

TELESCOPI

---

**Red de Observatorio de Buenas Prácticas de  
Dirección Estratégica Universitaria en América  
Latina y Europa**

**Buena Práctica No. 16**

**“Analítica para todos”: Configuración de entornos  
de trabajo para entender la Universidad Pública  
desde los datos**

**Convocatoria de Buenas Prácticas 2019**

**“Analítica para todos”: Configuración de entornos de trabajo para entender la Universidad Pública desde los datos**

**Palabras claves:** Big Data, Analítica, Datos.

<b>Institución responsable de la práctica</b>				
<b>Nombre de la institución</b>	<b>Ciudad</b>	<b>País</b>	<b>Teléfono / Fax</b>	<b>Sitio web de la institución/Facultad/Centro</b>
Universidad de Antioquia	Medellín	Colombia	219 5050	<a href="http://www.udea.edu.co/">http://www.udea.edu.co/</a>

<b>Datos de la persona responsable de la práctica</b>					
<b>Nombre y apellidos</b>	<b>Cargo</b>	<b>Unidad/Facultad/Escuela</b>	<b>Correo electrónico</b>	<b>Teléfono /Fax</b>	<b>Sitio Web de la Práctica</b>
Sol Mery Álvarez Argaez	Analista	Dirección de Planeación y Desarrollo Institucional- División de Planes y Proyectos	<a href="mailto:sol.alvarez@udea.edu.co">sol.alvarez@udea.edu.co</a>	219 5084	<a href="http://www.udea.edu.co/wps/portal/udea/web/inicio/institucional/data-udea">http://www.udea.edu.co/wps/portal/udea/web/inicio/institucional/data-udea</a>

<b>Criterios de Excelencia en los cuales se enmarca la buena práctica (Marcar dos criterios como máximo)</b>						
<b>Liderazgo</b>	<b>Estrategia</b>	<b>Personas</b>	<b>Alianzas y Recursos</b>	<b>Procesos, productos y servicios</b>	<b>Clientes</b>	<b>Responsabilidad Social</b>
				X		

## **1. Resumen ejecutivo de la Práctica:**

Existen múltiples variables que intervienen en la vida universitaria. Todas ellas, en constante evolución y provenientes de diversas fuentes (externas e internas), aumentan la complejidad tanto en el análisis como en la entrega continua de información enriquecida para el direccionamiento estratégico y la gestión de las universidades. La Universidad de Antioquia, compuesta por más de 50.000 integrantes vinculados en ámbitos de índole ciudadano, regional e incluso virtual, ha gestado diversas iniciativas respecto a la mejora de los procesos de gestión de información para la toma de decisiones, con el objetivo generar soluciones que permitan analizar y comprender las dinámicas institucionales y del entorno (p.e. permanencia estudiantil, vida laboral de los egresados, gestión de procesos académicos y administrativos, entre otros).

Durante los últimos 10 años dichas iniciativas se han concentrado, principalmente, en dos frentes: la configuración de observatorios institucionales y la implementación de plataformas tecnológicas de inteligencia institucional. Ambos acercamientos han resultado limitados para dar respuesta a las necesidades de integración de datos, análisis, procesamiento, visualización y agilidad en la divulgación y acceso a la información que la comunidad universitaria, en especial los tomadores de decisiones, han requerido.

Debido a esto, la Universidad habilita una plataforma de big data, configurada en su gran mayoría con software libre, usando los pilares clave para la próxima generación de plataformas de datos: Cloudera, JupyterHub, Hadoop, Hive, HBase, Impala, Apache Drill, Apache Presto, Python, R, Scala, Spark y Tableau. El primer producto publicado, gracias a la habilitación de esta plataforma, es denominado "Data UdeA", disponible en <http://data.udea.edu.co>. "Data UdeA" ha llevado a la Universidad a entenderse desde los datos. En este espacio del portal universitario se dispone información estadística histórica, con datos precisos, coherentes y de consenso institucional, visualizados como una infografía interactiva, que permite a todos los usuarios construir sus propios indicadores.

Mediante la práctica de disponer datos que estén al alcance de toda la comunidad universitaria, se ha generado una cultura del uso de datos que ha permeado las presentaciones oficiales de la Universidad, el reporte a entes externos, los procesos de autoevaluación, las presentaciones fundamentadas de las propuestas de los aspirantes a decanos ante el Consejo Superior y las estrategias de bienestar orientadas a la comunidad universitaria. Por último, gracias a esta práctica, se consolidó el cierre de la primera fase de la implementación de la Unidad de Analítica y Estudios Universitarios - Observatorio de la Universidad de Antioquia (UAO), cuya misión será la de "Proveer información y conocimientos de valor para la orientación y gestión de la Universidad de Antioquia, mediante la realización de investigaciones, la analítica de datos en infraestructura para big data y el desarrollo de estrategias culturales frente al uso y comprensión de los datos y la información; dirigidos tanto para los órganos de gobierno, instancias y dependencias universitarias como para instituciones y organizaciones del entorno".

## **2. Planificación de la Práctica**

Como punto de partida para la práctica, se destaca el proyecto de implementación de plataformas tecnológicas de inteligencia de negocios, liderado por la Oficina de Planeación en el Departamento de Organización y Sistemas, entre los años 2007 y 2008. Este proyecto estuvo orientado al montaje de bodegas de datos institucionales (datawarehouse) usando la tecnología disponible en ese momento (2007), que enlazaran los diferentes sistemas de información de la Universidad, procesaran los datos y entregarán información útil para la toma de decisiones, en los ámbitos de docencia, investigación, extensión y gestión administrativa. Sin embargo, esta iniciativa no tuvo acogida por parte de los usuarios finales y por lo tanto tampoco fue sostenible para la institución.

A pesar de esto, la necesidad de datos e información consolidados y válidos se mantuvo. La Universidad necesitaba contar con un sistema de analítica y observación institucional integrado, que permitiera enriquecer la información que se recopila y almacena diariamente, producto de los procesos académicos y administrativos, y así lograr que ese bien intangible le sirva de soporte para la toma de decisiones respecto a sus procesos misionales. Con base en las lecciones aprendidas de esa iniciativa (gestada en 2007), la Dirección de Planeación y Desarrollo Institucional en el año 2016 lidera la creación de un sistema de analítica y observación institucional, el cual permita el análisis de datos y la puesta a disposición de la comunidad universitaria de información enriquecida que facilite la toma de decisiones asociada a su gestión. Los objetivos definidos para el desarrollo de esta iniciativa fueron:

- Instalar una plataforma de big data para la Universidad de Antioquia
- Instalar capacidades para la observación y el análisis permanente de la gestión universitaria y de su entorno
- Difundir para el público general las principales cifras institucionales que dan cuenta de la gestión universitaria.

El punto de partida para el cumplimiento de estos fue el dimensionamiento y adquisición de infraestructura de hardware y herramientas de software, que permitiera el registro de información detallada desde múltiples fuentes de información, con múltiples opciones de visualización y acceso por parte de los usuarios finales. Como resultado de esta etapa, se adquirió tecnología con las siguientes características:

- Infraestructura escalable (horizontal y verticalmente) para analítica de datos, compuesta por cinco (5) servidores con 152 cores para procesamiento, 2,6 TB de memoria y 38 TB de almacenamiento.
- Software libre para analítica de datos, como Jupyterhub, Python, R, Scala, Apache Zeppelin, Apache Drill, Apache Presto, Spark, Hadoop, Apache HBase, Hive y Kafka.
- Herramientas para visualización, como Tableau y Supersets. Desde el punto de vista del recurso humano, participaron en el proyecto cuatro ingenieros con dedicación a tiempo completa, durante un periodo de tres años.

El cumplimiento de esos tres objetivos representó la generación de capacidades propias para el desarrollo de los estudios de comportamiento y tendencias en asuntos misionales y de soporte a la Universidad, entregados como un primer producto: “Data UdeA” (<http://data.udea.edu.co>). La inversión total en esta primera etapa fue de \$1.300.000.000.

### **3. Desarrollo y ejecución de la Práctica**

El logro de los objetivos planteados se dio como resultado de la ejecución de tres grandes etapas planeadas para un tiempo de tres años: 1ra: Referenciación y adquisición de la plataforma Se planificó un proceso de referenciación para identificar: qué plataformas eran usadas por empresas e instituciones educativas de la ciudad para atender necesidades similares a las identificadas y cuáles plataformas para big data eran las más usadas en el mundo. De los resultados obtenidos, se eligieron las herramientas que mejor se adaptaran al ecosistema tecnológico de la Universidad, tuvieran licencia de software libre y pudieran ser escalables y actualizadas en el tiempo.

Una vez adquirida y provisionada la infraestructura física de servidores, se iniciaron las pruebas de configuración y las pruebas de extracción, transformación y carga (ETL) de las bases de datos institucionales a la nueva plataforma. Es importante mencionar que para iniciar estos procesos de ETL se eligieron las bases de datos con mayor madurez y mayor cantidad de datos, en particular las que almacenan información del ciclo de vida académica. En nuestro caso, esas bases de datos corresponden a las usadas en los procesos de inscripción y admisión, matrícula y registro, donde se consignan los datos del principal actor universitario: los estudiantes. 2da: Configuración de la plataforma tecnológica, definición de datos y estadísticas Paralelo a la etapa 1, se planificó la identificación de las variables que la Universidad históricamente ha reportado en sus boletines estadísticos e informes de gestión y resultados sociales, que han sido de frecuente consulta interna y externa, y que han demandado la generación de reportes diferentes dependiendo de la instancia a la que se consultara.

Como resultado, se obtuvo un listado inicial de aproximadamente 20 variables relacionadas con las principales cifras institucionales y que dan respuesta a los ejes misionales. Luego, se inició con la definición y construcción de los tableros, para lo cual se trabajó en conjunto con las Vicerrectorías y Direcciones responsables de la información a fin de acordar qué, cómo y cuándo se continuarán presentando los datos. Este proceso de creación de los tableros con la intención de convertirse en un único punto de referencia institucional para contemplarse a sí misma desde los datos, implicó la implementación de procesos de estandarización y normalización de los datos, que permitieron llegar al consenso general de la estadística a presentar. 3ra: Visualización y montaje del micrositio Con el propósito de lograr vincular a todos los integrantes de la comunidad universitaria en este esfuerzo de “Analítica para todos”, se dio especial importancia a la identificación de las herramientas de software que permitieran visualizaciones agradables e intuitivas para el usuario, al punto que él mismo pudiera interactuar y generar sus propias gráficas partiendo de la base ofrecida. Como resultado, se eligió Tableau por permitir un fácil entendimiento de los datos presentados, trascendiendo los reportes estadísticos clásicos, al tiempo que permite al

usuario tener su propia experiencia en el reconocimiento de la universidad a partir de los datos institucionales, bien sea desde en el entendimiento de patrones históricos mediante series de tiempo, visualización de información georreferenciada, detalle en tablas, información adicional en leyendas, filtros personalizables, entre otros.

Experiencias y aprendizajes Enriquecer los datos exige:

- Enriquecer a los usuarios. No solo es importante divulgar, también lo es capacitar. Por esto, se adoptó una nueva etapa: “Talleres Data UdeA”, cuyo propósito es el de acercar a la comunidad al uso de este nuevo producto.
- Preparar los datos. Es necesaria la inversión de mucho tiempo del recurso humano, que debe estar capacitado en buenas prácticas para la limpieza de datos. En particular, identificamos que es fundamental validar de manera iterativa que los datos sean realmente correctos de acuerdo con la realidad histórica universitaria, para lograr la confianza del usuario en el dato publicado.
- Encontrar nuevos datos. La búsqueda continua, limpieza y combinación de diferentes tipos de datos permite la creación de nuevos grupos de datos con atributos adicionales.
- Blindarlos ante cambios en la infraestructura que los soporta. El software de analítica se encuentra en constante evolución y actualización, por lo que es importante mantenerse a la vanguardia de estos cambios tecnológicos y adoptar los que sean necesarios para que la infraestructura y el software que soporta los datos no resulte obsoleta. Por esta misma razón, es importante ampliar los tiempos destinados a la configuración y puesta en funcionamiento del entorno para analítica de datos, ya que esta deberá ser una actividad iterativa dentro de todo el tiempo que dure un proyecto de este tipo.
- Blindarlos ante accesos no autorizados. Por la naturaleza de la información, que hace parte del patrimonio público de la Universidad, es importante añadir una etapa para la identificación e implementación de sistemas de seguridad para la información almacenada en las bodegas de datos y expuesta en las plataformas de visualización.

#### **4. Resultados de la práctica**

Consideramos importante compartir, en un espacio como el habilitado por Telescopi, nuestros principales resultados:

- Hoy la Universidad de Antioquia cuenta con una plataforma escalable de analítica para Big Data, comparable a la de grandes empresas e instituciones de la región.
- El entorno tecnológico logrado permite abordar el análisis de diversas fuentes y tipos de datos, habilitando capacidades para procesar, interpretar, divulgar y visualizar tanto información de la gestión universitaria, como investigaciones y actividades de extensión que se adelantan en la institución.
- El logro de proveer datos y estadísticas avalados institucionalmente, facilitó la vinculación con grupos de investigación de la institución, cuyas agendas se cruzan con necesidades de

la gestión para aportar entendimiento a fenómenos universitarios que son de interés continuo, tales como la permanencia estudiantil, graduación, oferta de nuevos programas, entre otros.

- Integrar a Data UdeA con el portal universitario, brindando un único punto de acceso a la imagen digital mundial de la UdeA, donde se presentan las principales cifras institucionales, en un formato agradable, interactivo y fácil de usar, que permite a los miembros de la comunidad universitaria y públicos externos entender a la Universidad de Antioquia a través de sus datos.
- Validar el orden, excelencia y calidad que ha caracterizado a la UdeA en sus reportes estadísticos.
- Aportar Data UdeA como una herramienta que protege la tradición y patrimonio de información de la UdeA, al tiempo que habilita nuevas capacidades para que toda la comunidad aporte a la comprensión de los fenómenos históricos institucionales.
- Algunos datos que dan cuenta del uso de la información disponible en Data UdeA desde su primera publicación en octubre de 2018, a la fecha, son:
  - Tablero “Matriculados”: 1227 visitas
  - Tablero “Procedencia de los estudiantes”: 937 visitas
  - Tablero “Admitidos”: 827 visitas
  - Tablero “Inscritos”: 458 visitas
  - Tablero “Grupos de investigación”: 430 visitas
- También los espacios habilitados en el último mes para el acercamiento de los usuarios a Data UdeA:
  - Inscritos a “Talleres Data UdeA”: 65
- Cerca de siete (7) bodegas de datos instaladas en la plataforma tecnológica de Big Data de Data UdeA, con información institucional, en la mayoría de los casos desde 1994 a 2019 de los procesos de inscripción, admisión, matrícula y registro, contratación de los profesores, producción científica derivada de la asignación de puntaje, facturación, bienestar universitario, investigadores, grupos de investigación, revistas científicas, movilidad internacional, algunas de ellas provenientes de sistemas de información estructurados y otros de información semiestructurada en formato .csv o .xls e incluso en formato texto.

## **5. Evaluación y revisión de la práctica**

El producto disponible para la comunidad universitaria “Data UdeA” ha sido evaluado de manera cualitativa por nuestros principales usuarios que son los directivos universitarios, alcanzando siempre una valoración positiva expresada públicamente en todos los espacios de divulgación en donde ha sido presentado el producto. Dada la buena acogida que ha tenido el proyecto y el producto “Data UdeA”, cada vez se suman más fuentes de información en diversos formatos, lo que permite una integración de un alto volumen de datos concernientes a la gestión universitaria, siempre en función del entendimiento de la Universidad a través de los datos. Adicionalmente, por iniciativa de los usuarios, fue necesario incluir la etapa de capacitación a todas las dependencias administrativas y académicas sobre el uso de la herramienta y la cultura de los datos. De manera cuantitativa, la evaluación de la práctica ha sido bajo los estándares y protocolos de la Dirección de

Planeación y Desarrollo Institucional, principalmente de la División de Gestión Informática y de la División de Planes y Proyectos de la Universidad.

Se presentan los resultados obtenidos en la evaluación:

- Objetivo 1: Instalar una plataforma de big data para la Universidad de Antioquia. ○ Indicador: Porcentaje de la plataforma instalada y configurada. ○ Meta: 100%. ○ Valor alcanzado: 95%. ○ Porcentaje de cumplimiento: 95%.
- Objetivo 2: Instalar capacidades para la observación y el análisis permanente de la gestión universitaria y de su entorno. ○ Indicador: Número de tableros instalados en la plataforma institucional. ○ Meta: 15. ○ Valor alcanzado: 14. ○ Porcentaje de cumplimiento: 93%.
- Objetivo 3: Difundir para el público general las principales cifras institucionales que dan cuenta de la gestión universitaria. ○ Indicador: Número de estrategias de apropiación desarrolladas. ○ Meta: 5. Valor alcanzado: 2. ○ Porcentaje de cumplimiento: 40%.

## **6. Carácter Innovador de la práctica**

Se considera que los resultados son innovadores, puesto que se apropian tecnologías de tendencia mundial principalmente usadas en el sector privado y productivo al sector de Educación Superior Pública, reconociendo el volumen de datos que se generan y el potencial de ellos en función de la excelencia académica, la movilidad social y el carácter transformador de la sociedad, al que se debe la Universidad Pública. Se ha logrado que la información estadística de la Universidad de Antioquia, institución pública, se encuentre en un solo lugar. Que los datos y cifras presentadas sean el reflejo de la dinámica institucional y que exista un consenso entre las diferentes áreas sobre la consistencia de los datos presentados.

El uso de una plataforma de Big Data para la gestión de la información institucional ha permitido que la actualización, el cruce de variables, el tratamiento de datos, la visualización sea un proceso fácil de realizar con un alto nivel de automatización y con la posibilidad de ser escalable y modificable rápidamente, por el carácter de software libre, en la mayoría de sus componentes. Por último, la presentación de la información, con un formato de infografía interactiva más que de tablero de control, permite que el usuario construya sus propios indicadores, en donde prima el carácter histórico de la información presentada, con la intención de conocer muy bien el comportamiento de los datos y pensar en la proyección de los mismos.

## **7. Divulgación de la práctica**

En conjunto con la Dirección de Comunicaciones de la Universidad, se desarrolló la estrategia de divulgación entre las cuales destacamos, la presencia de Data UdeA, donde se presentan tableros interactivos con las principales cifras institucionales, soportados en la plataforma tecnológica de Big Data. Tener presencia en el portal Universitario nos permite acercarnos a diferentes públicos, no solo a los de nivel directivo. Como estrategias en los medios de divulgación contamos con la publicación permanente de la cifra de la semana en

los boletines digitales, contamos con vídeos que se publican en las redes sociales institucionales y que dan a conocer a nuestros públicos, qué encontrar en Data y qué usos posibles pueden darle a la información, también realizamos publicaciones en el suplemento Alma Mater, por último como parte del desarrollo de cultura alrededor del uso de los datos y la analítica, realizamos talleres permanentes para los analistas de la Vicerrectorías y Direcciones de nivel central y para la Red de comunicadores de la Universidad.

## **8. Fuentes complementarias**

<http://www.udea.edu.co/wps/portal/udea/web/inicio/institucional/data-udea/formacion/ubicacion>