

Instituto
International de
Ciencia de datos

Metodología para una Transformación Organizacional hacia la Toma de Decisiones Basada en Datos

Víctor Barrera Núñez

Introducción

En este artículo se describe una metodología para orientar estratégicamente a una organización a partir de lo que sugieren los datos. En primera instancia, se procede con la definición de macrotendencias y de la analítica de datos, posteriormente se describe la metodología propuesta, y finalmente se dan algunas recomendaciones generales.

Las macrotendencias y su efecto

Las macrotendencias pueden ser definidas como fuerzas de carácter social, tecnológico, económico, ambiental y político que logran cambiar el modo en que las personas se transportan, trabajan, deciden, viven, envejecen, etcétera; es decir, modifican el modo de vida de la humanidad.¹ Estas son una realidad y pueden ser verificadas tanto con sentido común, como con datos históricos, así como con hallazgos derivados de la experiencia. El comportamiento de la humanidad siempre ha estado influenciado por el efecto de las macrotendencias.

Por ejemplo, se espera que la población urbana se incrementará en un 72% hacia el año 2050; que se requerirán un 40% más de recursos (agua, energía y alimentos) hacia 2030 a pesar de que la población mundial sólo crecería un 15%; el dominio económico está pasando de occidente a oriente, países como China, Rusia, Indonesia y Turquía son economías emergentes que están creciendo cada vez más rápido²; la cantidad de dispositivos conectados a la Internet se duplicará en el 2020, se pasará de 25,000 a 50,000 millones de unidades³.

Las anteriores estadísticas sugieren que aquellas organizaciones que se encuentren indiferentes, reticentes o incrédulas respecto a la existencia y efecto de las macrotendencias en su modelo de negocio o modelo operativo, están destinadas a su eventual salida del mercado en el mediano o largo plazo.

Por el contrario y en aras de evitar el fracaso, las organizaciones deben realizar continuamente revisión de las macrotendencias con la intención de reinventarse, y así reflexionar con antelación sobre las señales emitidas por las macrotendencias, y debatir cómo éstas estarían influyendo en el mediano y largo plazo en el modelo de negocio de la organización. Para lograr este propósito, las organizaciones requieren de la adopción de una estrategia que combina aspectos de carácter tecnológico y cultural.

¹ Five megatrends and possible implications, Directors Edition, April 2014

² National Intelligence Council: Global Trends 2030: Alternative Worlds

³ Cisco Internet Business Solutions Group, April, 2011

La analítica de datos

Se conoce como Analítica de Datos, o “*Data Analytics*” o “*Data & Analytics*” a todas aquellas tareas orientadas a la exploración de los datos, con la intención de encontrar patrones o conocimiento útil, que permita optimizar o rentabilizar un proceso de negocio. La Analítica de Datos intenta responder preguntas del tipo:

- ¿Qué sucederá con el volumen de ventas si cambian las condiciones del mercado?
- ¿Bajo las actuales condiciones del mercado cuál debería ser el precio óptimo del producto?

Nótese que las preguntas descritas arriba no pueden ser resueltas solamente mediante los sistemas tradicionales de bases de datos. En tal sentido, se debe recurrir a otras herramientas que permiten realizar modelos estadísticos avanzados para responder a tales preguntas. En general para convertir los datos en conocimiento, se realizan tres tipos de análisis conocidos como analítica descriptiva, predictiva y prescriptiva. La primera hace referencia a la exploración de las variables, la segunda a predecir el comportamiento de las variables y la tercera a dar recomendaciones o planes de acción de acuerdo a los hallazgos dados por las etapas de descripción y predicción.

La analítica de datos y las macrotendencias

El estudio de efecto de las macrotendencias mediante la analítica de datos es una herramienta fundamental para orientar estratégicamente el curso de cualquier organización. Por ejemplo, combinar datos provenientes de las redes sociales, con datos de campañas de mercadeo, apoyarán a la organización a valorar de manera más adecuada la efectividad de las campañas, y por ende, la organización inherentemente estaría comprendiendo a lo largo del tiempo, el valor que el consumidor da al producto o servicio, y consecuentemente identificar cuando existan cambios en el comportamiento del cliente.

Metodología

Algunas de las etapas de la metodología deben ser realizadas una única vez, otras continuamente y otras cronológicamente (Figura 1). La frecuencia de realización de cada una y su relación con los demás es también mencionada en los siguientes párrafos.

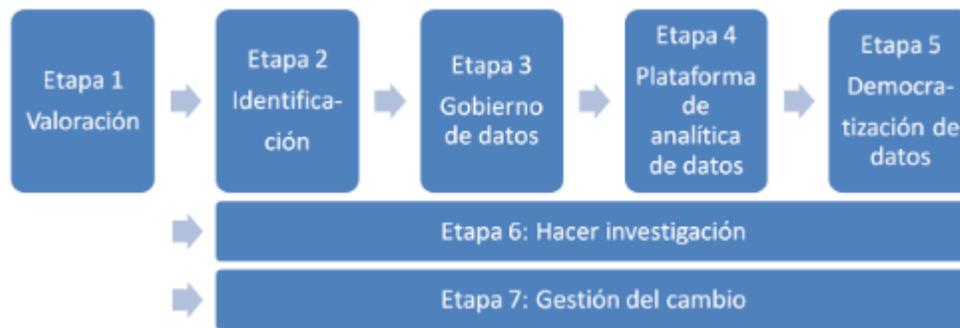


Figura 1. Metodología para una transformación organizacional hacia la toma de decisiones basada en datos

Etapa 1 - Valoración de la madurez de compañía respecto a la analítica de datos

El objetivo de esta etapa es identificar en qué nivel se encuentra la compañía, de acuerdo a lo establecido por un modelo de medición del grado de madurez de las prácticas de analítica de datos. Este modelo debe contemplar la valoración de aspectos de administración de datos, capacidad analítica, infraestructura, gobierno de datos y financiamiento relativos a la habilidad que tiene la organización para tomar decisiones desde una perspectiva basada en datos⁴.

La importancia de esta etapa radica en que establece el punto de partida de la organización. Por lo que al identificar su nivel de madurez, será posible generar una estrategia que lleve a la organización al siguiente nivel.

⁴ Analytics Maturity Model Guide, F. Halper & D. Stodder, TDWI, 2014-2015

Una valoración de esta naturaleza debe contemplar entrevistas y sesiones de trabajo con todas las áreas de la organización. La iniciativa debe estar apoyada por los tomadores de decisiones, de lo contrario, la intención de valoración del nivel de madurez tendría una alta probabilidad de fracasar.

Etapa 2 - Identificación de necesidades de analítica de datos

Una vez identificado el nivel de madurez de la organización, se recomienda identificar las necesidades de analítica de datos existentes a lo largo de las distintas áreas de la institución.

En tal sentido, el propósito es identificar las preguntas de negocio de corto, mediano y largo plazo que se pueden responder mediante la ayuda de modelos estadísticos o matemáticos. Estos modelos estarán acompañados posteriormente de los tableros (*dashboards*) que desplegarán las respuestas a las preguntas de negocio, o dicho en otras palabras, desplegarán la información extraída por los modelos a partir de los datos almacenados en los repositorios de la organización y/o repositorios externos. De manera similar, estos tableros deben ser actualizados periódicamente de acuerdo a la frecuencia en que la respuesta de negocio requiera ser respondida, esto es, en una base temporal diaria, semanal, mensual, trimestral, o según los requerimientos dilucidados durante las sesiones de trabajo de identificación de necesidades.

Las preguntas de negocio de corto plazo permitirán en etapas posteriores, mejorar la operación diaria y semanal debido a que éstas robustecen el proceso de toma de decisiones desde un enfoque basado meramente en la experiencia, por uno basado en la experiencia y en lo que los datos sugieren. Similarmente, las preguntas estratégicas de mediano y largo plazo, ayudarán a la conformación de líneas de investigación como elemento necesario para dar respuesta de manera recurrente dichas preguntas, de tal manera que se pueda controlar con mayor certeza y confianza el volante de la organización.

Los modelos estadísticos generados deben ser validados periódicamente para evitar desviación en la información que estos arrojen como resultado de cambios en la condiciones operativas del negocio, por ejemplo, cambio de regulaciones, recesiones, cambio brusco de divisas, ingreso de nuevos actores en el mercado, entre otros.

Etapa 3 - Implementación de una estrategia de gobierno de datos

Luego de la identificación de las necesidades de analítica de datos, es posible derivar a partir de éstas los datos requeridos para responder a las distintas preguntas de negocio. Algunos de estos datos son de carácter interno y otros, externo. Los primeros corresponden a datos alojados en los sistemas de información de la organización, mientras que los segundos, corresponden a aquellas fuentes de datos abiertas o de carácter privado, por ejemplo: INEGI, Profeco, NOAA, Nielsen, Kantar World Panel, etcétera. Las fuentes privadas deberían ser adquiridas por la organización para ser correlacionadas con sus datos internos.

La finalidad de esta etapa es identificar cuáles son los datos internos/externos que deberían integrarse en un único repositorio. Estos datos darán respuesta a las preguntas de negocio identificadas durante la etapa anterior. Adicionalmente, se deben

establecer las políticas de permisos de acceso a los datos y a los tableros que contienen las respuestas de negocio. Esta es la etapa inicial hacia una estrategia de democratización de datos, la cual será explicada más adelante en la etapa que lleva su nombre.

Es altamente recomendable realizar una estrategia de gobierno de datos antes de realizar la implementación de una plataforma de analítica de datos. De no realizarse, la nueva plataforma heredará prácticas y características no deseables, tales como: inconsistencia de datos a lo largo de la institución, vulnerabilidades de seguridad en datos e información, reprocesamiento ineficiente de datos (“*silos*”), ausencia de una definición de datos a nivel global de la organización, falta de confianza en los sistemas para apoyo a la toma de decisiones, entre otros.

El producto de esta etapa es un plan de acción para llevar a cabo la estrategia de gobierno de datos. Éste debería indicar cuáles son los actores que proveen de datos, qué modificaciones sufren los mismos durante su ciclo de vida, cómo éstos fluyen desde los actores que los proveen, hasta los distintos actores que los consumen, dónde residen. Las respuestas a las anteriores preguntas y otras, deberían estar contenidas en este plan de acción⁵.

Para alcanzar los objetivos de un proyecto de la naturaleza del descrito en esta etapa, es recomendable iniciar con la valoración de las actuales reglas, políticas y procedimientos de gobierno de datos con que cuenta la organización. Adicionalmente, es importante no perder de vista que transmitir a nivel organizacional los beneficios de una estrategia de gobierno de datos e involucrar a los principales interesados, es una actividad ineludible para garantizar la consecución satisfactoria de los objetivos del programa de gobierno de datos, lo cual se consigue con un componente relevante de gestión del cambio.

Etapa 4 - Implementación de una plataforma de analítica de datos

En este punto la organización conoce sus necesidades de analítica de datos y cómo se deben organizar y democratizar los datos que permitirán dar respuesta a las preguntas de negocio de corto, mediano y largo plazo.

La meta de esta etapa es implementar una plataforma tecnológica para el almacenamiento y procesamiento de datos. El objetivo de la plataforma es convertir los datos en conocimiento útil, es decir, convertirlos en respuesta a las preguntas de negocios identificadas en la etapa 2.

Hacia el año 2020 contar con plataforma de analítica de datos dejará de ser una ventaja competitiva para convertirse en un requerimiento. Consecuentemente, las organizaciones deben contemplar dentro de su plan estratégico, la inclusión de una plataforma analítica dentro de sus procesos productivos de datos.

Los siguientes son los beneficios de contar con una plataforma de analítica de datos y clasificados bajo distintos aspectos:

⁵ Make the information work to your advantage, Mayo, 2011

- Ingresos
 - Incremento del éxito de las campañas de mercadeo debido a la generación de estrategias centradas en el cliente.
 - Mejoramiento del entendimiento que se tiene del consumidor.
- Costos
 - Incremento de la productividad, facilidad a los empleados de la organización para acceder rápidamente a los datos.
 - Reducción de gastos operativos como consecuencia del reprocesamiento de datos (“*silos*”).
 - Reducción drástica de la cola de solicitud de datos al área de TI.
- Consistencia de datos
 - Existencia de una única fuente de la verdad con datos consistentes, validados, certificados y democratizados.
- Cultura organizacional
 - Desmitificación de los datos como arma política dentro de la empresa.
 - Incremento de la confianza de los empleados en los sistemas de información.

Las plataformas de analítica de datos también son conocidas como “*Data Lake*”, por hacer referencia al gran volumen de datos que son capaces de almacenar y procesar. En ocasiones también son referidos como “Refinería de datos”, en donde los datos en bruto corresponden a la materia prima, mientras que las respuestas a las preguntas de negocio como el producto. Las plataformas de analítica de datos suelen ser más económicas que un “*datawarehouse*”, en una relación de 1 a 10 aproximadamente.

La plataforma requiere ser operada por perfiles profesionales especializados en la administración y operación de infraestructura *hardware* y *software* de este tipo, por lo que el autor recomienda a las organizaciones adquirirla bajo un esquema de servicio, de esta manera, la organización podrá concentrarse en su negocio y ahorrar costos por concepto de devaluación de activos físicos y carga fiscal por consecuencia de contratación de personal. Adicionalmente, existe otro equipo de perfil profesional conocidos como Científicos de Datos, los cuales estarían encargados de la parte funcional de la plataforma, es decir, del entendimiento de las necesidades de analítica de datos, de la generación de los modelos estadísticos, y del diseño y mantenimiento de los correspondientes tableros de control. En este último caso, la organización debe también tomar la decisión de contratar o no, servicios externos de científicos de datos, o generarlos internamente dentro de la organización.

Etapa 5 - Implementación de una estrategia de democratización de datos

Cuando los datos ya han sido integrados en la plataforma analítica, deben ser puestos a disposición de los empleados de la organización siguiendo las políticas de gobierno de datos implementadas en el plan de acción de la etapa 3.

En resumen, la democratización de datos puede ser definida como la acción de poner a disposición de los empleados de la organización, *la mayor cantidad de información*

*dentro de lo legalmente posible*⁶. Esto es, brindar de herramientas de información, más allá de los cargos que tradicionalmente contaban con esta facilidad.

Para el despliegue de la estrategia de democratización de datos, es recomendable la implementación de soluciones de reporte con características “*self-service*” y “*responsive-design*”. La primera característica consiste en que los involucrados, además de poder acceder a los tableros de control desarrollados por los Científicos de Datos, también tengan la posibilidad de crear de manera independiente, tableros o correlacionar datos que den respuesta a nuevas preguntas de negocio que puedan surgir durante la operación diaria⁷. De otro lado, la segunda característica permitirá a los involucrados acceder a los datos, independientemente si lo hacen a través de un dispositivo móvil, o uno de escritorio, o distintos tipos de navegadores de Internet.

La democratización de datos traerá como beneficios principales: la reducción de solicitud de datos al área de TI y por tanto, un incremento en la eficiencia del personal al momento de tomar decisiones basándose en datos históricos. Una vez establecidas las etapas anteriores, las colas de espera para la entrega de datos en el área de TI serán cuestión del pasado, y los empleados podrán analizar los datos con autonomía y fácil acceso.

Luego de la democratización de los datos, el área de TI podría dedicar sus esfuerzos a generar un mayor valor estratégico a la organización, dando respuesta a las preguntas de negocio de mediano y largo plazo identificadas en la etapa 2.

También sería conveniente que la estrategia de democratización de datos, incluyese el envío de alertas y excepciones a los empleados correspondientes, cuando ocurra un cambio de interés en alguna variable operativa. Lo conveniente para iniciar con una estrategia de este tipo es iniciar con programas pilotos factibles y de alto impacto, de tal manera que al finalizarlos, afloren claramente las bondades de democratizar los datos.

Etapas 6 - Hacer investigación. Esta etapa debe realizarse continuamente una vez se hayan identificado las preguntas de mediano y largo plazo, es decir, desde el inicio del esfuerzo de alinear a la organización en una cultura de toma de decisiones basada en datos. Periódicamente a medida que se vayan respondiendo preguntas, se deben ir incluyendo otras, o profundizando las mismas, de acuerdo a las tendencias y macrotendencias del mercado.

La finalidad de esta etapa es proveer de información estratégica a los tomadores de decisión, de tal manera que éstos puedan orientar con una mayor fiabilidad el rumbo de la organización.

La identificación de riesgos que se presentarán en el futuro de mediano y largo plazo, así como las sugerencias para mitigarlos, son el producto esperado de esta etapa.

⁶ Data Driven: Creating a Data Culture, D. Patil & H. Mason, O'Reilly, 2015

⁷ The democratization of data, Infor Analytics, 2013

Es preferible que las actividades de investigación sean realizadas por un grupo interno o externo dedicado explícitamente a hacer estudios prospectivos, acerca de cómo las macrotendencias pueden incidir en el modelo de negocio, nuevos actores en el mercado, nuevas regulaciones, entre otros.

Etapas 7 - Realización de Gestión del cambio

Las actividades contenidas en las etapas anteriores suelen generar sentimientos de reticencia, miedo o amenaza por parte de las personas que se encuentran involucradas, tanto directa o indirectamente por los cambios derivados de un proyecto de transformación cultural, como el descrito a lo largo de este artículo.

Los sentimientos descritos arriba, pueden con una alta probabilidad conllevar a que un proyecto de este tipo fracase. Por tal motivo es importante, ante una transformación cultural de esta naturaleza, realizar actividades que permitan concientizar a las personas mostrando las ventajas de la transformación, tanto desde un punto de vista particular personal, como colectivo organizacional.

Una transformación hacia una cultura de toma de decisiones basada en datos, si bien contiene un componente tecnológico y de procesos, las actividades encaminadas a la gestión cambio, juegan uno de los papeles más importantes respecto al resto de actividades. En tal sentido, esta actividad debe ser llevada a cabo con la adecuada sensibilidad política, sentido común y tacto social.

Recomendaciones generales

Se ha descrito una metodología para la consecución de una transformación organizacional hacia la toma de decisiones basada en datos. La metodología está conformada por 7 etapas y contempla tanto aspectos tecnológicos, de procesos y culturales para convertir a una organización, en una institución que soporta sus decisiones en la experiencia y lo que los datos sugieren.

Una transformación de este talante no se da de la noche a la mañana, requiere del apoyo de expertos para la realización de un mapa de ruta, alcanzable, medible y ajustado a las necesidades, giro y entorno (social, económico, financiero y tecnológico) de la organización. Los aspectos culturales relativos a la gestión del cambio son de especial interés para lograr alcanzar los objetivos de una transformación de esta naturaleza.

La cálculo del retorno de la inversión de un proyecto de este tipo no es de fácil estimación, por lo que la consecución de recursos puede ser complicada de justificar ante el área financiera de la organización. Para lograr sortear esta situación, se sugiere iniciar con un proyecto piloto de alta factibilidad y alto impacto, del cual sus resultados permitan evidenciar las bondades de una transformación hacia una cultura de datos, y estimar posteriormente, el retorno de la inversión. Seguidamente, sería posible la consecución de recursos económicos para un despliegue de mayor alcance sobre la estructura organizacional.



Acerca del autor

Víctor Barrera Núñez es un experimentado científico de datos con maestría y doctorado, con especialización en proceso de toma de decisiones basados en datos.

Aplica métodos de inteligencia artificial, estadística y minería de datos en grandes bases de datos y datawarehouse a las áreas de administración de ingresos, a la mejora de eficiencia o productividad y a la extracción de ideas.

Actualmente trabaja en AXA México como Director de Datos, donde le corresponde ejecutar la implementación de principios y soluciones de análisis de datos, a través de teoría estadística multivariable y técnicas de inteligencia artificial para el crecimiento de ingresos. Su objetivo último en AXA México es convertir la empresa en una organización basada en datos, siguiendo tres principios: hacer de una función de datos una función global, generar una sola fuente de verdad y optimizar los procesos de análisis de datos.



Acerca del Instituto

El Instituto Internacional de la Ciencia de Datos busca resolver preguntas importantes de la sociedad a partir de los datos que genera la misma. Construye sinergias entre la creciente comunidad de entusiastas y científicos de datos y fomenta la correcta ejecución de la apertura de datos y conocimiento.



<https://www.linkedin.com/company/instituto-de-la-ciencia-de-datos>



@the_i2ds

© 2017 Instituto Internacional de Ciencia de datos. Las opiniones expresadas en este documento no reflejan necesariamente la posición oficial del i2ds. La información incluida en la publicación se obtiene de fuentes de información de terceros y de fuentes públicas y proporciona una guía general acerca de la industria y no debe utilizarse como consejo del i2ds, o como sustituto de servicios profesionales.

El i2ds no se será responsable de ninguna consecuencia, daño o perjuicio que pudieran derivarse de dicho uso de la información de este documento.