

licencia, podía devolverla a los fabricantes para resolver los problemas y no necesitaba abogados para usarlo”.

En el 2004 trabajó el sistema Nutch, un motor de búsqueda de código abierto basado en Lucene. Pero solo era empleado por académicos, conectado a Google, y la idea era que pudiera usar algo de código abierto. Para apoyar las búsquedas de ese motor, Cutting ideó Hadoop, cuyo nombre proviene del elefante de juguete de su hijo. Comenzó con cinco máquinas y luego 25 que podía dejar corriendo todo el día, pero se perdía información.

En el 2006 ingresó a Yahoo! adonde encontró ingenieros e investigadores con sus mismos intereses, que le proporcionaron más computadoras y un equipo que pasó de tres a doce personas. “Así pude contar con cientos de computadores para poder hacer este *software* viable”. Con su grupo puso a producir un sistema clúster de almacenamiento y análisis basado en Hadoop para las operaciones de misión crítica de negocio de la compañía.

En el 2007 llegó a su meta de tener miles de computadores trabajando en para-

“Siempre quiso producir algo que se pudiera aprovechar sin pagar, sin licencias, y ayudar a solucionar diversos problemas. Le gustaba escribir *software* y que la gente lo usara y descubrió que era una forma extraordinaria de trabajar”.

lelo (algo que no se había hecho y era más barato) y un *software* gratis. Antes, la gente pagaba por almacenar datos, por lo que muchas personas comenzaron a buscar el *software* para empresas y fábricas.

En el 2009, se retiró de Yahoo! para vincularse a Cloudera, una plataforma libre con soporte comercial que se emplea no solo para manejar datos, sino para procesar e interpretar los que generan sensores de las máquinas (luego hay que analizarlos para el mejoramiento de la empresa). En esa época era común que las organizaciones usaran *software* de tecnología Enterprise para ejecutar esas tareas, pero esta es muy costosa

e inflexible y a menudo no querían mezclarse con quienes operaban *open source*.

Era el momento en que, recuerda, se diseñaba la aplicación del *software* y luego se alimentaba con datos, lo cual no es productivo; es mejor introducir la información y hacer experimentos, mirar si funciona el programa para luego empezar a usarlo. A esta nueva manera de trabajar Cutting la llama “la transformación digital de los negocios”. Y en ella, los ingenieros que desarrollaron el *open source* pueden ayudar a quienes controlan la evolución del sistema en las empresas. Es una manera de mantener este proyecto constantemente vivo. ■

Big data contra el lavado de activos

Luis Edmundo Suárez, director de la Unidad de Información y Análisis Financiero (UIAF), habló de cómo el manejo de enormes volúmenes de datos les ha permitido mejorar los resultados en sus investigaciones.

La lucha contra el lavado de activos es global porque es un problema que está afectando todas las economías. Según las Naciones Unidas los ingresos de los criminales se acercan a 2.6 trillones de dólares, de los cuales introducen a la economía mundial 1.9 trillones de dólares, una cifra superior al Producto Interno Bruto (PIB) de cualquier país de América Latina.

Para luchar contra este flagelo se han conformado agrupaciones a nivel mundial y regional que comparten información so-

bre redes de tráfico de drogas, armas o terrorismo. Se destaca el grupo Egmont, con sede en París, que regula el tema y congrega unidades de inteligencia financiera de 149 países.

De los 2.6 trillones de ingresos producto de delitos se logra interceptar cerca de 0.20 %, explicó Luis Edmundo Suárez durante el Foro. Es decir 20 centavos de cada 100 dólares. Y en Colombia, de acuerdo con una investigación del Banco Mundial hecha entre 1983 y el 2013, hubo un marcado ascenso de 1999 al 2000 y los ingresos generados por bandas de crimen co-



Foto: Natalia Fernanda Madrid Vidales

Luis Edmundo Suárez, director de la Unidad de Información y Análisis Financiero (UIAF).

mún y narcotráfico estuvieron cerca al PIB. “En el país, hoy se lavan entre 16 y 20 billones de pesos cada año por narcotráfico”.

“ En el país, hoy se lavan entre 16 y 20 billones de pesos cada año por narcotráfico”.

En la UIAF trabajan en el nivel estratégico, táctico y operativo y han logrado quitarles a los lavadores, en medidas cautelares y sentencias, \$6.9 billones en cinco años. La cifra está 70 veces por encima del promedio mundial, dijo.

Para ello han cruzado información de más de 24 bases de datos, y según Suárez, el promedio de logros en el mundo al

interceptar solo 20 centavos de cada 100 dólares lavados no es suficiente, de modo que con el fin de aumentar el monto se necesitan más fuentes de información, no solo reportes sospechosos y de retroalimentación del sector financiero. Por eso en la UIAF crearon un modelo productivo, proactivo, amplio y de direcciones que obedece a un concepto de *big data*. Es decir que en el modelo *big data* colocaron más sensores y así pudieron aumentar las interceptaciones de dineros lavados en un 600 %.

Para la Unidad, la plataforma Hadoop ofrece con *big data* beneficios cruciales para la economía, con el lago de datos que permite tener rápidamente la información, con la analítica avanzada para poder hacer inteligencia artificial y análisis de texto y con el autoservicio y la escalabilidad. ■

MINE, nueva maestría del DISC

Este programa académico del Departamento de Ingeniería de Sistemas y Computación (DISC) se encamina a desarrollar en los ingenieros de sistemas y profesionales afines habilidades para adquirir, procesar, relacionar, integrar, analizar y visualizar información. Así esta se convierte en un activo de las organizaciones, apoya su estrategia de toma de decisiones y les añade valor ante la competencia.

En los últimos dos años, la humanidad ha producido la misma cantidad de información que generó en toda su historia. Con esta contundente frase la ingeniera y profesora Claudia Jiménez explica la importancia de la Maestría en Ingeniería de Información (MINE) que ofrece el DISC desde agosto del 2016.

La profesora Jiménez coordina el nuevo programa académico de Uniandes y no duda de la creciente relevancia de las tecnologías computacionales para que, a partir del ciclo de vida de la información,

entreguen soluciones oportunas y pertinentes, que diferencien a las organizaciones frente a sus competidores.

En las últimas décadas ese ciclo de vida ha cambiado: antes se circunscribía al histórico o la operación de cada empresa, mientras que ahora, con la convergencia tecnológica manifiesta en internet y en los dispositivos móviles, la toma de decisiones empresariales comienza por diseñar procesos que ayuden a descubrir información. “Esta puede venir de afuera, no la controla, no sé cuándo se empezó a generar ni cómo ni quién la producirá,

pero puede tener un impacto muy alto en mi negocio”, dice la profesora.

Tal cambio hizo que las tecnologías tradicionales, que siguen vigentes y son necesarias, resulten insuficientes para procesar, analizar, relacionar y contextualizar los datos y, finalmente, crear escenarios para tomar mejores decisiones de negocio, lo que constituye el eje de la Ingeniería de Información. Durante varias décadas se desarrollaron tecnologías orientadas al desarrollo de sistemas transaccionales confiables para apoyar la operación de las empresas, de suerte que, por ejemplo, el usuario tiene certeza de que su dinero no se pierde en las transacciones bancarias o de que sí tendrá cupo cuando compre un tiquete aéreo.

El reto es integrar nuevos tipos de información como video, imágenes, texto y datos de sensores desde infinidad de dispositivos, que se constituyen en fuente muy valiosa para la toma de decisiones. Detrás de esos procesos hay soluciones computacionales que se soportan mediante