contexto sobre la información que tenemos, es decir, dada la experiencia permitir dar un contexto, en nuestro caso sería por ejemplo realizar un análisis de los datos que se han obtenido, de tal forma predecir si va a haber más o menos contagios en el próximo 3 meses, obteniendo x resultados podemos valorar si va a haber más contagios o menos.

- **Sabiduría:** la sabiduría va más allá que el conocimiento, su objetivo es saber aplicar correctamente el conocimiento que se ha adquirido para hacer un juicio de valor. En nuestro caso, una vez que sabemos si va a haber más o menos contagios en 3 meses, podemos usar nuestra sabiduría para saber si tenemos que tomar medidas más restrictivas o no.

1.4. Pregunta 4

Describe la transformación digital que ha vivido HM Hospitales.

1.4.1. ¿Podemos asegurar que es una organización orientada al dato?

Antes de asegurar si es una organización orientada al dato debemos saber qué es exactamente un organización orientada al dato [6], podríamos decir que una organización orientada al dato es aquella en la que las decisiones estratégicas, es decir, la toma de decisiones están basadas en datos.

Además siguiendo la teoría, se podrían identificar tres principios que toda organización orientada al dato debería de cumplir [6]:

- Tiene que ser capaz de identificar y gestionar diferentes fuentes de datos, respecto a *HM Hospitales* esto sí que se cumple, aseguran que registran datos [9] de tratamientos, ingresos, pasos por la unidad de cuidados intensivos, pruebas diagnósticas por imagen... Como vemos, sí que se capturan datos de distinta índole, por lo que este punto sí que lo cumple.
- Tiene que ser capaz de realizar un análisis en profundidad, haciendo uso de herramientas como *machine learning* para así mejorar la toma de decisiones. Respecto a este punto *HM Hospitales* también lo cumple, ya que como indica su página web oficial [9], recopilan todos estos datos para obtener modelos predictivos de evolución y modelos epidemiológicos con el fin de poner fin al *COVID-19*.
- Tiene que ser capaz de hacer un cambio cultural, para que la organización se base en el pensamiento crítico y analítico para así poder tomar mejores decisiones. Este punto también lo cumple tal y como se aprecia en las diferentes noticias proporcionadas [3], [4] se ha realizado una gran inversión para que se produzca ese cambio cultural, de hecho un ejemplo es lo que sucedió con el *dataset* del *COVID-19*, recopilando muchos datos y liberándolos con el fin de mejorar la situación actual en la que vivimos.

En resumen, apoyándonos en lo que vimos en la teoría para determinar qué es una organización orientada al dato y qué la caracteriza, podemos concluir que *HM Hospitales* sí que es una organización orientada al dato, eso sí tal y como veremos en el siguiente punto no creo que a día de hoy sea una organización orientado al dato al cien por cien, todavía les queda un largo camino.

1.4.2. ¿En qué fase del modelo de madurez se encuentra?

Cuando se entra a analizar la madurez de una organización orientada al dato a partir del modelo DELTTA, debemos saber que hay una serie de factores de éxito (seis factores) y que cada factor puede estar en una fase diferente (cinco fases), siguiendo con la matriz DELTTA proporcionada en los apuntes de teoría [6] y ajustándonos a la organización *HM Hospitales* podemos observar los siguientes factores junto con sus respectivas fases:

- **Datos:** considero que este factor presenta un nivel de fase 5, ya que tal y como se menciona en las noticias proporcionadas para esta práctica, hay una incesante búsqueda de nuevos datos y métricas con el fin de mejorar la calidad del servicio [5]. Aunque *HM Hospitales* ha desarrollado proyectos para liberar al público como el *dataset* del *COVID-19*, su mayor interés es adquirir más métricas para ofrecer a los pacientes el mejor servicio posible, queda muy remarcada esta idea.
- **Organización:** respecto a este punto, creo que esta organización se encuentra en una fase 3 muy avanzada o fase 4. Tal y como se puede leer [3] el proyecto empezó en 2019 y está planeado que dure hasta 2024, por lo que me cuesta crear que a nivel de organización haya ya una cultura analítica y crítica en el que todo esté centralizado o en red, he llegado a esa conclusión ya que tal y como menciona su CIO, Alberto Estirado [4] *HM Hospitales* está viviendo una revolución digital, sin embargo, se vieron afectados por la pandemia ya que no estaban preparados (tuvo que conectar a 6000 doctores vía VPN para realizar las consultas) si realmente estuvieran en una fase avanza no les habría pillado por sorpresa en algo a priori tan sencillo.
- **Liderazgo:** aquí consideró que están en fase 5 ya que hay un fuerte liderazgo para que se adquiriera una cultura analítica, creo que es así porque es el propio presidente de *HM Hospitales* el que habla y apoya esta idea de cultura orienta al dato [3], [5], y no solo apoya sino que también habla de las diferentes tecnologías como: *machine learning*, *big data*, entre otras, para alcanzar sus objetivos a partir de la mejora en la toma de decisiones.
- **Objetivos:** aunque es verdad que sus objetivos se quieren basar en la capacidad analítica de la organización, creo que todavía no han llegado a su máximo esplendor por lo que he comentado ya anteriormente, la transformación está a mitad de camino de su objetivo, por lo que es

imposible que a día de hoy puedan basarse al cien por cien en la capacidad analítica de la empresa. Es por todo ello que considero que están en fase 4.

- **Tecnología:** al igual que me sucedía con el anterior punto, están a mitad de camino de sus objetivos, y una evidencia clara es lo que comenté anteriormente (tuvieron que conectar a 6000 médicos vía VPN debido a la pandemia, si estuvieran preparados ya lo hubieran hecho así desde hace tiempo), por otro lado su CIO también menciona que tuvieron que desarrollar tecnologías de una forma rápida porque se acababan los plazos [4], por lo que considero que en este factor se encuentran en fase 3 un poco avanzada, es decir, han comenzado el despliegue de iniciativas a nivel corporativo y en determinadas áreas.
- **Personas:** sinceramente, creo que según las noticias leídas todavía no están rodeados de un alto equipo profesional, lo cual es algo normal, esta transformación todavía tiene que durar 3 años más. Sin embargo, considero que sí que tienen una buena base analítica y profesionales capacitados, ya que de lo contrario no hubieran hecho ni el proyecto del *dataset* del *COVID-19*, ni buscarían nuevos datos y métricas para mejorar la calidad del servicio. Es por ello por lo que considero que este factor está en fase 4, es decir, hay analistas altamente capacitados organizados centralmente o en red.

En resumen, creo que en determinados aspectos están muy avanzados, sobre todo a nivel de datos y liderazgo, pero considero que no podemos hablar de una organización que esté al cien por cien transformada digitalmente, les queda un largo recorrido a nivel tecnológico, sin embargo van por el buen camino ya que según lo leído, los altos mandos de la organización tienen claro que quieren una cultura enfocada al dato.

2. Bibliografía

- [1] «Open Data En El Ámbito Sanitario Y Su Compatibilidad Con La Privacidad Del Paciente | Andreu Martínez | Revue Internationale des Gouvernements Ouverts». <https://ojs.imodev.org/index.php/RIGO/article/view/200/330> (accedido mar. 06, 2021).
- [2] C. 31-12-2020 | 16:00 H. actualización 31-12-2020 | 16:00 H, «HM Hospitales aúna investigación y Big Data para cercar a la Covid-19», *La Razón*, dic. 31, 2020. <https://www.larazon.es/salud/20201231/yzm7vnrlzgrbpvc4f76fbak7q.html> (accedido mar. 06, 2021).
- [3] «Telefónica y HM Hospitales digitalizan el sector salud», *El Español*, abr. 30, 2019. https://www.elespanol.com/invertia/disruptores-innovadores/innovadores/20190430/telefonica-hm-hospitales-digitalizan-sector-salud/393962093_0.html (accedido mar. 07, 2021).
- [4] «En HM Hospitales hemos vivido una revolución digital a todos los niveles», *ComputerWorld*. <https://www.computerworld.es/entrevistas/en-hm-hospitales-hemos-vivido-una-revolucion-digital-a-todos-los-niveles> (accedido mar. 07, 2021).
- [5] «“No queremos curar, queremos llegar antes que la enfermedad” | Opinno». <https://opinno.com/es/insights/no-queremos-curar-queremos-llegar-antes-que-la-enfermedad> (accedido mar. 07, 2021).
- [6] «PID_00209854.pdf». Accedido: mar. 07, 2021. [En línea]. Disponible en: https://materials.campus.uoc.edu/daisy/Materials/PID_00209850/pdf/PID_00209854.pdf.
- [7] «Qué es un stakeholder: definición completa del concepto», *Rock Content - ES*, ago. 21, 2019. <https://rockcontent.com/es/blog/que-es-un-stakeholder/> (accedido mar. 07, 2021).
- [8] «El Papel del CIO en la Empresa Actual: Como influye | TicNegocios.es», *Tecnología para los negocios*. <https://ticnegocios.camaravalencia.com/servicios/tendencias/cual-es-el-papel-que-debe-tener-un-cio-en-la-empresa-actual/> (accedido mar. 07, 2021).
- [9] «Covid Data Save Lives». <https://www.hmhospitales.com/coronavirus/covid-data-save-lives> (accedido mar. 08, 2021).
- [10] N. Figuerola, «Gestión del Conocimiento (Knowledge Management) Pirámide D-I-K-W», p. 9.

- [11] «screen-shot-2013-04-06-at-13-42-12.png (507×356)». <https://legoviews.files.wordpress.com/2013/04/screen-shot-2013-04-06-at-13-42-12.png> (accedido mar. 08, 2021).