

**ACTIVIDAD DE APRENDIZAJE**

.  
GESTIÓN DE LA CALIDAD DE LA INFORMACIÓN

Innoo

**ACTIVIDAD DE APRENDIZAJE No. 4:**

**Calidad de los datos**

|  |
| --- |
| Programa académico |
| Ingeniería de sistemas |

|  |
| --- |
| Asignatura |
| Gestión de la calidad de la información |

|  |  |  |
| --- | --- | --- |
| Estrategia didáctica | Duración estimada | |
| En aula o en plataforma | Trabajo autónomo |
| Simulación | 5 horas | 20 horas |

**CONTEXTUALIZACIÓN**

En un mundo centrado en la generación y procesamiento de datos es de gran importancia desarrollar un gran diferencial frente a los demás competidores en el sector empresarial. El conocimiento más una acción, es lo que llamo una ventaja competitiva; y explotar los datos para generar el conocimiento es lo que se debe hacer para apoyar a las empresas en tres perspectivas: la Ciberseguridad, la gestión y Monitorización de la infraestructura tecnológica y en la Inteligencia de negocio.

Sin embargo, las empresas apenas empiezan a comprender el valor de los DATOS. La transformación de varios sectores como el de transporte y el turístico, en empresas como Uber o AirBnB, hace que las compañías comiencen a entender cuál es el valor en el manejo de información sensible y confidencial a través de plataformas digitales. Aunque queda mucho camino por recorrer, ya se habla de que el valor de una empresa se mide, también, por los datos que maneja.

Lo que sí es cierto es que el 79% de las empresas afirman que el análisis de datos mejora la toma de decisiones, según información arrojada por el Centro de Seguridad y Vigilancia Digital (CSVD) de la compañía A3Sec. Asimismo las innovaciones de Big Data aplicadas en los sectores financieros y salud, como, por ejemplo, el desarrollo de herramientas de automatización de tareas que ayudan a reaccionar de forma rápida, eficaz y oportuna han permitido identificar oportunidades y riesgos basados en los datos de la compañía.

**Al finalizar esta actividad estará en capacidad de:**

Resultado cuatro: Explicar los referentes teóricos de los datos, sus dimensiones, su correlación y la calidad de los mismos teniendo en cuenta su utilización en las organizaciones y su impacto social y global.

**ACTIVIDAD**

Usted ha sido contratado como director de calidad de la información en una compañía internacional que ha venido teniendo problemas con sus clientes en los últimos meses.

Es por esto que le han solicitado generar un **plan de acción para implementar un proyecto de calidad de datos.** La solución a estos problemas pasa por implementar procesos de calidad adaptados a las necesidades de la empresa, basados en el control del ciclo de vida de la calidad de los datos mediante un perfilado de la estructura y el contenido y una limpieza. Por lo general, los pasos se dividen en distintas acciones, que se llevan a cabo siguiendo el orden lógico: descubrir, analizar/definir, desarrollar, revisar/analizar y monitorizar. Idealmente, además, la solución ha de ser global, no departamental, si no desde un principio, al menos, sí a largo plazo.

Además de implementar un proyecto de calidad de datos, transformar pérdidas en beneficios requiere contar con recursos humanos idóneos, tecnologías de vanguardia.

En esta **tercera etapa** se solicita consultar y articular un documento que resuma:

1. Conceptos de datos, su estructura, sus dimensiones, dentro de la organización.

**Se definen los datos como toda información o documentación que es manejada en este caso en el interior de una empresa, es decir todo tipo de documento que es usado dentro de una organización con el fin de apoyar algún proceso, en cuanto a su estructura, es la forma en la que los datos (o información),se encuentra organizada y de que forma se puede gestionar, es decir la estructura es la que nos permite definir de qué forma esta información va a ser almacenada, consultada y modificada.**

**En cuanto a las dimensiones de los datos, dentro de un sistema de información, estos datos pueden ser presentados en forma de tablas, los cuales son vectores de múltiples dimensiones, es decir es una colección consecutiva de los datos. Esto sobre la marcha, puede ser representado por medio de una base de datos en la que se almacene toda la información que es gestionada por la empresa, y a su vez esta información ya ubicada en tablas, debe definirse de que forma se administrara, por medio de la estructura de los datos.**

1. Métodos de estimación de la dimensionalidad de los datos al interior de la empresa

**Estos métodos de estimación de la dimensionalidad de los datos son importantes que se lleven a cabo, debido a que de esta forma, estamos haciendo un barrido sobre toda la información que use la empresa, con el fin de determinar qué información no es relevante para sus procesos, muchas veces mas es menos y en este caso para la empresa también aplica, es decir no por tener una gran cantidad de datos, significa que se tiene un buen sistema de información, el fin de estos métodos es lograr una reducción de la dimensionalidad de los datos, para tener un sistema de información más ágil y mejor consolidado, menos complejo y que facilite el rendimiento del mismo.**

**Dentro de estos métodos, los mas usados son la selección de variables, que vendrían siendo el encapsulamiento de los datos, en este caso se procede a definir que variables son las mas importantes dentro del sistema de información, por medio de la correlación y asociación de los datos, también a su vez se encuentra lo que es la identificación de que variables son redundantes, es decir identificar que variables o datos no son relevantes, ya que puede haber otro dato que contenga esta información.**

**Todo esto para la empresa se ve reflejado como pulir toda la información que actualmente se maneja para a su vez entregar un mejor rendimiento al sistema de información, y que de esta forma se facilite la administración del mismo.**

1. Importancia los procesos de reducción de la dimensión de los datos dentro de la organización.

**Como se explico en el punto anterior, esta reducción de la dimensión de los datos, o la información, es bastante importante, debido a que este proceso permite que el sistema de información sea mas efectivo y preciso en la administración de los datos, se puede poner como ejemplo, un celular que tiene 85 aplicaciones instaladas, las cuales ocupan mas del 80% del almacenamiento del celular, pero sucede que, de esas 85 aplicaciones, solo usamos 20. De esta forma nos damos cuenta que tenemos 65 aplicaciones que nos están generando afectación en el rendimiento del celular y que no son fundamentales de acuerdo a nuestro requerimiento. Si procedemos a eliminar estas 65 aplicaciones sin usar, tendríamos como resultado, una ocupación del 30% aproximadamente, que a su vez esto mejorara el rendimiento de nuestro celular. Así mismo sucede en una empresa, y los procesos de reducción pueden significar un gran aspecto diferencial, frente a otros sistemas de información.**

1. Algoritmos locales de reducción de dimensiones de los datos dentro de la organización.

**La tarea de estos algoritmos de reducción, son básicamente automatizar el reproceso de datos, con el fin de impactar positivamente la estructura de los datos y de esta forma lograr una reducción en la variables que son manejadas en el sistema de información, es decir, gracias a estos algoritmos, se nos permite optimizar la cantidad de variables o datos que son usados en el sistema de información, es decir, realizaran una búsqueda para determinar que variables son redundantes o no son necesarias para una correcta gestión de los datos. A su vez este proceso, permitirá una mejor respuesta y rendimiento. Lo cual se traduce como una mejora implementada al sistema de información, entonces, estaremos entregando un valor agregado, o un diferenciador frente a otros sistemas de información.**

1. Interpreta los métodos globales de reducción de la dimensionalidad de los datos dentro de los sistemas de información y su impacto en la empresa.

En la **COMPAÑÍA MULTINACIONAL (de Confecciones)** se realizó un mapeo del conjunto de los datos a subespacios provenientes de los datos originales, de menor dimensión, a través del sistema ERP para permitir llevar a cabo el análisis descriptivo de datos multivariados mediante el método de reducción de variables que dentro del Plan de Acción tiene los siguientes objetivos:

* Generar nuevas variables que expresen la información contenida en un conjunto de variables.
* Reducir la dimensión del espacio donde están inscritos los datos.
* Eliminar las variables que aporten poco al estudio del problema.
* Facilitar la interpretación de la información contenida en los datos.

1. Importancia de la correlación de las dimensiones dentro del análisis de los datos de la organización.

La ordenación de la información que incorpora una componente principal permite la normalización de los datos y por ende la correlación de las dimensiones dentro del análisis de los datos de la compañía.

Como Directores de Calidad de **COMPAÑÍA MULTINACIONAL (de confecciones)** obtuvimos información de variables correlacionadas entre sí en mayor y menor grado. Por esta razón se implementó el método de reducción de variables para pasar a un nuevo conjunto de variables (no correlacionadas) llamadas componentes principales, es decir variables incorrelacionadas entre sí, para poder ordenarlas por la información que llevan incorporadas.

Como medida de la cantidad de información incorporada en una componente se utiliza su varianza. Es decir, cuanto mayor sea su varianza mayor es la información que lleva incorporada en dicha componente.

1. Importancia de la calidad de los datos dentro de la organización.
2. Importancia del procesamiento de los datos como generador de información dentro de la organización.

**Entregables:**

1. Informe en Word con presentación en video

**REFERENTES CONCEPTUALES**

***Unidad 1. Conceptos de datos, su estructura, sus dimensiones, dentro de la organización.***

***Unidad 2. Métodos de estimación de la dimensionalidad de los datos al interior de la empresa***

***Unidad 3. Importancia los procesos de reducción de la dimensión de los datos dentro de la organización.***

***Unidad 4. Algoritmos locales de reducción de dimensiones de los datos dentro de la organización.***

***Unidad 5. Interpreta los métodos globales de reducción de la dimensionalidad de los datos dentro de los sistemas de información y su impacto en la empresa.***

***Unidad 6. Importancia de la correlación de las dimensiones dentro del análisis de los datos de la organización.***

***Unidad 7. Importancia de la calidad de los datos dentro de la organización***

***Unidad 8. Importancia del procesamiento de los datos como generador de información dentro de la organización.***

**ROLES Y FUNCIONES**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Roles asociados a la actividad | Funciones | Tareas asociadas a las funciones | Tiempos estimados para el desarrollo de la tarea | |
| Presencial / plataforma | Autónomo |
| Estudiante | Generar el plan de acción para implementar un proyecto de calidad de datos | Consulta |  |  |
| Docente | Acompañar el proceso de construcción de los documentos | Acompañamiento |  |  |

**ESPACIOS DE SOCIALIZACIÓN**

Encuentro sincrónico

**EVALUACIÓN**

|  |  |
| --- | --- |
| Criterios de evaluación | Evidencias asociadas (desempeño, producto y conocimiento) |
| CE 16. Maneja los conceptos de datos, su estructura, sus dimensiones, dentro de la organización.  CE 17. Explica los métodos de estimación de la dimensionalidad de los datos al interior de la empresa  o de la comunidad.  CE 18. Demuestra la importancia los procesos de reducción de la dimensión de los datos dentro de la  organización.  CE 19. Ilustra los algoritmos locales de reducción de dimensiones de los datos dentro de la  organización.  CE 20. Interpreta los métodos globales de reducción de la dimensionalidad de los datos dentro de los  sistemas de información y su impacto en las empresas.  CE 21. Demuestra la importancia, de la correlación de las  dimensiones dentro del análisis de los datos de la organización.  CE 22. Ilustra la importancia de la calidad de los datos dentro de la organización, y en la comunidad.  CE 23. Demuestra la importancia de los procesamiento de los datos como generador de información dentro de la organización. | Informe en Word con presentación en video |

|  |
| --- |
| Instrumentos de evaluación |
|  |

**RECURSOS NECESARIOS PARA LA ACTIVIDAD**

|  |  |  |
| --- | --- | --- |
| Recursos | | |
| Técnicos, tecnológicos y materiales | **Computador** | |
| Ambientes de aprendizaje | **Plataforma** | |
| Recursos bibliográficos | **Obligatorios** | Los presentes en plataforma |
| **Complementarios** | • Opinión - Ángel García-Muñoz Rodrigo, Executive MBA EADA. Profesor colaborador de la Asociación de la Industria Navarra (AIN). Director Gerente de Aticaredex.  • Presentación – El proceso de Comunicación. Guía de buenas prácticas de comunicación interna  • El impacto de las tic en las familias. Universidad de Alcalá |