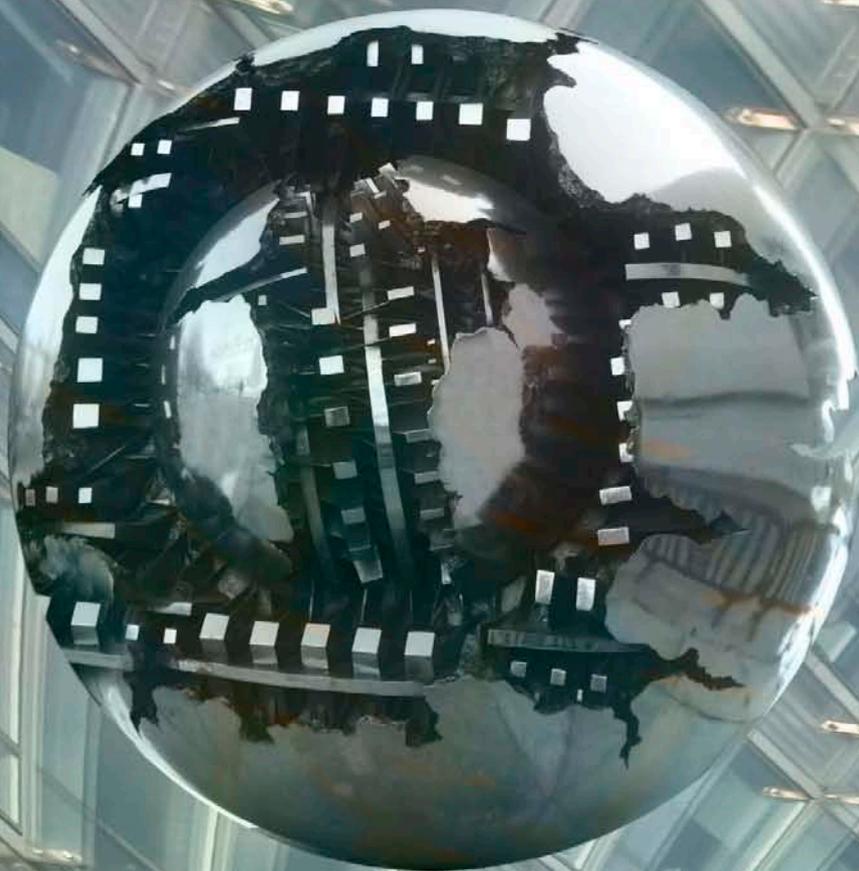


NOVÁTICA

Revista de la Asociación de Técnicos de Informática

Nº 237, julio - octubre 2016, año XLII



Big Data



La **vigesimotercera edición de las Jornadas sobre Enseñanza Universitaria de la Informática (JENUI 2017)** están promovidas por la Asociación de Enseñantes Universitarios de la Informática (AENUI) y organizadas por el Departamento de Ingeniería de Sistemas Informáticos y Telemáticos y la Escuela Politécnica de la Universidad de Extremadura.

El objetivo de JENUI es promover el contacto, el intercambio y la discusión de conocimientos y experiencias entre profesores universitarios de informática y grupos de investigación. El debate sobre el contenido de los programas y los métodos pedagógicos empleados y los enfoques innovadores que se materializan en este foro están orientados a mejorar la docencia de la informática en las universidades.

Las JENUI se articulan mediante contribuciones de tipo ponencia completa, póster y demostración de recursos docentes, así como tertulias y otras actividades que estimulan el contacto personal entre los docentes.

Fechas importantes

23/01/2017	Fecha límite para envío de resúmenes
13/02/2017	Fecha límite para envío de trabajos
07/04/2017	Notificación del resultado del proceso de revisión
24/04/2017	Fecha límite para envío de trabajos definitivos
13/04/2017 al 30/05/2017	Inscripción temprana
31/05/2017 al 03/07/2017	Inscripción normal
04/07/2017	Taller JENUI 2017
05/07/2017 al 07/07/2017	JENUI 2017



Organizadores



AENUI
Asociación de
Enseñantes
Universitarios de la
Informática

Colaboradores



**Re
Visión**

Revista de
Investigación en
Docencia Universitaria
de la Informática



Asociación de Técnicos
de Informática



Revista de la Asociación de Técnicos de Informática



Sociedad Educación
CAPÍTULO ESPAÑOL



Sistedes
Sociedad de Ingeniería de Software
y Tecnologías de Desarrollo de Software



Novática, revista fundada en 1975 y decana de la prensa informática española, es el órgano oficial de expresión y formación continua de **ATI** (Asociación de Técnicos de Informática), organización que edita también la revista **REICIS** (Revista Española de Innovación, Calidad e Ingeniería del Software).

<<http://www.ati.es/novatica/>>
<<http://www.ati.es/reicis/>>

ATI es miembro fundador de **CEPIS** (Council of European Professional Informatics Societies), representa a España en **IFIP** (International Federation for Information Processing) y es miembro de **CLEI** (Centro Latinoamericano de Estudios de Informática) y de **CECJA** (Confederation of European Computer Associations). Asimismo tiene un acuerdo de colaboración con **ACM** (Association for Computing Machinery) y colabora con diversas asociaciones informáticas españolas.

Consejo Editorial

Guillermo Alstina González, Pere Lluís Barabà, Miquel García-Menéndez (presidente del Consejo), Ernest Gijón Gil, Juan Hernández Basora, Silvia Leal Martín, David Moya Alvarez, Francesc Noguera Puig, Andrés Pérez Payeras, Víkto Pons i Colomer, Daniel Raya Demidoff, Jordi Roca i Marimón, Jorge Daniel Vigo López, Juan Carlos Vigo López

Coordinación Editorial

Llorenç Pagés Casas <pages@ati.es>

Composición y autoedición

Impresión Olfset Derra S. L.

Traducciones

Grupo de Lengua e Informática de ATI <<http://www.ati.es/gi/lengua-informatica/>>

Administración

Tomás Brunete, María José Fernández, Enric Camarero

Secciones Técnicas - Coordinadores

Accesibilidad

Emmanuelle Gutiérrez y Restrepo (Fundación Sidar), <emmanuelle@sidar.org>

Loïc Martine Normand (Fundación Sidar), <loic@sidar.org>

Acceso y recuperación de la información

José María Gómez Hidalgo (Pragsis Technologies), <jmgomez@pragsis.com>

Enrique Puertas Sanz (Universidad Europea de Madrid), <enrique.puertas@universidadeuropea.es>

Administración Pública electrónica

Francisco López Crespo (MAE), <flc@ati.es>

Sabastià Justicia Pérez (Diputación de Barcelona) <sjusticia@ati.es>

Arquitecturas

Enrique F. Torres Moreno (Universidad de Zaragoza), <enrique.torres@unizar.es>

José Filich Cardo (Universidad Politécnica de Valencia), <jfilich@disca.upv.es>

Auditoría SITIC

Marina Tourinho Troitino, <marinatourinho@marinatourinho.com>

Sergio Gómez-Landero Pérez (Endesa), <sergio.gomezlandero@endesa.es>

Derecho y tecnologías

Elena Davara Fernández de Marcos (Davara & Davara), <edavara@davara.com>

Enseñanza Universitaria de la Informática

Cristóbal Pareja Flores (DSIP-UCM), <cpareja@sip.ucm.es>

J. Ángel Velázquez Irujo (DLSI I, URJC), <angel.velazquez@urjc.es>

Entorno digital personal

Andrés Marín López (Univ. Carlos III), <amarin@it.uc3m.es>

Diego Gachet Páez (Universidad Europea de Madrid), <gachet@uem.es>

Estándares Web

Encarna Quesada Ruiz (Virat), <encarna.quesada@virat.com>

José Carlos del Arco Prieto (TCP Sistemas e Ingeniería), <jcarco@gmail.com>

Gestión del Conocimiento

Joan Baiget Solé (Cap Gemini Ernst & Young), <joan.baiget@ati.es>

Gobierno corporativo de las TI

Manuel Palao García-Suelto (ATI), <manuel@palao.com>

Miguel García-Menéndez (ITI) <mgarciamenendez@itirendsinstitute.org>

Informática y Filosofía

José Ángel Olivás Varela (Escuela Superior de Informática, UCLM), <joangel.olivas@uclm.es>

Roberto Feltre Orea (UNED), <rfeltre@gmail.com>

Informática Gráfica

Miguel Chover Selles (Universitat Jaume I de Castellón), <chover@lsi.uji.es>

Roberto Vivó Hernando (Eurographics, sección española), <rvivo@dsic.upv.es>

Ingeniería del Software

Luis Fernández Sanz, Daniel Rodríguez García (Universidad de Alcalá), <luisfernandez.daniel.rodriguez@uah.es>

Inteligencia Artificial

Vicente Botti Navarro, Vicente Julián Inglada (DSIC-UPV), <vbotti.vinglada@dsic.upv.es>

Interacción Persona-Computador

Pedro M. Latorre Andrés (Universidad de Zaragoza, AIPO), <platorre@unizar.es>

Francisco L. Gutiérrez Vela (Universidad de Granada, AIPO), <fgutierrez@ugr.es>

Lengua e Informática

M. del Carmen Ugarte García (ATI), <cugarte@ati.es>

Lenguajes Informáticos

Oscar Belmonte Fernández (Univ. Jaime I de Castellón), <obelform@lsi.uji.es>

Inmaculada Coma Talay (Univ. de Valencia), <inmaculada.coma@uv.es>

Lingüística computacional

Xavier Gómez Guinovart (Univ. de Vigo), <xgg@uvigo.es>

Modelado de software

Jesus Garcia Molina (DIS-UM), <jmolina@um.es>

Gustavo Rosca (LIFA-UNLP Argentina), <gustavo@sol.info.unlp.edu.ar>

Mundo estudiantil y jóvenes profesionales

Federico G. Mon Trotti (RITSJ), <fgm.fede@gmail.com>

Mikel Salazar Peña (Area de Jóvenes Profesionales, Junta de ATI Madrid), <mikelbo_uni@yahoo.es>

Seguridad

Rafael Fernández Calvo (ATI), <rflcalvo@ati.es>

Miguel Sarrías Griño (ATI), <miguel@sarrias.net>

Redes y servicios telemáticos

Juan Carlos López López (UCLM), <juancarlos.lopez@uclm.es>

Ana Pont Sanjuán (UPV), <apont@disca.upv.es>

Robótica

José Cortés Arenas (Sopra Group), <joscortea@gmail.com>

Juan González Gómez (Universidad Carlos III), <juan@iearobotics.com>

Seguridad

Javier Arellano Bertolin (Univ. de Deusto), <jarellito@deusto.es>

Javier López Muñoz (ETSI Informática-UMA), <jlm@cc.uma.es>

Sistemas de Tiempo Real

Alejandro Alonso Muñoz, Juan Antonio de la Puente Alfaro (DIT-UPM), <[@dit.upm.es](mailto:aalonso.jpunte)>

Software Libre

Jesús M. González Barahona (GSYC-URJC), <jgb@gysc.es>

Fernando Tricas García (Universidad de Zaragoza), <tricas@unizar.es>

Tecnologías para la Educación

Juan Manuel Dodero Beardo (UC3M), <dodero@inf.uc3m.es>

César Pablo Córcoles Briongo (UOC), <ccorcoles@uoc.edu>

Tecnologías y Empresa

Didac López Vinas (Universitat de Girona), <didac.lopez@ati.es>

Alonso Álvarez García (TID) <aag@tid.es>

Tendencias tecnológicas

Gabriel Martí Fuentes (Interbits), <gabi@atinet.es>

Juan Carlos Vigo (ATI) <juancarlosvigo@atinet.es>

TIC y Turismo

Andrés Agayo Maldonado, Antonio Guevara Plaza (Univ. de Málaga), <[@cc.uma.es">agayo.guevara@cc.uma.es](mailto:agayo.guevara)>

Las opiniones expresadas por los autores son responsabilidad exclusiva de los mismos. **Novática** permite la reproducción, sin ánimo de lucro, de todos los artículos, a menos que lo impida la modalidad de [cc-by](https://creativecommons.org/licenses/by/4.0/) o [cc-by-nc](https://creativecommons.org/licenses/by-nc/4.0/) elegida por el autor, debiéndose en todo caso citar su procedencia y enviar a **Novática** un ejemplar de la publicación.

Coordinación Editorial, Redacción Central y Redacción ATI Madrid
Gutiérrez de Cetina 24, 28017 Madrid • Tlf: 91 4029391 • novatica@ati.es

Administración y Redacción ATI Cataluña

Calle Avila 50, 3a planta, local 9, 08005 Barcelona

Tlf: 934 125235 <secretgen@ati.es>

Redacción ATI Andalucía <secretand@ati.es>

Redacción ATI Galicia <secretgal@ati.es>

Suscripción y Ventas <novatica.subscriptions@atinet.es>

Publicidad Gutiérrez de Cetina 24, 28017 Madrid

Tlf: 91 4029391 <novatica@ati.es>

Imprenta: Impresión Olfset Derra S.L., Lluís 41, 08005 Barcelona.

Depósito legal: B 15.154-1975 -- ISSN: 0211-2124; CODEN NOVAC

Portada: "La decisión" - Concha Arias Pérez / © ATI

Diseño: Fernando Agresta / © ATI 2003

editorial

La hora del Big Data

> 02

Periodicidad de Novática desde julio de 2016 hasta junio de 2017

noticias de ATI

Nombramiento de la nueva Directora de Novática

> 03

en resumen

Un agradecimiento muy especial para todos nuestros colaboradores

> 03

Llorenç Pagés Casas

noticias de IFIP

Asamblea General de IFIP

> 04

Ramon Puigjaner Trepap

WITFOR 2016

> 05

Ana Pont Sanjuán

Noticias del TC9: ICT and Society

> 05

Ignacio Gil Pechuán

Reunión anual del TC2 "Software: Theory and Practice"

> 06

Antonio Vallecillo Moreno

actividades de ATI

X Edición del Premio Novática: Entrega del premio al autor ganador

> 06

monografía

Big Data

Editores invitados: José María Gómez Hidalgo y Ricardo Baeza-Yates

Presentación. Big Data: Conceptos y aplicaciones

> 09

José María Gómez Hidalgo, Ricardo Baeza-Yates

Datos masivos en la Web

> 12

Ricardo Baeza-Yates

Big Data: Preprocesamiento y calidad de datos

> 17

Salvador García, Sergio Ramírez-Gallego, Julián Luengo, Francisco Herrera

Internet de las Cosas: La minería de flujos de datos masivos en tiempo real

> 24

Albert Bifet, Jesse Read

Análisis Big Data en sistemas de computación de alto rendimiento: Tecnologías, herramientas y ejemplos

> 31

Alexey Cheptsov, Bastian Koller

Big Data y sistemas de recomendación

> 39

David C. Anastasiu, Evangelia Christakopoulou, Shaden Smith, Mohit Sharma, George Karypis

Estudio sobre la escalabilidad del algoritmo de agrupamiento estructural paralelo para redes en Big Data

> 46

Weizhong Zhao, Gang Chen, Venkata Swamy-Martha, Xiaowei Xu

Introducción a la analítica de texto con Spark

> 53

José María Gómez Hidalgo

Cómo mejorar el conocimiento de tu audiencia: Experiencias de la CCMA en un entorno Big Data

> 60

Xavier Ferrándiz Bofill, Alberto Alejo Marcos

Privacidad en la analítica masiva de datos

> 65

José María del Álamo Ramiro, Esmeralda Saracibar Serradilla, Emilio Aced Féliz

secciones técnicas

Tendencias Tecnológicas

¿Nos está haciendo felices la tecnología?

> 70

Dorian Peters

Referencias autorizadas

> 72

sociedad de la información

Programar es crear

El problema del robot de exploración de Marte

> 78

(Competencia UTN-FRC 2014, problema 5, enunciado)

Julio Javier Castillo, Diego Javier Serrano, Marina Elizabeth Cárdenas

Discos duros

> 79

(Competencia UTN-FRC 2015, problema A, solución)

Julio Javier Castillo, Diego Javier Serrano, Marina Elizabeth Cárdenas

asuntos interiores

Coordinación editorial / Programación de Novática / Socios Institucionales

> 80

Monografía del próximo número: "Seguridad digital"

La hora del *Big Data*

Informática puede venir de **información**, y la información se nutre de **datos**. Los datos en sí mismos suelen ser poco significativos dado que su valor se manifiesta cuando los relacionamos, porque es entonces cuando se transforman en información.

Esta reflexión, posiblemente elemental, configura en parte la historia de nuestra informática, que consiste, entre otros aspectos, en ofrecer información elaborada. Aunque para completar el objetivo que todo proceso informático requiere faltará un tercer elemento que es la **tecnología**, absolutamente necesaria para poder acceder a la información de forma eficiente.

En esta etapa actual de nuestra profesión impera la necesidad de obtener información eficiente y en este cometido toma papel predominante la cultura "*Big Data*". Por ello, el posicionamiento de los profesionales de la Informática que necesitan estar al corriente de las novedades del sector y de sus posibilidades, se debe orientar a entender la utilidad del "*Big Data*" en la empresa, y a traducir en términos de rentabilidad las posibilidades de su aplicación como concepto, así como también conocer el uso de la tecnología que acompaña al mismo. Y esto requiere formación y recursos.

¿Más formación y más recursos todavía? dirán algunos informáticos seniors, hartos del cambio tecnológico continuo. Y así es. En nuestra profesión, lo cotidiano es cambiar por lo general, nuestras competencias en sistemas operativos, lenguajes de programación y conocimientos de estructuras de información, todo ello a muy corto plazo, y sin tregua hacia un recorrido que no tiene por qué tener un final.

Las estructuras de datos siempre han sido motor de cambio en nuestra profesión. De los sistemas de ficheros a los modelos de bases de datos, de las primeras estructuras, jerárquicas o Codasyl, a las aceptadas como más eficaces, y basadas en el modelo relacional. Ésto solo fue un inicio que evolucionó en términos de eficacia, ya que aparecieron los modelos documentales y los orientados a objetos. Todo ello en paralelo con la evolución del hardware y los lenguajes de programación.

El "*Big Data*" es el protagonista del momento actual en gestión de los datos, porque dispone de muchas herramientas para tratar la información de forma masiva, que es lo que define este nuevo concepto. Y por ello se va a imponer probablemente como una de las herramientas más útiles para tomar decisiones y prever futuras operaciones de largo alcance relacionadas con la información obtenida.

Pero existe otra faceta que suscita reflexión sobre esta nueva tecnología, que es la gestión de unos datos, donde va a ser difícil averiguar si los resultados obtenidos se dan a conocer por el gran público, y si atentan a la seguridad de las personas, empresas y de la sociedad en general. Y también si los datos manejados son obtenidos con el consentimiento de su uso por parte de los ciudadanos implicados.

Si las reglas de juego son laxas, el poseedor de la información elaborada a través de procesos de tratamiento de información de forma masiva, y que además también domina parte de determinadas redes sociales, nos puede llegar a conocer y manejar sin que seamos conscientes de ello, con más intensidad de lo que ya lo debe estar haciendo actualmente.

No hay que ser demasiado observador para ver paseando por nuestras calles ciudadanos que realizan, por lo general, una actividad casi frenética: llenar de forma continua y a través de sus *smartphones* ficheros de datos e imágenes, algunos a bases de datos privadas, y otros a determinados bancos de datos. Y lo que se está publicando en estos bancos de datos son referencias a quienes somos, que nos apetece, cuáles son nuestras intenciones personales a corto y medio plazo, y por qué estaríamos dispuestos a jugarlos el tipo. Todo un manjar para los servicios de inteligencia y las empresas que comercializan productos dirigidos al gran público.

Por lo demás, centrándonos más ya en la evolución de los productos y conceptos para tratar la información, y aprendiendo del curso de la Informática y de nuestra experiencia como profesionales, el "*Big Data*" no tiene por qué ser el final de ninguna historia, ya que sabemos que lo único estable que tiene la gestión de datos es su evolución continuada.

Podemos imaginar ya un futuro de evolución del "*Big Data*" con la aparición de nuevos conceptos, del tipo: "*Very Large Big Data*" o "*Super Big Data*" que pueden estar a la vuelta de la esquina.

La Junta Directiva General de ATI

Periodicidad de Novática desde julio de 2016 hasta junio de 2017

Con la publicación de este número, el cual significa la despedida de Llorenç Pagés Casas, quien ha venido dirigiendo la revista desde julio de 2006, se inicia una fase de remodelación de la revista, acorde con los nuevos tiempos y con los retos presupuestarios que debe afrontar ATI.

Para facilitar esta transición, ya iniciada desde el nombramiento de Encarna Quesada Ruiz como nueva

directora de la revista, y poder realizar mejor todas las adecuaciones necesarias, hemos decidido cambiar provisionalmente la periodicidad de publicación de la revista, pasando a ser ésta cuatrimestral entre julio de 2016 y junio de 2017. Por lo tanto, en este periodo además del número actual se publicarán otros dos números más: Novática 238 (noviembre 2016-febrero 2017) y Novática 239 (marzo 2017-junio 2017).

Es de preveer que a partir de julio de 2017 podamos reanudar la periodicidad trimestral que nos ha sido habitual desde 2012, de modo que Novática 240 y Novática 241 se puedan publicar también a lo largo de 2017.

Nombramiento de la nueva Directora de Novática

Nos complace informaros de que en la última reunión de Junta Directiva General, y tras el pertinente concurso, celebrado en fechas recientes, se acordó nombrar a **Encarna Quesada Ruiz** como nueva Directora de *Novática*, en sustitución de **Llorenç Pagés Casas**, después de que éste comunicase a esta JDG, hace unos meses, su intención de ir cerrando un ciclo que ya se prolongaba por espacio de diez años.

Encarna Quesada es especialista en procesos de Transformación Digital. La investigación, la divulgación y el asesoramiento son los tres ejes principales de su labor, en relación a los avances e innovación en tecnolo-

gía que afectan tanto a las empresas, como a la sociedad en su conjunto.

Encarna fue Directora de W3C España, un consorcio internacional que trabaja en el desarrollo de protocolos y pautas que permiten asegurar el crecimiento futuro de la Web y hacer posible la comunicación humana, el comercio y las oportunidades para compartir conocimiento. Asimismo, ha trabajado con una serie de organismos y empresas nacionales e internacionales en proyectos de transformación digital, lo que le ha permitido adquirir una visión estratégica en innovación tecnológica, en diferentes sectores.

De forma constante, Encarna investiga y analiza la evolución de la tecnología, asesorando y tomando parte activa en proyectos empresariales para la transformación tecnológica de procesos, negocios y personas.

Finalmente, no queremos dejar pasar la ocasión sin agradecer a Llorenç su labor al frente de *Novática* durante todos estos años. Vayan para él nuestros mejores deseos para cuantos retos, personales y profesionales, aborde de aquí en adelante.

La Junta Directiva General de ATI

En resumen Un agradecimiento muy especial para todos nuestros colaboradores

Llorenç Pagés Casas

Coordinación Editorial de *Novática*

Cuando en el año 2006, **Rafael Fernández Calvo** me dio a conocer que me consideraba un candidato idóneo para sucederle al frente de *Novática*, he de confesar que me dio un vuelco el corazón al tiempo que se abrían ante mí un montón de ilusiones y buenas expectativas. Se me vino a la mente en aquel momento el modo en el que, a finales de los años 1970 y sin haber cumplido todavía los 18 años, esperaba con devoción la llegada vía postal de la revista a casa para devorarla con fruición tratando de descubrir las claves de esa “mitad ciencia-mitad arte” que empezaba ya a ser mi profesión.

Una vez que acepté ese para mí increíble reto y pude ganar el pertinente concurso, he tenido la suerte y el privilegio de haber podido dirigir la revista en un periodo tan largo como son algo más de 10 años. En los cuales hemos podido publicar, contando con éste, 56 números de la revista, más de la mitad de los cuales en conjunción con **UPGRADE**, una revista promovida por CEPIS donde publicábamos básicamente una versión inglesa del monográfico de *Novática*.

Han sido para mí años de intenso trabajo y mucha dedicación procurando dar lo mejor de mí mismo para mantener ese nivel y prestigio en la revista que le habían ido otorgando con el paso del tiempo todos y cada uno de mis antecesores. Ahora, en el momento de la despedida, puedo decir que, aun a pesar de que quedaron muchas cosas en el tintero, y otras muy mejorables en diversos aspectos, me voy con la satisfacción del deber cumplido ante ese reto que solo a base de entusiasmo e ilusión he conseguido que no me viniera excesivamente grande.

Pero me despido sobre todo con la enorme satisfacción de haber conocido y tratado

con personas maravillosas que me han facilitado muchísimo el trabajo y sin los cuales todo ese entusiasmo e ilusión no hubieran servido para nada. Así, he podido coincidir tanto con excelentes profesionales como con muchas personas que voluntariamente han cedido su tiempo en beneficio de que *Novática* publicase unos contenidos mejores y más adecuados. Ha sido para mí una experiencia impagable en este sentido.

Empezando por el principio, solo puedo tener palabras de agradecimiento para **Fernando Piera** y **Rafael Fernández Calvo**, quienes me dieron “acogida” en el proyecto en su calidad, por entonces, de vocal de *Novática* en la JDG y antecesor mío, respectivamente. Posteriormente, Rafael ha continuado colaborando con *Novática* hasta fechas bien recientes en su calidad tanto de mentor, colaborador y en muchos casos inspirador mío, haciendo que ese gratísimo recuerdo vaya a permanecer eternamente en mi memoria.

Por otra parte, tuve también la suerte de trabajar con **Jorge Llácer** (q.e.p.d) una excelente persona de quien uno piensa que la vida no lo trató como él merecía. Jorge fue maquetador de la revista durante más de 20 años y cuidó perfectamente de todos los detalles como gran profesional que era.

Otra magnífica profesional es **Concha Arias**, portadista de la revista con la misma antigüedad que la mía, y que ha tenido la virtud de irme sorprendiendo día tras día, portada tras portada, con su ingenio, imaginación y buen hacer. Estoy seguro de que, en virtud de su trabajo, nuestros lectores aprecian hoy especialmente la calidad de nuestras portadas.

He de agradecer también el apoyo encontrado en las secretarías, a lo largo de todos estos años, por parte de **Enric, Feli, Pepa** y **Tomàs**, así como en Derra, S.L., con especial mención para **Josep** y **Lola**, siendo ésta

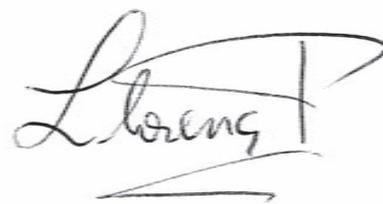
la empresa que cuidó inicialmente de la impresión de la revista, para posteriormente hacerse cargo de la maquetación, en ausencia de Jorge.

Y finalmente, los voluntarios. Esas personas que han conseguido sacar tiempo de sus quehaceres habituales para favorecer la calidad de nuestros contenidos. Especialmente mencionable es el grupo de coordinadores de sección que ha cumplido con la misión tanto de cuidar de las revisiones de artículos no solicitados como de redactar esas impagables “referencias autorizadas” que aparecen regularmente en cada número. Así como el grupo de traductores voluntarios que, con el paso del tiempo, se ha ido nutriendo de extraordinarios colaboradores, y sin cuyo trabajo no hubiéramos podido publicar los varios centenares de artículos que nos llegaron originalmente en inglés.

A todas esas personas maravillosas, muchas gracias por vuestra colaboración.

Ahora, la revista queda en unas excelentes manos. Por una parte, mi sucesora, **Encarna Quesada Ruiz**, una gran profesional, experta en comunicación, y por otra, **Miguel García Menéndez**, en su calidad de vocal de *Novática* en la JDG. Ambos han sido magníficos colaboradores de la revista durante años, y por lo tanto están preparados para los grandes retos que afronta *Novática* en el futuro. Como ya saben ellos, siempre podrán contar con mi modesta ayuda, allá donde sea necesario.

Un fuerte abrazo a todos,



Asamblea General de IFIP

Ramon Puigjaner Trepal

ex-Vicepresidente de IFIP; Catedrático Emérito de la Universitat de les Illes Balears; ex-Presidente de ATI

<putxi@uib.cat>

Entre los días 16 y 17 de octubre de 2016, se celebró en San José (Costa Rica) la Asamblea General de IFIP. A continuación, reseñamos los principales puntos tratados en la misma.

Presupuestos y finanzas

Como asuntos de trámite se aprobaron las cuentas del año 2015 y se puso de manifiesto que el informe de la auditoría no detectó ningún fallo en el tratamiento contable de los datos.

Se analizaron también las previsiones para 2016 y se observó que el déficit previsto probablemente se reducirá por la existencia de gastos no realizados según lo previsto, pero que habrá que esperar a fin de año para tener certeza de ello.

La asamblea aprobó el presupuesto de 2017 que proponía el tesorero. La Comisión de Finanzas manifestó, no obstante, el riesgo que representa plantear presupuestos con déficit con la esperanza de que los resultados del capital permitan reducirlo o transformarlo en un pequeño superávit.

Renovación del Board

Al haber llegado al final de su segundo período como vice-presidente, Ramon Puigjaner (exoficio, antiguo representante de ATI) no pudo presentarse a la reelección.

Quien se presentó a la reelección fue Max Bramer (exoficio y antiguo *chair* del TC12). Para esas dos vice-presidencias se presentaron dos candidatos Max Bramer y la representante de CLEI, Gabriela Marín, en aquel momento *councillor*. No hizo falta proceder a ninguna votación al haber tantos candidatos como posiciones. A Min Tjoa (Austria) actual secretario fue reelegido para este cargo.

Respecto de los *councillors* había 4 puestos a elección:

- Franz Rammig (exoficio y anterior representante de Alemania) elegido por la Asamblea General, finalizaba su primer período y era reelegible.
- Forrest Lin (representante de China), elegido por las sociedades miembro, finalizaba su segundo período y no era reelegible.
- Michael Goedicke (*Chair* del TC2) elegido por la *Technical Assembly*, finalizaba su primer período y era reelegible.
- Anthony Wong (representante de Australia), elegido en representación del IP3, finalizaba su primer período y era reelegible.

Para esos cargos se presentaron en tiempo y forma adecuados, respectivamente:

- Franz Rammig.
- Maria Raffai (representante de Hungría).
- Michael Goedicke.
- Anthony Wong.

Al haber un solo candidato para cada cargo, fueron elegidos automáticamente. Sin embargo, al haber pasado Gabriela Marín a Vice-presidente se admitieron candidaturas para ese cargo de *councillor*.

Los candidatos fueron:

- Jee-In Kim (representante de Corea).
- Andrew McGettrick (representante del Reino Unido).
- Jacques Sakarovitch (*Chair* del TC1).

De ellos fue elegido con claridad el representante de Corea Jee-In Kim.

También se reeligió como Tesorero a Declan Brady representante de Irlanda, al ser único candidato.

Con todos estos cambios la composición del *Board* queda como sigue:

Presidente:

- Michael Hinschey (exoficio, antiguo *chairman* del TC1).

Presidente Anterior:

- Leon Strous (Países Bajos).

Vice-presidentes:

- Max Bramer (exoficio, antiguo *chairman* del TC12).
- Gabriela Marín (CLEI).
- Yuko Murayama (Japó, y *chairwoman* del TC11).
- Kai Rannenberg (Alemania).

Secretario:

- A Min Tjoa (Austria).

Tesorero:

- Declan Brady (Irlanda).

Councillors:

- Maria Raffai (Hungría).
- Michael Goedicke (*chairman* del TC2).
- Franz Rammig (exoficio, anterior representante de Alemania).
- Anthony Wong (Australia).
- Jee-In Kim (Corea).
- Jerzy Nawrocki (Polonia).
- Raimundo Macedo (Brasil).
- Yayas Abeywickrama (Sri Lanka y *chairman* del InterYIT)

Propuesta de nueva estructura de cuotas anuales de las sociedades miembro

Como paso siguiente a la nueva estructura de miembros de IFIP, se trató de la nueva estructura de cuotas anuales de las sociedades miembro. Se presentaron tres alternativas.

Este acuerdo debía tomarse por una mayoría calificada de tres cuartas partes de los miembros presentes. Ninguna de las tres alternativas alcanzó ese nivel de votos favorables, por lo que las tres fueron rechazadas. En consecuencia para el año 2017 se mantienen las mismas cuotas que para el año 2016.

Acontecimientos generales

Respecto de los acontecimientos generales:

- Se escucharon los detalles referentes a los resultados de WITFOR 2016 realizado en San José de Costa Rica del 12 al 14 de septiembre de 2016 <www.witfor2016.org> los días anteriores a la Asamblea General. La presentación corrió a cargo de los dos *co-chairs*, Gabriela Marín y Ramon Puigjaner. Se analizaron los números de trabajos presentados y aceptados. También se dio una aproximación de los resultados económicos del evento, estando estos resultados pendientes del cierre final de las cuentas. Se acordó que de la experiencia de WITFOR con la estructura de conferencia convencional debían sacarse conclusiones para la organización de futuras ediciones.
- Se escuchó la propuesta de organización del WCC 2018 en Poznan (Polonia) a realizar en los locales de la Universidad Politécnica de Poznan. En principio, se dio luz verde a esta organización.
- Se procedió a despedir a los miembros de la Asamblea General que la dejaban por algún motivo. En particular se agradeció, entregándole un recuerdo, la tarea de Ramon Puigjaner por su labor como vice-presidente durante 7 años. También se despidieron Jan Gulliksen (*chair* del TC13), Augusto Casaca (representante de Portugal), Joe Turner (USA-ACM) y Dong-Yoon Kim (representante de Corea).
- Finalmente se procedió al relevo de la presidencia. Leon Strous (representante de Países Bajos) pasó a presidente anterior y fue substituido por el nuevo Presidente Mike Hinchey (exoficio, antiguo *Chair* del TC1). Se hizo entrega a Leon Strous de un recuerdo de sus 6 años de presidente.

WITFOR 2016

Ana Pont Sanjuán

Universidad Politécnica de Valencia; representante de ATI en el TC6

<apont@disca.upv.es>

Este evento, celebrado en San José de Costa Rica del 12 al 14 de septiembre, ha sido la sexta edición en la historia del *World Information Technology Forum* de IFIP y la primera que se ha organizado, en parte, como conferencia académica. En WITFOR 2016 se han discutido y presentado proyectos y estrategias bajo el común denominador de *ICT for Promoting Human Development and Protecting the Environment*.

Los profesores Ramon Puigjaner (vicepresidente saliente de IFIP) y Gabriela Marín (nueva vicepresidenta de IFIP) fueron los promotores y responsables del evento actuando ambos como *General-Chairs*.

El desarrollo del programa de WITFOR estuvo a cargo del profesor costarricense Francisco J. Mata y de la profesora Ana Pont, actual representante de España en el TC6 de IFIP.

Se recibieron para su evaluación un total de 45 trabajos procedentes de 25 países distintos representando a todos los continentes. Cada uno de ellos fue cuidadosamente evaluado por al menos 3 revisores de un comité científico internacional, formado por especialistas en diferentes campos de las TIC y el desarrollo. Finalmente, se aceptaron 22 trabajos que han sido publicados por Springer dentro de la serie de libros *IFIP Advances in Information and Communication Technology*

bajo el título *ICT for Promoting Human Development and Protecting the Environment*

Los trabajos seleccionados fueron clasificados en 7 categorías: medio ambiente y sostenibilidad, educación, salud, equidad digital, género, desarrollo económico e infraestructura y, gobierno y ciudades inteligentes.

Además de los mencionados trabajos, el programa se completó con conferencias, mesas redondas y demostraciones.

Las conferencias presentadas por Robert Atkinson, Mario Cimoli, Dave Hoelscher, Chrisanthi Avgerou, Erick Mata y, Malcolm Johnson añadieron un nivel de excelencia

al programa de WITFOR. Suscitaron gran interés en el público por la temática abordada, crearon controversia sobre el uso de las nuevas tecnologías como oportunidad para el desarrollo y, animaron a la reflexión sobre los riesgos que también conlleva.

WITFOR 2016 ha demostrado que esta conferencia puede jugar un importante papel para difundir la cooperación internacional, estimular la investigación, fomentar programas y aplicaciones específicos para el desarrollo, apoyar la educación como uno de los pilares de futuro, así como para intercambiar y diseminar información relacionada con todos los campos de la innovación tecnológica.



Promotores de WITFOR. De derecha a izquierda: Leon Strous (Presidente de IFIP), Eduard Dunder (Secretario de IFIP), Ana Pont, y Francisco Mata (TPC Chairs), y Gabriela Marín y Ramon Puigjaner (Vicepresidentes de IFIP y General Chairs).

Noticias del TC9: *ICT and Society*

Ignacio Gil Pechuán

Universidad Politécnica de Valencia; representante de ATI en el TC6

<igil@doe.upv.es>

IFIP colabora con un Bachillerato Internacional

El *International Baccalaureate* (IB) (Bachillerato Internacional) es un programa internacional de formación que los colegios pueden ofrecer (tras un proceso de autorización a partir del cual se les pasa a denominarse "colegios del mundo") y es para el mantenimiento de una línea de calidad formativa homogénea a nivel internacional.

En relación con las Tecnologías de la Información (TI) y a través de una colaboración con el Comité Técnico de IFIP 9 (TC9), centrado en Informática y Sociedad, se ofrece un programa TISG.

De este modo, la Tecnología de la Información en una Sociedad Global (TISG) formará parte de un Programa (IB) de dos años ofrecido a estudiantes de entre 16 y 19 años. ITGS es el único curso de TI preuniversitario centrado en los "individuos y sociedad" en lugar de en la propia tecnología.

ITGS abarca tres aspectos:

- Importancia social y ética: una discusión de los aspectos sociales y éticos relacionados con la evolución de las TI.
- Aplicación de las TI a situaciones específicas que utiliza el método de escenarios para abordar las evoluciones de los temas concretos.

- Sistemas de TI, que explora la terminología, conceptos y herramientas relativas a la evolución de las TI.

HCC12: 12th IFIP TC9 Human Choice and Computers Conference

Del 7 al 9 de septiembre de 2016 tuvo lugar la 12ª conferencia internacional HCC12 en Manchester, con el tema "*Technology and Intimacy: Choice or Coercion*", y cuyas aportaciones pueden obtenerse en el libro de Springer "*IFIP Advances in Information and Communication Technology*" Volume 474, 2016. <<http://link.springer.com/book/10.1007%2F978-3-319-44805-3>>.

Reunión anual del TC2 “Software: Theory and Practice”

Antonio Vallecillo Moreno

Universidad de Málaga; representante de ATI en el TC2

<antoniovallecillomoreno@gmail.com>

La reunión anual del *Technical Committee 2* de IFIP (TC2: “Software: Theory and Practice”, <<https://ifip-tc2.paluno.uni-due.de>> se celebró este año el 6 de julio de 2016, en Viena, Austria.

Siguiendo la tradición de los últimos años, se hizo coincidir junto a una conferencia internacional relacionada con los temas del comité, que en este caso fue STAF 2016 (*Software Technologies, Applications and Foundations*, <<http://staf2016.conf.tuwien.ac.at/>>).

El TC2 fue creado en 1962, siendo el encargado dentro de IFIP de todos los temas relativos a la programación de los sistemas de software. El objetivo del TC2 es “mejorar la calidad del software mediante el estudio de todos los aspectos relacionados con el proceso de desarrollo del software, de forma tanto teórica como práctica”.

El comité está organizado en 16 grupos de trabajo, y sus principales áreas de actuación son:

- Los modelos formales de los conceptos del software.
- Los lenguajes y técnicas de programación.
- Los modelos de almacenamiento y procesamiento de la información.
- Los entornos de programación.
- Las interfaces de usuario.
- La calidad del software.

Las reuniones anuales del comité se centran principalmente en el seguimiento de las actividades de los grupos de trabajo, la aprobación de las propuestas de incorporación de nuevos investigadores a dichos grupos, y en la preparación de los temas que han de ser discutidos en la asamblea general de IFIP que normalmente tiene lugar en octubre.

A la reunión de este año asistieron 15 representantes nacionales, y otros cinco delegaron su voto en alguno de los asistentes. La reunión comenzó con la conferencia inaugural que dio el Prof. Gul Agha, de la Universidad de Illinois en Urbana-Champaign, en la conferencia SEFM, uno de los congresos que forman parte de STAF <<http://staf2016.conf.tuwien.ac.at/sefm/keynotes/>>.

Dentro del capítulo de organización de los grupos de trabajo del TC2, se nombraron nuevos *Chairs* para los grupos de trabajo WG 2.7 (José Creissac Campos, Universidade do Minho), WG 2.9 (Jane Cleland-Huang, DePaul University) y WG 2.15 (Jean-Christophe Filliâtre, CNRS, LRI, Université Paris-Sud). Igualmente, se nombraron 25 nuevos miembros para los diversos grupos de trabajo del TC2 y se repasaron los informes de cada uno de los 16 grupos de trabajo que componen el comité. La estructura y composición de dichos grupos está disponible en: <<http://www.ifip.org/bulletin/bulltcs/memtc02.htm>>.

En cuanto a los premios, Kevin Ryan y Lambert Meertens recibieron el *Outstanding Service Award* en 2015.

Asimismo se informó que la conferencia principal de IFIP (*IFIP World Computer Congress*) tendrá lugar en Poznan (Polonia) en 2018, organizada por Jerzy Nawrocki, *vice-chair* del TC2.

Un punto especial de este año fue la discusión sobre el papel que pueden jugar las sociedades nacionales de informática en su relación con IFIP, y los beneficios mutuos que podrían obtenerse de posibles colaboraciones. A la reunión asistió el Presidente de la Sociedad Austriaca de Informática, que cuenta con más de 2.000 socios, muchos de ellos alumnos, y se expusieron diferentes líneas de trabajo. Dado el interés y relevancia del tema, se quedó en debatirlo con mayor profundidad a lo largo de este año.

La reunión concluyó con un debate sobre el lugar de la próxima reunión del TC2, que se anunciará más adelante.

Actividades de ATI

X Edición del Premio Novática: Entrega del premio al autor ganador



Imagen del momento de la entrega del diploma a Ramon López de Mántaras (izq.) por parte de Daniel Raya (derecha).

El pasado día 25 de octubre se celebró en los locales de ATI en Barcelona la ceremonia de entrega del X Premio Novática, correspondiente al año 2015, a **Ramon López de Mántaras Badía**, Director del Institut d'Investigació en Intel·ligència Artificial del CSIC (Consejo Superior de Investigaciones Científicas).

La entrega del diploma acreditativo del premio corrió a cargo de Daniel Raya Demidoff, presidente de ATI.

Cabe recordar que el artículo escrito por el autor galardonado fue publicado en la Sección “Tecnología” del número 234 de *Novática* (octubre-diciembre de 2015) con el título “*Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial*”.

Una vez más queremos dar nuestra más sincera enhorabuena a Ramon por tan magnífico artículo.



Co-funded by the
Erasmus+ Programme
of the European Union

El proyecto **AppSkill**, financiado por la Unión Europea a través del programa “*Erasmus+ Sector Skills Alliances*” (554271-EPP-1-2014-1-UK-EPPKA2-SSA), constituye una Alianza Sectorial TIC Europea, cuyo objetivo es la mejora del nivel y calidad de la formación, las habilidades y la empleabilidad en el sector del desarrollo de aplicaciones móviles.

La mayoría de los programadores de *apps* mantiene sus competencias de desarrollo de *software* actualizadas; pero no siempre disponen de conocimientos avanzados en diseño para hacer que sus *apps* resulten atractivas y fáciles de utilizar. Además, en muchos casos, carecen también de las necesarias habilidades de gestión y orientación al cliente. Ese es el motivo por el que miles de *apps* móviles fracasan cada año. Ello hace necesario adecuar la formación para que responda a las necesidades del mercado. Objetivo que se persigue a través de la eficaz colaboración de catorce socios -entre ellos la **Asociación de Técnicos de Informática, ATI**- procedentes de Reino Unido, Italia, Bélgica y España, y ocupados en diferentes áreas de actividad, desde autoridades reguladoras en materia de Educación, hasta representantes del Sector TIC, pasando por centros académicos de diferente tipo.

El proyecto, iniciado en enero de 2015, ha permitido conocer las principales competencias requeridas por los creadores de aplicaciones móviles.

A partir de una revisión preliminar de la situación, se obtuvo un mapa competencial basado en el estándar europeo e-CF (*e-Competence Framework*) y un análisis comparativo de la brecha existente entre las competencias reales de los creadores de *apps* móviles y las verdaderamente demandadas por el mercado. Posteriormente, **ATI** ha liderado un estudio cualitativo y cuantitativo, a gran escala, realizado para contrastar el mapa competencial obtenido en la fase inicial. Entre las conclusiones alcanzadas por el referido estudio cabe citar que el conocimiento técnico en materia de programación móvil, programación dirigida por eventos, interoperabilidad, pensamiento computacional o prototipado, prima sobre el conocimiento general de los sistemas operativos o de la arquitectura *hardware* de los dispositivos móviles. De igual modo, destaca la importancia de aquellas habilidades relacionadas con el diseño técnico de las plataformas, con el desarrollo de aplicaciones y con el soporte a las mismas, identificándose la creación de interfaces de usuario y la gestión de la seguridad como las más demandadas. En relación al despliegue de aplicaciones, resulta esencial o muy requerido saber cómo enviar *apps* a la tienda virtual de aplicaciones. Y, finalmente, otra importante conclusión es que las habilidades relacionadas con el negocio móvil y la gestión del mismo se consideran, también, esenciales o muy importantes, a pesar de no formar parte de la mayoría de los programas educativos existentes en la actualidad. (Habilidades relacionadas con el márketing, como promover o mejorar la visibilidad de las *apps* dentro de la tienda virtual -lo que en inglés se conoce como *App Store Optimization, ASO*- o comprender los mercados, fueron consideradas por la mayoría de los expertos consultados, como importantes o bastante importantes).

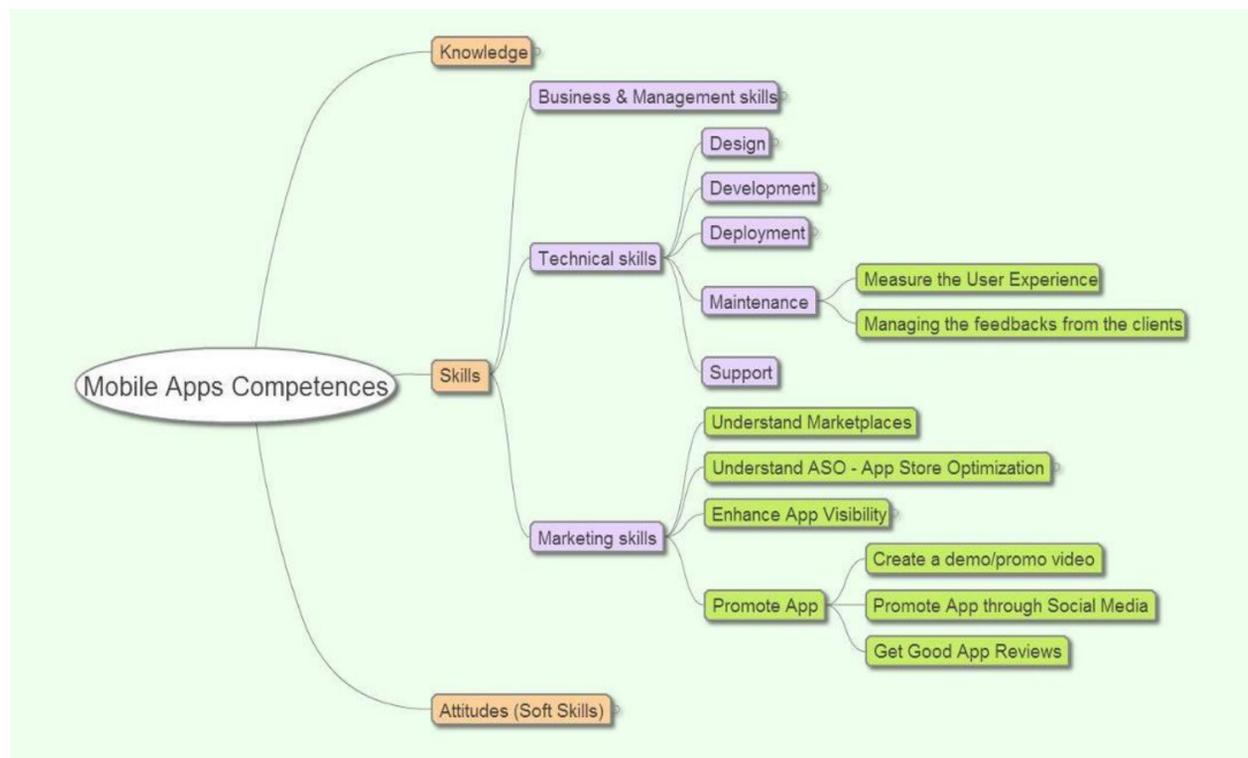
Todos estos resultados han permitido obtener el **mapa competencial** final y definir el **programa formativo del MOOC** para creadores de aplicaciones móviles en cuyos contenidos se está trabajando actualmente y que, con 60 horas de duración, se prevé esté disponible para el último trimestre de 2017 en una oferta multilingüe: inglés, francés, alemán, italiano y español.

Obtenga más información sobre el proyecto en: <http://appskil.eu/>.





Co-funded by the
Erasmus+ Programme
of the European Union



Mapa competencial para creadores de apps móviles

Module & learning objective	Learning outcome <i>On completion of the learning topic, the learner will be able to:</i>
1. Development of the App	Collect and report adequate information to decision makers
	Identify the required technical attributes and capabilities of the app, especially from an architecture and integration point of view
	Utilise Agile methodologies (Feature Driven Development – FDD -, compliance with Scrum, use of the Extreme Programming methodology - XP -) and associated items
	Conduct a <u>test driven</u> development (create and conduct a test plan, monitor and evaluate the test process, conduct tests, report and analyse results, prepare and conduct actions)
2. Marketing channels	Describe the use of the app in the context of social media
	Identify marketplaces to sell, promote and publish the app
	Use and interpret mobile app analytics (market analysis)
	Evaluate and monitor the customers review / feedback process
3. Project Management	Evaluate the time / money / quality triangle
	Predict and assess risks during development and deployment
	Monitor the change request process (analyse the impact if changes on users, predict all actions required to mitigate the impact of changes – training, documentation, new processes...)
	Monitor and report on project activities and progress against the project plan
	Monitor software quality assurance of products for usefulness and impact
	Monitor, assess and conduct actions upon quality indicators
	Deploy the app: organise and monitor trial days, submit app to the store.

Programa formativo del MOOC para creadores de aplicaciones móviles

Obtenga más información sobre el proyecto en: <http://appskil.eu/>.



José María Gómez Hidalgo¹,
Ricardo Baeza-Yates²

¹*Analytics Manager de Pragsis Bidoop, coordinador de la sección técnica "Acceso y recuperación de información" de Novática;*
²*Director de Tecnología (CTO) de NTENT, Catedrático part-time de la Universitat Pompeu Fabra y la Universidad de Chile*

<jmgomez@pragsis.com>, <rbaeza@acm.org>

Presentación. *Big Data:* Conceptos y aplicaciones

1. Introducción al Big Data

Big Data, como *Cloud*, *Machine Learning* y otras palabras clave de moda, aparece por todas partes hoy en día. Está en la prensa y en las noticias, en publicaciones de informática y de investigación, tiene sus propias conferencias, y no es raro que un amigo no informático nos pregunte, en una charla de café: "Oye, ¿qué es eso del *Big Data*?".

Existen dos visiones de lo que es el *Big Data*, una general y otra técnica. Por un lado, el público en general es consciente, a través de las noticias, de que las grandes compañías de Internet, y también los gobiernos, están acumulando gran cantidad de datos sobre nuestro comportamiento, con fines comerciales (la publicidad dirigida), o para reforzar la seguridad.

Para alguien de la calle, decir *Big Data* es decir "Gran Hermano". En parte al menos, los responsables de esta visión son los medios de comunicación, que dan una visión sesgada y alarmista de un conjunto de tecnologías que, a todos los efectos, se pueden considerar casi imprescindibles en la gestión de datos de hoy y mañana en cualquier tipo de organización o sector.

Por otro lado, *Big Data* hace referencia a volúmenes de datos tan grandes que han requerido el diseño de nuevos algoritmos y estrategias para manejarlos. Por asimilación, entendemos que también hace referencia al conjunto de tecnologías que, basadas en la idea de la distribución de los datos y los recursos para procesarlos, permiten una gestión del volumen, procesamiento y análisis de datos con una profundidad y escalabilidad sin precedentes, y virtualmente ilimitada.

Las tecnologías *Big Data*, especialmente a través del ecosistema Hadoop, cubren la adquisición, el almacenamiento, el acceso, el procesamiento (ya sea por lotes - *batch* o en tiempo real), la analítica y la visualización de inmensas cantidades de datos.

Una de las ideas más celebradas de *Big Data* es la de distribuir los datos y los programas de tal manera que los programas no tratan

Editores invitados

José María Gómez Hidalgo ha sido profesor e investigador en la Universidad Complutense de Madrid y la Universidad Europea de Madrid, durante 16 años, y Director de I+D en la empresa multinacional de seguridad de Optenet (ahora Allot Communications). Actualmente es *Analytics Manager* en la empresa de *Big Data* Pragsis Technologies, donde realiza consultoría de *Big Data Analytics* para empresas de banca y finanzas, turismo, industria, medios de comunicación, etc. A nivel de investigación, José María se centra principalmente en el Procesamiento del Lenguaje Natural y Aprendizaje Automático sobre datos textuales, con aplicaciones en el acceso a la información de actualidad y biomédica, y la Recuperación de Información con Adversario (filtrado de correo basura, filtrado Web y protección del menor en Internet). Es autor de numerosos trabajos de investigación en estas áreas, coordinador de la sección técnica de "Acceso y recuperación de información" de *Novática*, y fue coeditor de la monografía del número 185 de la revista, titulada "*Buscando en la Web del Futuro*". Asimismo, José María es miembro de SEPLN, ACM y ATI.

Ricardo Baeza-Yates es Director de Tecnología (*Chief Technology Officer*, CTO) de NTENT <<http://www.ntent.com>>, una compañía de tecnología de búsqueda semántica basada en New York y California, desde junio de 2016. Antes fue Vicepresidente de Investigación de Yahoo Labs, primero en Barcelona y luego en Sunnyvale, California, desde enero del 2006 hasta febrero del 2016. Entre 2008 y 2012 también supervisó Yahoo Labs Haifa y entre 2012 y 2015 estuvo a cargo de Yahoo Labs Londres. Hasta 2005 fue director del Centro de Investigación de la Web <<http://www.cwr.cl/>> en el Departamento de Ciencias de la Computación <<http://www.dcc.uchile.cl/>> de la Escuela de Ingeniería <<http://www.fcfm.uchile.cl/>> de la Universidad de Chile <<http://www.uchile.cl/>>; y catedrático ICREA y fundador del Grupo de Investigación de la Web <<http://wrg.upf.edu/>> en el Dept. de Tecnologías de la Información y las Comunicaciones <<http://www.upf.edu/dtecn/>> de la Universitat Pompeu Fabra <<http://www.upf.edu/>> en Barcelona (España). Mantiene vínculos con ambas universidades como catedrático jornada parcial. Obtuvo su doctorado en Ciencia de la Computación en la Universidad de Waterloo (Canadá) en 1989. Sus intereses de investigación incluyen algoritmos y estructuras de datos, recuperación de información, búsqueda y minería de datos en la Web además de *data science* y visualización de datos. Tiene más de 500 publicaciones donde destaca el libro *Modern Information Retrieval*, cuya segunda edición fue publicada en 2011 por Addison-Wesley. Ha obtenido varios premios, incluyendo distinciones de la Organización de Estados Americanos, el Centro Latinoamericano de Estudios en Informática, el Instituto de Ingenieros de Chile y la Universidad de Waterloo. Es *Fellow* de la ACM y de la IEEE.

de acceder a los datos, provocando un cuello de botella, sino que los datos se ubican de manera natural a lo largo de un *cluster* o grupo de ordenadores, en los lugares donde residen los programas que los precisan.

La distribución de datos se realiza de forma replicada, aumentando asimismo la fiabilidad y disminuyendo el riesgo de pérdida de información. El uso de múltiples servidores de relativas bajas prestaciones individuales en un *cluster* hace además que el factor de disminución del coste sea un beneficio estratégico de esta tecnología.

Éstas son sólo algunas de las ventajas que proporciona el *Big Data*. En esta monografía

pretendemos que se muestren éstas y muchas otras a través de los artículos que la componen.

2. Esquema de la monografía

Nuestra idea al preparar esta monografía, ha sido dar a sus lectores la oportunidad de obtener una visión global de cómo funcionan estas tecnologías, de cómo afectan a los enfoques tradicionales de análisis de datos, y cómo están contribuyendo a promover y a posibilitar la transformación digital de la empresa.

Para ello, contamos con una serie de artículos que tocan distintas áreas del *análisis de datos* (características de los datos, preprocesamiento, *streams* o flujos, arquitecturas), varias *aplicaciones* (sistemas de recomenda-

ción, análisis Web, redes sociales, analítica de texto), un caso de uso real y práctico que nos muestra una visión plenamente aplicada y sus beneficios de *negocio*, además de un artículo específicamente dirigido hacia la *privacidad*.

En lo que se refiere al *análisis de datos*, **Ricardo Baeza-Yates**, uno de los editores de esta monografía, saca partido de su gran experiencia en la búsqueda y la minería de datos en la Web, para exponernos en su artículo todos los problemas de los datos de la Web, como por ejemplo su calidad, escalabilidad, los sesgos, la dispersión y la privacidad de los mismos.

A continuación, **Salvador García, Sergio Ramírez-Gallego, Julián Luengo y Francisco Herrera** nos describen los requisitos de los algoritmos de preprocesamiento de datos (limpieza, normalización, transformación, etc.) en entornos *Big Data*. También nos muestran un caso de uso en la tarea concreta de selección de atributos, donde se describe el algoritmo Fast-mRMR, desarrollado por su grupo para garantizar la escalabilidad y la eficiencia en esta tarea.

En el mismo ámbito, **Albert Bifet y Jesse Read** realizan una discusión de los requisitos necesarios para procesar grandes cantidades de datos en tiempo real (*streams* o flujos), con especial atención a una de sus fuentes principales: la Internet de las Cosas (*Internet of Things*). En particular, nos explican cómo los algoritmos usados tradicionalmente para la resolución de la evolución de conceptos, la clasificación y el agrupamiento, se pueden modificar para operar en ámbitos *Big Data*. También nos introducen algunas herramientas de código abierto que se pueden usar para procesar grandes cantidades de datos en tiempo real.

Adicionalmente, **Alexey Cheptsov y Bastian Koller** presentan su trabajo sobre la evolu-

ción de los sistemas de supercomputación (*High Performance Computing*) para poder hospedar tecnologías *Big Data*, con la idea de traer lo mejor de ambos mundos: la rapidez de la supercomputación y la escalabilidad del *Big Data*. Dado que las plataformas de supercomputación no pueden soportar por defecto arquitecturas *Big Data*, los autores proponen, por ejemplo, un cambio en los sistemas de paso de mensajes que afecta a la arquitectura de estas plataformas, desarrollado en el proyecto europeo Juniper.

En lo que se refiere a las *aplicaciones*, **David C. Anastasiu, Evangelia Christakopoulou, Shaden Smith, Mohit Sharma y George Karypis** nos ofrecen un panorama de los avances más recientes en la adaptación de los sistemas de recomendación a entornos *Big Data*, con el fin de dotarles de la necesaria escalabilidad para realizar recomendaciones a millones de usuarios en cuestión de segundos. No debemos olvidar que la personalización, la recomendación y la publicidad son áreas críticas para el negocio hoy en día, ya sea digital o físico.

Por otra parte, **Weizhong Zhao, Gang Chen, Venkata Swamy-Martha y Xiaowei Xu** presentan un artículo muy interesante en el que se discuten los algoritmos de análisis de grafos, habitualmente empleados en el ámbito de la Minería de Redes Sociales. Debido a las deficiencias de los algoritmos tradicionales, los autores han tenido que extender un algoritmo para el análisis social a un entorno paralelo; este nuevo algoritmo, llamado PS-CAN, es una contribución importante para esta monografía.

Asimismo, el otro editor de esta monografía, **José María Gómez Hidalgo**, nos presenta las bondades que ofrece una de las tecnologías más extendidas en el ámbito *Big Data* (la plataforma Spark) para el análisis de información textual. En su artículo, se discuten

dos de los elementos que permiten construir de manera rápida sistemas de clasificación de texto, que son la extensión Spark Streaming y la biblioteca Spark MLlib. Adicionalmente, nos demuestra la tecnología por medio de un ejemplo de aplicación al análisis de sentimiento de texto.

Con el fin de demostrar el potencial de las tecnologías *Big Data* para transformar el *negocio*, contamos con el artículo proporcionado por **Xavier Ferrándiz Bofill y Alberto Alejo Marcos**, que nos detalla las motivaciones para realizar esta transformación en la Corporación Catalana de Medios Audiovisuales. Xavier y Alberto también nos explican qué casos de uso concretos han abordado hasta el momento, con qué objetivos y tecnologías, y cuáles son los resultados operativos que han obtenido.

Es importante subrayar que los beneficios obtenidos en esta iniciativa incluyen lecciones organizativas que les permiten avanzar en una transformación digital completa en el análisis y procesamiento de la información, con beneficios específicos para el negocio.

Para finalizar, y dada la preocupación en los medios y en los usuarios finales por los aspectos relacionados con la *privacidad*, hemos pedido a **José María del Álamo Ramiro, Esmeralda Saracibar Serradilla y Emilio Aced Féliz**, que contribuyan a esta monografía con un análisis de la privacidad en entornos *Big Data*.

En su artículo, nos ofrecen una panorámica general sobre los principales riesgos que la analítica masiva de datos entraña para la privacidad y la protección de datos personales, y también presentan varios enfoques jurídicos, organizativos y técnicos para sacar el máximo provecho de los datos mientras se respetan los derechos individuales.

Referencias útiles

A continuación se proporcionan algunas referencias sobre *Big Data*, acompañadas por artículos científicos recientes y enlaces con herramientas útiles para profundizar en los distintos enfoques sobre el tema.

Libros y revistas

■ **Bart Baesens.** “*Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*”. Wiley, 2014. ISBN: 978-1-118-89270-1. Un libro orientado a sacar partido de los datos y del *Big Data* en el marco empresarial.

■ **Viktor Mayer-Schönberger, Kenneth Cukier.** “*Big Data: A Revolution That Will Transform How We Live, Work, and Think*”. Houghton Mifflin Harcourt, 2013. ISBN-10: 1848547927. Este libro presenta una visión general de cómo estas tecnologías afectarán a la sociedad.

■ **Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills.** “*Advanced Analytics with Spark: Patterns for Learning from Data at Scale*”. O’Reilly Media, 2015. ISBN-10: 1491912766. En este libro se presenta cómo desarrollar aplicaciones de analítica de datos con Spark.

■ **Tom White.** “*Hadoop: The Definitive Guide, 4th Edition. Storage and Analysis at Internet Scale*”. O’Reilly, 2015. ISBN-10: 1491901632. Un libro que detalla la arquitectura y funcionamiento de Hadoop.

Artículos y estudios

■ **Richard M. Burton, Dolly Mas-trangelo, Fabrizio Salvador (editores).** “Big Data and Organization Design Special Issue”. *Journal of Organization Design, Vol. 3, No. 1, 2014*. Monografía sobre cómo *Big Data* afecta a la organización de las empresas.

■ **Jeffrey Dean, Sanjay Ghemawat.** “MapReduce: simplified data processing on large clusters”. *Communications of the ACM - 50th anniversary issue: 1958 - 2008, Volume 51 No 1*, enero 2008, pp. 107-113. Descripción del paradigma de programación funcional distribuida MapReduce.

■ **Sanjay Ghemawat, Howard Gobioff, Shun-Tak Leung.** “The Google File System”. *Proceedings of the nineteenth ACM Symposium on Operating Systems Principles*, ACM, New York, NY, EEUU, pp. 29-43, 2003. El artículo clásico en el que se presenta el sistema de archivos de Google, que inspiró Hadoop.

Eventos

■ **Apache Big Data Europe 2016.** 14-16 de noviembre, 2016, Sevilla, España. La conferencia de la Fundación Apache sobre *Big Data* en Europa. <<http://events.linuxfoundation.org/events/apache-big-data-europe>>.

■ **Big Data Spain 2016.** 17-18 de noviembre, 2016, Madrid, España. La principal conferencia técnica sobre *Big Data* en España. <<https://www.bigdataspain.org/>>.

■ **IEEE BigData 2016.** 5-8 de diciembre, 2016, Washington D.C., EEUU. Una de las más importantes conferencias científicas sobre *Big Data*. <<http://cci.drexel.edu/bigdata/bigdata2016/index.html>>.

■ **Data Beers Madrid y Barcelona.** Los dos *meetups* más importantes sobre *Data Analytics* y *Big Data* en España. <<http://databeers.tumblr.com/>>, <<http://databeersbcn.com/>>.

Distribuciones y software

■ **Apache Hadoop.** La página oficial del proyecto Hadoop y de muchos otros que conforman su ecosistema. <<http://hadoop.apache.org/>>.

■ **Cloudera.** Probablemente la distribución de Hadoop más popular. <<http://es.cloudera.com/>>.

■ **Hortonworks.** Otra distribución muy popular de Hadoop. Hortonworks contribuye con frecuencia a proyectos del ecosistema Hadoop. <<http://es.hortonworks.com>>.

■ **MapR.** Otra de las distribuciones de Hadoop más extendidas. <<https://www.mapr.com/>>.

■ **SoIR.** La herramienta de búsqueda del ecosistema Hadoop. <<https://lucene.apache.org/solr/>>.

Proyectos y organizaciones a seguir

■ **Indra.** Es miembro fundador de la principal iniciativa de investigación europea en *Big Data*:

<<http://www.indracompany.com/es/noticia/indra-miembro-fundador-principal-iniciativa-investigacion-europea-big-data>>. INDRA ha colaborado en la puesta en marcha de una Asociación Público-Privada (APP) que marcará la estrategia de I+D+i sobre macrodatos.

■ **Pragsis Bidoop.** Participa en un proyecto europeo de seguridad: <http://pragsis.com/blog/pragsis_bidoop_participa_en_un_proyecto_europeo_de_seguridad>. Pragsis Bidoop participa en el proyecto DANTE que tiene como fin la detección de actividades terroristas en el ámbito digital usando tecnologías *Big Data*.

■ **Banco Santander.** *El Big Data aportará 2.500 millones en ingresos al Banco Santander*: <<http://www.muycomputerpro.com/2016/03/03/el-big-data-aportara-2-500-millones-en-ingresos-al-banca-santander>>. Una noticia que indica cuáles son los planes de futuro del Grupo Santander para su transformación digital.

■ **BBVA.** Proyecto *Big Data* del Centro de Innovación BBVA: <<http://www.centrodeinnovacionbbva.com/proyectos/big-data>>. Toda la información y conocimiento que se está generando en torno a este tema en BBVA.

■ **Unión Europea.** Proyecto Juniper. Un proyecto del Séptimo Programa Marco de la Unión Europea, para el desarrollo de una arquitectura que permita el procesamiento de grandes volúmenes de datos en tiempo real. <<http://www.juniper-project.org/>>.

■ **Telefónica, Universitat de Barcelona.** Telefónica y la *Universitat de Barcelona* aplican por primera vez *Big Data* al mundo de la gastronomía con Appetit. Una iniciativa para aplicar las tecnologías del *Big Data* al mundo de la gastronomía. <<https://www.telefonica.com/es/web/sala-de-prensa/-/telefonica-y-la-universitat-de-barcelona-aplican-por-primera-vez-big-data-al-mundo-de-la-gastronomia-con-appetit>>.

Privacidad en la analítica masiva de datos

José María del Álamo Ramiro¹, Esmeralda Saracibar Serradilla², Emilio Aced Féliz³

¹Departamento de Ingeniería de Sistemas Telemáticos, Escuela Técnica Superior de Ingeniería de Telecomunicación, Universidad Politécnica de Madrid; ²ECIX Group; ³Unidad de Evaluación y Estudios Tecnológicos, Agencia Española de Protección de Datos

<jm.delalamo@upm.es>, <esmeralda.saracibar@ecixgroup.com>

1. Introducción

La analítica masiva de datos, o *Big Data*, es el nombre que se le da al conjunto de tecnologías, algoritmos y sistemas empleados para recolectar y almacenar datos y extraer información de valor de ellos mediante sistemas analíticos avanzados, todo ello llevado a cabo sobre ingentes y variados volúmenes de datos [1]. Es fácil ver que la capacidad de tratar grandes colecciones de datos (muchos de ellos personales), y desarrollar inferencias y detectar correlaciones lleva aparejadas enormes posibilidades de desarrollo y a la vez retos importantes para la privacidad y la protección de datos personales a los que hay que hacer frente.

En efecto, la analítica masiva de datos puede producir importantes beneficios sociales en muchos campos como la investigación científica y médica, la mejora de la asistencia sanitaria, la detección del fraude, o la mejora en la asignación de recursos por parte de organismos públicos, entre otros. Y, por supuesto, beneficios económicos a las organizaciones que la empleen para conocer mejor a sus clientes y usuarios, y establecer sus estrategias comerciales de acuerdo con ese conocimiento.

Por ejemplo, en el sector de la distribución permite obtener conocimiento para anticiparse al consumidor evitando situaciones de desabastecimiento de productos y falta de suministro; en el sector sanitario permite reducir el tiempo de ingreso hospitalario o predecir futuras enfermedades y riesgos sanitarios; la investigación médica se ha beneficiado del desarrollo de herramientas que permiten el procesamiento de la enorme cantidad de datos generados en relación con el genoma humano; y, se prevé un uso extensivo de estas tecnologías para la gestión y prevención de las incidencias de tráfico y excesos de contaminación en las ciudades.

Por otro lado, también surgen dudas y preocupaciones sobre posibles usos que, o bien no sean lícitos por realizarse sin respaldo legal para ello, o bien generen incertidumbre y desconfianza en la sociedad, derivados del hecho de que estas tecnologías no sólo

En memoria de nuestro compañero Emilio Aced Féliz que falleció mientras este artículo estaba en proceso de publicación.

Resumen: La analítica masiva de datos promete revelar relaciones entre datos hasta ahora ocultas, predecir tendencias generales y ofrecer nuevo conocimiento inferido mediante la recolección de enormes y complejos conjuntos de datos y la aplicación de novedosos algoritmos de análisis. Este avance tecnológico trae consigo grandes posibilidades para organizaciones de distintos tipos y tamaños, que ven en él una posibilidad para mejorar sus procesos, resultando tanto en beneficios sociales como económicos. Sin embargo, los avances tecnológicos normalmente llevan asociados nuevos desafíos para las sociedades que los adoptan, y la analítica masiva de datos plantea notables riesgos para la privacidad de las personas por su capacidad para revelar información que se creía privada. En este artículo se ofrece una panorámica general sobre los principales riesgos que la analítica masiva de datos entraña para la privacidad y la protección de datos personales, y se ofrecen distintos enfoques jurídicos, organizativos y técnicos que las organizaciones pueden adoptar para sacar el máximo provecho de los datos de que disponen a la vez que respetan los derechos individuales.

Palabras clave: Análisis de datos, Big Data, GDPR, inferencia, privacidad, procesamiento, protección de datos personales, recolección, regulación.

Autores

José María del Álamo Ramiro es doctor e ingeniero de telecomunicación, y actualmente es profesor en la Universidad Politécnica de Madrid (UPM) donde enseña distintos aspectos de la ingeniería de sistemas y servicios telemáticos a nivel de grado, máster y doctorado. Investiga en gestión de la identidad, privacidad y confianza desde 2005, colaborando con empresas en sectores como las telecomunicaciones (Teléfono, Ericsson), banca y servicios financieros (Santander), redes sociales (Vodafone, Twitter) o transporte y logística (Transfesa – Deutsche Bahn AG), participando en consorcios y proyectos nacionales e internacionales, y representando a la UPM ante organismos de estandarización como OASIS o W3C. Organiza el *IEEE International Workshop on Privacy Engineering* en San José, California, y es autor de más de 40 publicaciones y 5 patentes internacionales en la temática.

Esmeralda Saracibar Serradilla es socia del área de *Governance, Risk & Compliance* de Ecix. En su trayectoria profesional cuenta con un extenso recorrido tras 15 años dedicada a la asesoría jurídica de empresas en asuntos de Derecho y nuevas tecnologías, privacidad de datos, *Compliance*, *Big Data*, *Cloud Computing*, *Corporate Compliance*, Comercio Electrónico, etc. Es licenciada en Derecho por la Universidad de Deusto (2001), Máster Derecho de las Telecomunicaciones ICAI-ICADE (2002) y Máster en Dirección de Seguridad Global de la Universidad Europea Madrid (2004). Ostenta el certificado de ISMS Forum como analista de riesgos de Seguridad de la Información (2009) y el CDPP *Certified Data Privacy Professional* del Data Protection Institute (2011). Forma parte del capítulo español del CSA *Cloud Security Alliance* y del Comité Operativo del DPI y participa en el Grupo de Trabajo del DPI para la elaboración de una guía de buenas prácticas ante iniciativas de *Big Data*. Participa como ponente en cursos especializados en instituciones de reconocido prestigio como el Instituto de Empresa en el que imparte el Programa Especializado de *Compliance*. Ha sido ponente en el Master Ejecutivo de Dirección de Seguridad Global de la Universidad Europea de Madrid (2003-2006) y en el curso especializado en Seguridad de la Información del Ilustre Colegio de Abogados de Madrid). Ha publicado como co-autora las obras: *Compliance. Cumplimiento normativo y seguridad en la empresa y Protección de datos personales* (Thomson - Aranzadi). Es colaboradora habitual en revistas y prensa especializada: Diario Jurídico, El Derecho, Legal Today, El Economista, Economist & Jurist, etc.

Emilio Aced Féliz fue Jefe de Área en la Agencia Española de Protección de Datos (AEPD). Durante los últimos años desempeñó su actividad en el ámbito de la protección de datos personales tanto en la AEPD como Subdirector General de Inspección y Adjunto al Director y encargado de las relaciones internacionales de la AEPD; como en la Agencia de Protección de Datos de la Comunidad de Madrid en la que fue Subdirector General de Inspección y Tutela de Derechos y Subdirector General de Registro de Ficheros y Consultoría. Representó a la AEPD en el Grupo de Trabajo de Autoridades Europeas de Protección de Datos (Grupo de Trabajo del Artículo 29) y el Comité Consultivo del Convenio 108 del Consejo de Europa y fue miembro de las Autoridades Comunes de Control de Schengen, Sistema de Información Aduanero y Europol, de la cual fue Vicepresidente y Presidente.

“ Las iniciativas de *Big Data* conllevan la adopción de soluciones de cumplimiento en materia de protección de datos, desde un punto de vista jurídico, organizativo y técnico, con el fin de llevar a cabo un adecuado respeto de los derechos de los individuos y una idónea gestión de riesgos de incumplimiento por parte de las organizaciones ”

son capaces de describir el pasado y predecir el futuro con cierto nivel de incertidumbre, sino incluso de condicionarlo con las decisiones adoptadas [2].

Hoy en día son conocidos casos desafortunados donde el uso de técnicas de *Big Data* ha resultado en la violación de los derechos de privacidad y protección de datos de individuos.

Por ejemplo, el empleo de herramientas de analítica de datos para la identificación automática de clientes a término de su periodo de gestación y el posterior envío de material publicitario relacionado desencadenó una oleada de críticas al desvelar el embarazo de una menor de edad de forma imprevista por la afectada [3].

También han sido muy criticados los experimentos de condicionamiento inducido a usuarios de redes sociales a partir de la recolección de ingentes cantidades de datos generados por ellos y sus contactos, su análisis para la identificación de estados de ánimo, y su publicación de forma selectiva para condicionar estados de ánimo particulares en los receptores [4].

Este artículo aborda la problemática, y presenta algunos enfoques de solución, para mantener el equilibrio entre un uso legítimo y adecuado de las tecnologías de *Big Data* por parte de las organizaciones y el respeto a la privacidad y la protección de datos de los individuos.

Para ello, la siguiente sección presenta los riesgos más destacados asociados al análisis masivo de datos y, a continuación, se introducen distintos enfoques de solución, considerando aspectos legales, organizativos y técnicos. Por último, y dada la necesidad de abordar soluciones integrales ante un problema tan complejo, se ofrecen referencias para profundizar en aspectos particulares.

2. Principales riesgos para la privacidad

Los proyectos de *Big Data* siempre se concretarán en escenarios complejos desde el punto de vista de su interrelación con el

derecho fundamental a la protección de datos personales. Se trata de procesamientos masivos que afectan a millones de personas, para finalidades que puede que no estuvieran previstas cuando se recogieron los datos y cuyos resultados, en muchos casos y si no se utilizan adecuadamente, pueden tener efectos muy adversos sobre la vida de las personas.

De hecho, el nuevo Reglamento 2016/679 (UE), General de Protección de Datos (RGPD) [5]¹, aunque no habla en ningún momento de *Big Data* como tal, sí refleja la preocupación del legislador por dos de los resultados habituales de los tratamientos de *Big Data*: la utilización de sus resultados para la toma de decisiones automatizadas sobre las personas y la utilización de sus resultados para la elaboración de perfiles, estableciendo previsiones sobre la necesidad de informar al ciudadano y entregar esta información cuando se ejerce el derecho de acceso. Por lo tanto, existen una serie de riesgos genéricos que los tratamientos de *Big Data* tienen para la protección de datos (independientemente de los específicos de cada proyecto concreto).

La falta de transparencia u opacidad de los mismos es uno de los más importantes. En efecto, los tratamientos de *Big Data* se llevan a cabo en la mayoría de las ocasiones con los datos existentes en las organizaciones (complementados con datos de fuentes externas en muchos casos) sin que los afectados tengan ningún conocimiento de ello y no puedan expresar, como mínimo, su oposición a la utilización de los mismos.

El otro riesgo importante, conectado de alguna manera con el anterior, es la utilización de los datos con fines incompatibles respecto de los que se indicaron cuando se recogieron los datos y sin una legitimación adecuada para hacerlo.

Otra consideración muy importante es la calidad de los datos que se usan en los análisis. En muchas ocasiones no existe ninguna garantía de que los mismos tengan una mínima calidad y, por tanto, las predicciones que se realicen pueden ser completamente

erróneas, que es otro de los riesgos: el creer a pies juntillas que los resultados son algo más que correlaciones estadísticas entre variables que pueden no reflejar la realidad.

Finalmente, hay que mencionar el potencial discriminatorio que tiene el *Big Data* y el riesgo de que se nos encasille en diversos perfiles que tengan efectos muy relevantes sobre nuestra vida en el acceso a servicios como el crédito, los seguros o la salud.

Además, también es necesario tener en cuenta la temporalidad de las predicciones. Una predicción o una inferencia que se ha hecho hoy no necesariamente sigue siendo válida un año después, entre otras cosas porque la utilización de la predicción altera la realidad y puede transformarla en un plazo no demasiado grande de tiempo.

Teniendo en cuenta estos riesgos, las organizaciones pueden adoptar una serie de medidas legales, organizativas y técnicas que contribuyan a mitigarlos o evitarlos completamente, y que se introducen en el siguiente apartado.

3. Enfoques de solución

Las iniciativas de *Big Data*, si bien aportan valor a las organizaciones, conllevan la adopción de soluciones de cumplimiento en materia de protección de datos, desde un punto de vista jurídico, organizativo y técnico, con el fin de llevar a cabo un adecuado respeto de los derechos de los individuos y una idónea gestión de riesgos de incumplimiento por parte de las organizaciones.

En general, en primer lugar, es necesario valorar si la iniciativa requiere necesariamente el tratamiento de datos personales, es decir, asociados a una persona concreta, identificable o que pudiera ser identificada. En el caso de que sea necesario o se tenga previsto utilizar datos personales, habrá de tenerse en cuenta la aplicación de la normativa de protección de datos.

Habida cuenta del contexto en el que nos encontramos, habrán de observarse tanto la normativa local actual (Ley Orgánica de Protección de Datos –LOPD, y el Regla-

“ La disociación es el proceso por el cual se hacen anónimos distintos datos personales, es decir, se generan datos que no permiten la identificación del afectado o interesado con el que estaban relacionados ”

mento que la desarrolla - RLOPD) como la reciente normativa europea (RGPD), aplicable a partir de mayo de 2018.

Por el contrario, si la iniciativa no requiere el uso de datos personales, se reducirán de forma considerable los riesgos asociados. Por ello, a continuación describimos por separado cada uno de estos casos.

3.1. Análisis que no requiere datos personales

En los procesos de análisis masivos de datos que no requieren el uso de datos personales se puede establecer un proceso preliminar denominado de disociación o anonimización.

La disociación es el proceso por el cual se hacen anónimos distintos datos personales, es decir, se generan datos que no permiten la identificación del afectado o interesado con el que estaban relacionados. La disociación es un proceso relevante para la gestión de la privacidad en los procesos de *Big Data* ya que permite anonimizar conjuntos de datos de forma que éstos dejen de contener información considerada personal, y pueda ser tratada para generar conocimiento para el negocio sin las restricciones asociadas a los datos personales. Dicho de otra forma, al disociar la información que una organización maneja se reducen los riesgos de infringir la normativa de protección de datos, puesto que la organización reduce la cantidad de información personal que debe gestionar.

Existen distintas técnicas que permiten llevar a cabo la disociación de un conjunto de datos, que básicamente se pueden categorizar en técnicas de supresión, generalización y aleatorización.

Las primeras son las más burdas, y se centran en suprimir los registros más comprometidos. Las técnicas de generalización buscan diluir valores particulares de un dato de forma que ese valor sea compartido por un número suficiente de afectados para evitar que se pueda singularizar a un afectado (y el resto de sus datos asociados) conociendo exclusivamente ese valor.

Por ejemplo, en lugar de usar cinco dígitos para representar el código postal se pueden utilizar los cuatro más significativos, tres o incluso sólo dos, con lo que el número de afectados que comparten el mismo valor

para este atributo crece, reduciendo el riesgo de identificar a una persona en particular al conocer su código postal. Por su parte, las técnicas de aleatorización modifican el valor del dato para un individuo pero buscando conservar las propiedades observables para el conjunto de datos.

Como cabe esperar cada técnica presenta sus ventajas e inconvenientes. Primero, su aplicación lleva asociada una pérdida de información distinta, lo que finalmente reduce la utilidad, y por lo tanto el valor, del conjunto de datos resultante. Por otro lado, los resultados ofrecen distintas garantías frente a la re-identificación, por ejemplo, al combinarse con otros conjuntos de datos.

Es importante destacar que para garantizar la irreversibilidad se habrán de considerar tanto (1) las potenciales fuentes de información externas disponibles en los diferentes medios, especialmente en internet, como (2) la tecnología aplicable, no solo por parte del responsable del tratamiento sino por cualquier otra persona. De lo contrario, se corre el riesgo de que, como ha ocurrido en el pasado [6][7], conjuntos de datos considerados anónimos realmente no lo sean.

Las organizaciones deben evaluar los beneficios e inconvenientes de cada técnica y decantarse por aquella que en cada caso resulte óptima [GTA29].

Para llevar a cabo un correcto proceso de disociación las organizaciones deberían implantar una serie de medidas adicionales:

■ **Definición de una política de disociación:** resulta aconsejable definir una política de disociación que se encuentre documentada y actualizada, de manera que refleje de forma justificada las actuaciones a llevar a cabo para proteger la privacidad de los interesados y se encuentre accesible al personal implicado en el tratamiento de datos disociados, así como un protocolo de actuación del proceso de disociación.

■ **Uso de sellos de tiempo:** adicionalmente se habrá de tener en cuenta la posibilidad de utilizar en el proceso de anonimización algoritmos de sello de tiempo, con el fin de garantizar la fecha y hora en la que se realizó, o incluso algoritmos de firma electrónica para garantizar la identidad electrónica de quien ha realizado el proceso.

■ **Establecimiento de garantías jurídicas adicionales:** La política de disociación, el protocolo de actuación y las medidas tecnológicas adoptadas respecto de los procedimientos de anonimización habrán de reforzarse con las garantías jurídicas necesarias para preservar los derechos de los interesados, tales como (1) acuerdos de confidencialidad y cláusulas contractuales que garanticen la privacidad de la información incluso cuando haya brechas de re-identificación; (2) compromisos de mantenimiento de la anonimización de la información suscritos con los posibles destinatarios de la misma así como de no realizar ninguna acción para re-identificarla, o (3) auditorías de uso de la información anonimizada.

Por último, sería recomendable la realización de un proyecto piloto con una pequeña muestra de datos de prueba (no reales) del que puedan extraerse, de forma objetiva, conclusiones con respecto a la viabilidad de las técnicas de anonimización propuestas y del procedimiento de disociación.

3.2. Análisis que utiliza datos personales

Aquellas organizaciones que desarrollan análisis masivos de datos en los que intervienen datos personales deberán llevar a cabo procedimientos conducentes a mitigar o evitar los riesgos para la privacidad de los afectados o interesados, y con ello reducir sus propios riesgos de incumplimiento de la legislación vigente.

Para ello, como primera medida, la organización debería establecer una política de protección de datos desde el diseño y por defecto [9], para que desde las etapas iniciales de análisis de viabilidad de un proyecto se tenga en cuenta la necesidad de que el desarrollo respete los principios de protección de datos y lleve embebidos todos los requerimientos en esta materia.

De hecho, desde hace años, son muchas las voces que reclaman que cualquier proceso de desarrollo de servicios o sistemas debe tener en cuenta los aspectos de privacidad desde sus fases iniciales y a lo largo de todo el ciclo de vida [10]. Por ejemplo, el RGPD indica que “...el responsable del tratamiento debe [...] aplicar medidas que cumplan en particular los principios de protección de datos desde el diseño...”.

“ Como primera medida, la organización debería establecer una política de protección de datos desde el diseño y por defecto para las etapas iniciales de análisis de viabilidad de un proyecto ”

Si bien hay acuerdo en que las consideraciones de privacidad no deben realizarse una vez que se ha concluido un desarrollo, el proceso concreto a seguir para desarrollar un servicio o sistema que respete los principios básicos de privacidad no está tan claro.

En la actualidad, existen múltiples metodologías que ofrecen soluciones parciales para distintos dominios de negocio y/o contextos [11], varias técnicas que abordan problemas de privacidad particulares, e incluso herramientas de soporte. Aun así, la disciplina que permita identificar y abordar de forma sistemática y óptima los desafíos de privacidad a los que un ingeniero se enfrenta está todavía en desarrollo [12], aunque con resultados prometedores.

Por ejemplo, Apple ha anunciado recientemente la introducción de herramientas novedosas para conjugar sus necesidades de explotación de datos con el respeto a los derechos de privacidad y protección de datos [13].

La política de protección de datos debe ser vinculante y aplicable, con el fin de asegurar y evidenciar la verificación de cumplimiento de las exigencias legales y contemplando, como mínimo, los siguientes aspectos:

- **Limitación del tratamiento** de datos personales al mínimo imprescindible, y a las finalidades informadas y, en su caso, consentidas.
- **Licitud del tratamiento** en cuanto a que su origen sea legítimo, su tratamiento sea proporcional y no excesivo.
- **Transparencia**, en cuanto a la actitud del personal en la asunción del compromiso de actuar conforme a los objetivos de cumplimiento que previamente se hayan establecido.
- **Formación al personal** en el contenido de la misma.
- **Mecanismos de evaluación** de su fiabilidad y efectividad.
- **Auditorías** externas y/o internas de su cumplimiento.

De cara a observar algunos de estos aspectos se deberán introducir cláusulas de consentimiento informado al interesado, transparentes, expresas, precisas e inequívocas sobre las finalidades previstas, tales como, la elaboración de perfiles y modelos predictivos basados en la información del individuo, sus patrones de conducta, hábitos

y preferencias, así como, en un futuro, sobre los criterios lógicos, o algoritmos utilizados para la creación de su perfil y de las consecuencias que la creación de esos perfiles tendrán para los mismos y, a partir de la cual, se obtenga el consentimiento libre, específico, informado e inequívoco, de los afectados.

Además, deberán tenerse en cuenta aquellos supuestos en los que los datos se hayan obtenido de fuentes distintas del interesado, por ejemplo, fuentes de acceso público o de bases de datos comercializadas por terceros.

Especial relevancia adquieren los supuestos en los que las iniciativas de *Big Data* sean implantadas para decidir, por ejemplo, sobre el perfeccionamiento, finalización y/o renovación de la relación contractual suscrita, basándose únicamente en un tratamiento de datos destinados a evaluar determinados aspectos de su personalidad, a partir de valoraciones automatizadas, donde será necesario informar al cliente/usuario sobre la posibilidad de impugnar dichas valoraciones.

Otro aspecto muy importante, que se puede considerar como una herramienta para implantar las políticas de protección de datos desde el diseño, son las evaluaciones de impacto en la protección de datos (EIPD, o PIA —*Privacy Impact Assessment*— por sus siglas en inglés).

Una evaluación de impacto en la protección de datos personales es, básicamente, un ejercicio de análisis de riesgos para detectar los que puede entrañar el nuevo producto o servicio para las personas cuyos datos trata.

Concretamente el RGPD, establece la necesidad de realizar la EIPD siempre que se lleven a cabo elaboraciones de perfiles, en especial si sobre el resultado del tratamiento se basan decisiones que produzcan efectos jurídicos sobre el individuo, o pueden afectar de manera significativa a los individuos. Consecuencia de lo anterior, las organizaciones habrán de realizar una EIPD ante iniciativas de *Big Data* que involucren datos personales.

Para ayudar a las organizaciones a llevar a cabo estas evaluaciones de impacto, la Agencia Española de Protección de Datos (AEPD) publicó ya en 2014 una *Guía para una evaluación de impacto en la protección de datos* [14]².

La gran ventaja derivada de la realización de una EIPD es que permite identificar los posibles riesgos y corregirlos anticipadamente, antes de que el producto esté en el mercado y pueda lesionar los derechos de las personas y también tener un coste importante para la reputación de la organización.

En función de los resultados del análisis de riesgos (riesgos respecto de los derechos fundamentales de los interesados), habrá que implantar las medidas técnicas que se consideren idóneas para asegurar la privacidad. Entre las medidas técnicas que pueden contribuir a mitigar y corregir estos riesgos se encuentran las técnicas de disociación ya mencionadas en el apartado anterior, pero también otras como las técnicas de cifrado u otras para protección de la información, correspondientes al nivel de seguridad de la tipología de datos personales que sean objeto del tratamiento. En este sentido, al tratarse de la creación de perfiles basados en las características del individuo, serán de aplicación, como mínimo, las medidas de seguridad de nivel medio.

Las técnicas de cifrado buscan garantizar la confidencialidad de los datos personales, al impedir su acceso por quienes no dispongan de las claves de descifrado.

Estas técnicas pueden aplicarse cuando los datos están almacenados (por ejemplo, en un fichero o base de datos) y también mientras están en tránsito (por ejemplo, mediante uso de protocolos de comunicación seguros). Los sistemas comerciales de procesamiento masivo de datos suelen integrar este tipo de soluciones para minimizar el riesgo de revelación de datos, aunque de momento trabajan con los datos descifrados.

Para mitigar el riesgo de fuga de datos durante el procesamiento se están desarrollando soluciones que son capaces de realizar operaciones sobre datos cifrados (englobadas en el denominado cifrado homomórfico), lo que sin duda contribuirá a reducir más aún los riesgos.

Cabe mencionar que existen otras medidas técnicas que el responsable de la protección de datos debe considerar como un adecuado control de acceso a los datos, la trazabilidad de las operaciones realizadas, la implementación de procesos automáticos de monitorización y depuración de los datos,

“ Otro aspecto que deberían tener en cuenta todas aquellas organizaciones que llevan a cabo este tipo de tratamientos es la necesidad de contar con un equipo adecuado para ello ”

la definición de los periodos de retención y la aplicación automática de mecanismos de borrado una vez transcurridos éstos, etc.

Otro aspecto que deberían tener en cuenta todas aquellas organizaciones que llevan a cabo este tipo de tratamientos es la necesidad de contar con un equipo adecuado para ello.

En primer lugar, se debería nombrar un delegado de protección de datos para contar con asesoramiento de calidad sobre privacidad y protección de datos y para que supervise que todas las operaciones se llevan a cabo de manera respetuosa con la normativa. Este nombramiento será obligatorio con el RGPD cuando la actividad principal del negocio consista en operaciones de transformación que requieran de un seguimiento regular o sistemático de los interesados a gran escala.

Además, se requiere la presencia en el equipo de diseño y desarrollo de personas con conocimientos y experiencia en privacidad y protección de datos para asesorar y validar los resultados según se vaya avanzando en las fases del proyecto y en todo el ciclo de vida del mismo.

Por último, no hay que olvidar que los proyectos de análisis masivos de datos están sujetos a las mismas obligaciones que el resto de proyectos que tratan datos personales, y que no pueden dejarse de lado, tales como:

- Notificación del fichero a la AEPD cuya finalidad sea la implantación de iniciativas de *Big Data* o declarar esta finalidad en aquellos ya declarados que se encuentren afectados por dichos fines.
- Disponer de protocolos de actuación de ejercicios de derechos de acceso, rectificación, cancelación u oposición al tratamiento y/o de revocación del consentimiento para la gestión de las solicitudes de ejercicio de tales derechos derivadas de esta tipología de iniciativas. La posibilidad de revocación del consentimiento requiere de controles organizativos y técnicos para hacerlo efectivo.
- Cuando la iniciativa cuente con la participación de colaboradores externos o prestadores de servicios, se deberá regular contractualmente el acceso a los datos, con referencia a las medidas de seguridad que colaborador deberá observar.

Todas estas herramientas y metodologías que se han descrito tienen también la función de poner de manifiesto un elemento

crucial del nuevo RGPD: la responsabilidad activa de las organizaciones en la demostración del cumplimiento de la norma, implantando los elementos necesarios para demostrar que han puesto en marcha políticas adecuadas de protección de datos.

4. Conclusiones

Este artículo ha presentado los principales riesgos para la protección de datos personales derivados de procesos de análisis masivos de datos, presentando un conjunto de medidas que pueden ser útiles para las organizaciones que llevan a cabo estos procesos.

Estas medidas son tanto de tipo jurídico, como organizativo y también técnico, basadas en los principios de protección de datos desde el diseño y por defecto, y soportadas por una evaluación de impacto para la protección de datos.

Con el objeto de profundizar en estos temas y de ofrecer más detalle que puedan ayudar a las organizaciones en el respeto de los principios de protección de datos a la vez que desarrollan sus procesos de negocio con soporte de tecnologías de *Big Data*, el *Data Privacy Institute* de la Asociación Española para el Fomento de la seguridad de la información, ISMS Forum Spain <<http://www.ismsforum.es>> ha promovido y está próximo a publicar una guía de protección de datos para proyectos de *Big Data* en el cual se abordan con más detalle los aspectos que deben tenerse en cuenta ante iniciativas de esta tipología y que puede ser un documento útil y de interés para muchas empresas. Invitamos a todos los lectores interesados en la temática a consultar este documento.

Agradecimientos

Este artículo ha contado con la colaboración de la Asociación Española para el Fomento de la seguridad de la información, ISMS Forum Spain, a través de su iniciativa *Data Privacy Institute*, en la que profesionales de la seguridad y la protección de datos, representantes de empresas, universidades e instituciones, han trabajado conjuntamente para desarrollar la primera guía en el ámbito nacional de protección de datos para proyectos de *Big Data*.

Referencias

- [1] M.A. Beyer, D. Laney. *The importance of 'big data': a definition*, Gartner, pp. 2014-2018, 2012.

- [2] The White House. *Big Data and Privacy: a technological perspective*. Executive Office of the President, President's Council of Advisors on Science and Technology, 2014.

- [3] C. Duhigg. *How companies learn your secrets*. *The New York Times* 16, 2012. <<http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>>.

- [4] A.D. Kramer, J.E. Guillorey, J.T. Hancock. *Experimental evidence of massive-scale emotional contagion through social networks*, *Proceedings of the National Academy of Sciences* 111.24, pp. 8788-8790, 2014.

- [5] EUR-Lex. *Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos)*. *Document 32016R0679*. <<http://data.europa.eu/eli/reg/2016/679/oj>>.

- [6] A. Narayanan, V. Shmatikov. *How to Break Anonymity of the Netflix Prize Dataset*, The University of Texas at Austin, 2006.

- [7] L. Sweeny. *k-anonymity: A Model for Protecting Privacy*, *International Journal on Uncertainty, Fuzziness and Knowledge based Systems* 10, pp. 557-570, 2012.

- [8] Comisión Europea. *Grupo de Trabajo sobre Protección de Datos del Artículo 29, Dictamen 05/2014 sobre técnicas de anonimización*, 10 de abril de 2014. <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_es.pdf>.

- [9] G. D'Acquisto, J. Domingo-Ferrer, P. Kikiras, V. Torra, Y.A. de Montjoye, A. Bourka. *Privacy by design in big data: an overview of privacy enhancing technologies in the era of big data analytics*. ENISA, diciembre 2015. <<https://www.enisa.europa.eu/publications/big-data-protection>>.

- [10] 32nd International Conference of Data Protection and Privacy Commissioners. *Resolution on Privacy by Design*. Jerusalem, 2010.

- [11] N. Notario, A. Crespo, Y.S. Martín-García, J.M. Del Álamo, D. Le Métayer, T. Antignac, A. Kung, I. Kroener, D. Wright. *PRIPARE: Integrating Privacy Best Practices into a Privacy Engineering Methodology*, *Actas de IWPE15 en IEEE Security and Privacy Workshops (SPW)*, pp. 151-158, San Jose (CA), 21 May 2015.

- [12] S. Gürses, J.M. Del Álamo. *Privacy Engineering: Shaping an Emerging Field of Research and Practice*. *IEEE Security and Privacy*, vol. 14, no. 2, pp. 40-46, 2016.

- [13] K. Conger, N. Lomas. *What Apple's differential privacy means for your data and the future of machine learning*. *TechCrunch*, junio 2016. <<https://techcrunch.com/2016/06/14/differential-privacy/>>.

- [14] Agencia Española de Protección de Datos. *Guía para una evaluación de impacto en la protección de datos personales*, 2014.

Notas

¹ Toda la legislación de la Unión Europea está accesible en Eur Lex <www.eurlex.eu>.

² Se pueden consultar y descargar éste y otros documentos en el apartado *Publicaciones del Canal Resoluciones y Documentos* del sitio web de la AEPD <www.agpd.es>.

Dorian Peters

Creative Leader, Positive Computing Lab, Universidad de Sidney (Australia), coautora del libro "Positive Computing: Technology for Wellbeing and Human Potential"

<dorian.peters@sydney.edu.au>

¿Nos está haciendo felices la tecnología?

1. Introducción

Probablemente haya Vd. oído que el dinero no compra la felicidad. Más concretamente, los economistas no han encontrado incrementos significativos de nuestra felicidad, a nivel nacional, a pesar de la abundancia. A medida que se escala en la abundancia, las tecnologías digitales se han deslizado sigilosamente en cada rincón de nuestra existencia (desde la sala de juntas al dormitorio) presumiblemente para mejorar nuestras vidas. Por lo que la pregunta aún persiste: ¿Por qué no somos ciertamente más felices?

Por supuesto, si la tecnología no está guiando el camino para dar un gran salto y superar ese vacío psicológico, es porque no se ha diseñado para ello.

2. No diseñamos para el bienestar

Durante décadas, los ingenieros han realizado diseños para cosas sensibles como la productividad, la eficiencia o la certeza. Esas son fáciles de cuantificar y valorar en el lugar de trabajo donde surgieron los ordenadores. Pero los ordenadores han pasado de ser grandes mamuts del tamaño de oficinas a ser discretos y omnipresentes mediadores de las experiencias que nos modelan.

En respuesta a ello, algunos investigadores tecnológicos han comenzado a explorar cosas menos fáciles de cuantificar pero más críticas, como las emociones humanas. En realidad, muchos de los avances de nuestras vidas digitales han llegado de diseñadores que se están esforzando ahora por hacer tecnologías, no solamente funcionales, sino agradables, simpáticas e incluso significativas.

Aún son pocas las industrias tecnológicas que han llegado más allá de "mejorar el bienestar psicológico" en sus informes. Unas pocas han utilizado las ventajas del rico conocimiento generado por campos como la psicología positiva, para producir tecnología que de forma deliberada nos ayuda a crecer.

Antes de que se me pueda acusar de ridícula, me gustaría apuntar que la innovación tecnológica es mi pan y mi mantequilla. Como diseñadora digital estoy completamente de acuerdo en que la tecnología debería mejorar el bienestar mundial. No solo pienso que no ha sucedido aún, sino que creo firmemente que ahora es el momento perfecto para cambiar este hecho.

Traducción: Miguel Angel Jiménez Martín (socio sénior de ATI)

Resumen: Durante décadas, los ingenieros han realizado diseños para cosas sensibles como la productividad, la eficiencia o la certeza. Esas medidas funcionaban bien para el lugar de trabajo, pero fallan ahora que las tecnologías penetran cada vez más íntimamente en nuestras vidas, mediatizando cada experiencia que nos modela. En respuesta a ello, ha emergido un nuevo campo de trabajo llamado "Computación Positiva" (Positive Computing) para investigar cómo la psicología del bienestar puede ayudar al diseño de la tecnología para crear experiencias digitales que den mayor soporte a nuestro bienestar. En última instancia, la visión de la computación positiva es un futuro en el cual todas las tecnologías, desde los medios sociales al correo electrónico y los juegos, sean diseñados para ayudarnos a prosperar.

Palabras clave: Bienestar, computación positiva, experiencia de usuario, interacción persona-ordenador, psicología.

El investigador en Ingeniería del Software Rafael Calvo y yo escribimos el libro, *Computación positiva* para ayudar a catalizar este cambio. Entre otras cosas, llamamos a los tecnólogos, psicólogos y consumidores a esperar más de las tecnologías que impregnan nuestras vidas. Después de todo, ¿por qué no podría la tecnología dar soporte a nuestros mayores objetivos en la vida: ayudarnos en nuestra propia prosperidad, la de nuestras familias y la de nuestras sociedades?

3. ¿Cómo reestructurar nuestro futuro tecnológico?

El bienestar psicológico es un concepto amplio y puede ser difícil definirlo y concretarlo. Por lo tanto, recomendamos a los tecnólogos comenzar por referirse a los factores que contribuyen a ello. Estos "determinantes del bienestar" son cosas como la compasión, la plenitud mental, la gratitud, la autonomía, la competencia y la capacidad de relacionarse, que se ha comprobado que incrementan el bienestar psicológico.

Podemos ver a un número creciente de tecnologías diseñadas a dar soporte a algunos de estos determinantes (por ejemplo, programas y apps en línea sobre consciencia plena o gratitud), que señalan el comienzo de un largo camino. Pero ¿por qué parar en herramientas específicas?

4. Cada tecnología que usamos debería considerar su impacto en nuestro bienestar

El mayor cambio solo llegará cuando todas las tecnologías, desde las apps de medios sociales, al correo electrónico, o los juegos sean apropiados para el bienestar psicológico.

Imaginemos una red social rediseñada para fomentar la inteligencia emocional¹, un procesador de textos diseñado para dar soporte a flujos (algunas vistas de Microsoft Word podrían ser un paso en esa dirección), o videojuegos diseñados para promover auto reglas, conciencia o compasión². ¿Cómo crearíamos un futuro en el que todas nuestras tecnologías diarias den soporte óptimo a mentes más sanas?

En primer lugar, los tecnólogos tendrían que añadir medidas de validación de bienestar al ciclo de diseño (al igual que hacen los economistas para calificar la política nacional). Los tecnólogos necesitarían además (y esto es crítico) trabajar a medias con investigadores del bienestar como psicólogos y sociólogos, para asegurarse de que el proceso tiene una base científica y no pierde rigor.

Sería muy fácil elevar una tecnología como "positiva" porque es divertida y por lo tanto hace feliz a la gente, pero sabemos que el significado genuino de bienestar es mucho más complejo que eso. Por lo tanto, incluso las mejores intenciones, si no están bien formuladas, pueden fallar. Por ello, las categorizaciones deben ser validadas por investigadores y evaluadas por expertos en las ciencias del bienestar de que se trate.

5. El poder al usuario

Aunque la mayoría de nosotros no somos creadores de tecnología, somos todos consumidores de la misma, y como tales podemos ayudar a catalizar el camino pidiendo a nuestras compañías tecnológicas (grandes, pequeñas y start-ups) que comiencen a medir el impacto en el bienestar y diseñen para mejorarlo.

Por ejemplo, los creadores de tecnología podrían, como mínimo, tratar a esos compromisos con el bienestar como errores, y por lo tanto rediseñarlos.

A lo sumo, podríamos llegar a la siguiente generación un mundo en el que los humanos comiencen a ser psicológicamente mejores por el uso de la tecnología, no a pesar de ella.

¿Pero cómo podría ser este nuevo mundo? He aquí algunos ejemplos de lo que se ve en el horizonte.

6. Una ojeada al futuro de la computación positiva

■ Tecnología para la autoconciencia

En la era del “autoseguimiento” llega una ráfaga de nuevas oportunidades para la reflexión y el incremento de la autoconciencia. Por supuesto, el seguimiento no es incondicionalmente bueno para nosotros, por lo que el investigador del bienestar necesitará informarse de esos esfuerzos, aunque tienen gran potencial. Por ejemplo, imaginemos una generación de herramientas de autoconciencia informadas por un investigador en “autocompasión”³.

■ Tecnología para la conciencia plena

Los ejemplos más obvios de ejemplos de tecnologías de conciencia plena se encuentran entre la creciente lista de *apps* de meditación como *SmilingMinds*, *Breathe* y *Buddhify*. Estos, como muchos otros programas, están diseñados para ayudar a la gente a aprender y mantener una práctica de conciencia plena. Menos conocidos son herramientas más experimentales como *Tenacity*, un juego de conciencia plena en el trabajo⁴, *HAPIFork* (una versión o *fork* vibrante)⁵ o *Sonic Cradle* (una cavernícola inmersión en sonidos)⁶ cada una diseñada para facilitar la conciencia del momento actual. La idea de que las propias tecnologías que de forma notoria nos evaden del presente podrían ser usadas en sentido contrario, para traernos de vuelta, es ciertamente gratificante.

■ Tecnología para la gratitud

Desde los diarios de agradecimientos hasta las *apps* de notas y los *posts* en redes sociales, las oportunidades de dar gracias abundan. Aunque muchos sitios están cambiando el botón “Me gusta” por otras alternativas agradecimiento como “Aprecio esto” o “Gracias”. El diseño cuidadoso de las pequeñas cosas será parte importante en el largo camino hacia una experiencia digital más agradable.

■ Tecnología para la empatía

En términos de entender lo que otros piensan y sienten, es difícil competir con los esfuerzos de los juegos de rol *immersivos*, o los juegos y mundos virtuales. El juego del pacificador, por ejemplo, permite a los jugadores ponerse

en los dos lados del conflicto Israel/Palestina, mientras que *Migrant Trail*⁷ les permite experimentar los peligros que rodean a los inmigrantes y a los guardas fronterizos. Pero ¿por qué parar en la empatía hacia humanos?. En *Shelter*⁸ el jugador debe actuar como una tejió hembra cobijando a sus cachorros del frío, y en un mundo virtual en *Stanford*⁹ puedes ser transformado en una pieza de coral y verte a ti mismo desintegrarte a medida que el océano se vuelve más ácido. La experiencia empática ya no está limitada por el espacio, el tiempo, o siquiera las especies.

■ Compasión y altruismo

Dado que muchos videojuegos se han hecho famosos por la violencia, los investigadores han mostrado que son simplemente capaces de promocionar comportamientos pro-sociales.¹⁰ ¿Qué podría ser más excitante que un juego que indujera al altruismo? Un grupo investigador de *Stanford*¹¹ encontró que la gente que pasa mucho tiempo salvando vidas como un superhéroe en un mundo virtual era más propensa a ayudar a otros en el mundo real después de todo. En el juego *Half-the-Sky* los jugadores se hacen pasar por mujeres en varios países en desarrollo que deben encontrar formas de conseguir medicinas, libros o comida que necesitan sus comunidades. Más allá del aprendizaje de otras formas de vida, los jugadores tienen oportunidades de hacer donaciones a importantes proyectos de ayuda en el mundo real a medida que progresan en el juego.

■ Tecnología para la resiliencia

La tecnología que da soporte a la compasión dará también soporte a la resiliencia¹², porque según han mostrado los investigadores¹³, la compasión y la resiliencia tienen relación¹⁴. Pero además hay herramientas que abordan directamente la rebelión, incluyendo *SuperBetter*¹⁵, una *app* similar a un juego que soporta usuarios creando recursos físicos, sociales, mentales y emocionales. Otros ejemplos incluyen *apps* para familias militares, programas de resiliencia para el lugar de trabajo, y trabajos realizados con el *HopeLab*¹⁶.

En el Laboratorio de Computación Positiva (*Positive Computing Lab*) de la Universidad de Sydney medimos y diseñamos para el bienestar mediante determinantes como la autonomía, la competencia, y el significado de los contextos de enfermedades crónicas, el bienestar en el lugar de trabajo y la salud mental juvenil.

Las posibilidades son incontables y los primeros ejemplos ofrecen inspiración. Por supuesto, cada nuevo camino conlleva nuevos cambios y necesitaremos negociar e hilar fino en lo relativo a la privacidad, seguridad y diversos valores a lo largo del camino.

Nunca dijimos que fuera fácil, pero merece la pena. Si un futuro en el que la tecnología dé soporte a la resiliencia, la capacidad de estar conectado y la compasión suena muy ambicioso, recordemos que conseguimos llegar a la Luna. Y, cuando del bienestar del mundo se trata, me atrevo a decir que errar no es una opción.

Así es que sigamos adelante y *re-imagínemos* un mundo digital tal que en una década a partir de ahora podamos decir de manera aplastante y, sin duda, que nuestras experiencias digitales nos están ayudando a prosperar.

Notas

¹ Ver *Facebook's Compassion Research Day* para empezar: <<https://www.facebook.com/events/758398997575624/>>.

² Ver *Games for Change* para ejemplos sencillos: <<http://www.gamesforchange.org/>>.

³ <http://greatergood.berkeley.edu/gg_live/science_meaningful_life_videos/speaker/kristin_neff/the_three_components_of_self_compassion/>.

⁴ <<http://www.glsstudios.com/gls/games.php>>.

⁵ <<https://www.hapify.com/product/hapifork>>.

⁶ <<http://www.jayvidyarthi.com/tagged/human+computer+interaction>>.

⁷ <<http://www.gamesforchange.org/play/the-migrant-trail/>>.

⁸ <<http://mightanddelight.com/?page=shelter>>.

⁹ <<https://vhil.stanford.edu/>>.

¹⁰ <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2678173/>>.

¹¹ <<https://vhil.stanford.edu/>>.

¹² “Resiliencia: Capacidad de adaptación de un ser vivo frente a un agente perturbador o un estado o situación adversos”. Fuente: Diccionario de la Real Academia, www.rae.es.

¹³ <http://greatergood.berkeley.edu/article/item/how_to_transform_stress_courage_connection>.

¹⁴ Ver “Compassion v. Empathy: Designing for Resilience”. <http://greatergood.berkeley.edu/article/item/how_to_transform_stress_courage_connection>.

¹⁵ <<https://www.superbetter.com/>>.

¹⁶ <<http://www.hopelab.org/>>.

A continuación presentamos las habituales referencias que desde 1999 nos ofrecen los coordinadores de las Secciones Técnicas de nuestra revista.

Sección Técnica “Acceso y recuperación de información” (José María Gómez Hidalgo, Enrique Puertas Sanz)

Tema: Congreso: 40ª edición del congreso ACM SIGIR

SIGIR es el congreso más relevante a nivel internacional en el área de la Recuperación de Información. Son cinco días con charlas y presentaciones en los que se muestran resultados de investigación, demostraciones de nuevos sistemas, y técnicas y algoritmos relacionados con la recuperación de información.

Esta edición va a ser la número 40, y se celebrará del 7 al 11 de agosto de 2017 en Tokio, Japón. Actualmente están confirmadas charlas plenarias de *Yoelle Maarek*, VP de Investigación en Yahoo!, y *Stephen E. Robertson*, creador del modelo probabilístico de recuperación de información.

En la web del congreso está ya abierto el período de recepción de contribuciones de artículos científicos y demostraciones sobre temas como la búsqueda y análisis de *queries*, búsquedas sociales, representación de documentos, o sistemas de filtrado y recomendación, entre otros tópicos relacionados con la recuperación de información.

Más información en: <<http://sigir.org/sigir2017/>>.

Tema: Recurso: UIPath Studio

UIPath Studio es una *suite* de herramientas de construcción de robots lógicos para la extracción automática de datos de diversas fuentes. La herramienta permite construir sistemas de *web scraping*, para recuperar datos de la web o automatizar la extracción de datos de entrada y salida de programas de escritorio mediante la automatización de ejecuciones.

El software permite automatizar tareas repetitivas mediante la construcción de *scripts* de forma visual, simplificando y acelerando la tarea de extracción de datos. Para ello el sistema cuenta con un editor visual que permite definir flujos de ejecución de programas (o webs), e interacciones, con el fin de crear un robot de recuperación automática de datos.

UIPath Studio cuenta con una versión comercial, con soporte por parte de la compañía, y una versión “Community” gratuita, ideal para proyectos no demasiado grandes o para aprender a crear flujos de extracción de datos.

Más información en: <<https://www.uipath.com>>.

Tema: Recurso: Usando Cognonto para crear modelos word2vec para dominios específicos

Word2vec es una red neural de dos capas utilizada para minería de texto y recuperación de información, que aprende relaciones entre palabras dentro de un corpus de documentos.

El corpus de textos utilizado por word2vec para aprender las relaciones entre las palabras es lo que se conoce como corpus de entrenamiento. De cara a automatizar la tarea de creación de corpus específicos para diferentes dominios podemos usar la base de conocimientos de Cognonto para crear automáticamente corpus de

entrenamiento específicos de dominios de alta precisión que puedan ser utilizados por word2vec para generar modelos de relación de palabras (para un dominio concreto).

Más información en: <<http://fgiasson.com/blog/index.php/2016/09/28/using-cognonto-to-generate-domain-specific-word2vec-models/>>.

Sección Técnica “Derecho y Tecnologías” (Elena Davara Fernández de Marcos)

Tema: El Gobierno avisará por SMS o correo electrónico de las multas y notificaciones

Tal y como señala la Ley del Procedimiento Administrativo Común, el Gobierno debe avisar al ciudadano por medio de un SMS o correo electrónico a la hora de enviar notificaciones sobre cualquier acto administrativo, ya sea éste una multa, o un requerimiento cuando vaya a recibir una multa. Adicionalmente, también se recibirá en papel.

La Administración comenzará enviando un aviso al teléfono móvil o a la dirección de correo electrónico que tenga del ciudadano, informándole de que tiene a su disposición una notificación en la sede electrónica que haya habilitado para ello. Este trámite se llevará a cabo con independencia de que el ciudadano haya manifestado su preferencia por el aviso por medios telemáticos.

Igualmente, nos encontramos con nuevos sujetos obligados a realizar actuaciones con la Administración mediante medios electrónicos, que hasta ahora era obligatorio para empresas. En este sentido, se añade esta obligación para los profesionales que exijan una colegiación, para los funcionarios y para aquellas profesiones que representen a un interesado.

Más información en:

<http://economia.elpais.com/economia/2016/09/26/actualidad/1474882001_399918.html>.

Tema: Multa de 150.000 euros a Google por derecho al olvido

La Agencia Española de Protección de Datos (AEPD) ha impuesto una sanción de 150.000 euros a Google, debido a una infracción del principio del deber de secreto estipulado en el artículo 10 de la Ley Orgánica de Protección de Datos (LOPD)

Esta sanción es consecuencia directa del ejercicio del derecho al olvido por parte del afectado. El procedimiento seguido, al ejercitar un usuario o interesado su derecho al olvido en Google, mediante la página que tiene al efecto, era que Google enviaba una comunicación al medio censurado, informándole de que se había solicitado la eliminación del contenido en cuestión. De esta manera, se ponía en riesgo la identidad del afectado que reclamaba ese derecho al olvido.

Al recibir Google la denuncia presentada ante la AEPD y proceder a la contestación a la misma, Google se defendió reclamando el derecho de la web censurada de conocer que esta censura se había llevado a cabo, que se había eliminado la indexación.

Google ha indicado que está estudiando la resolución y que procederá a analizar las distintas vías de actuación que considere convenientes.

Más información en:

<<http://www.europapress.es/portaltic/internet/noticia-multa-150000-euros-google-poner-riesgo-identidad-demandante-derecho-olvido-20160919145222.html>>.

Tema: *La AEPD e INCIBE publican 18 fichas prácticas para concienciar de la importancia de proteger su información personal*

La Agencia Española de Protección de Datos ha publicado en su web una serie de fichas (18 en total) que ha realizado en colaboración con el Instituto Nacional de Ciberseguridad (INCIBE)

En cada una de estas fichas, se ofrecen consejos y recomendaciones prácticas sobre la seguridad de los dispositivos, contraseñas, copias de seguridad, protección de datos, *cookies* o redes sociales, entre otros.

Estas fichas van encaminadas a tratar de resolver o arrojar algo de luz sobre las preguntas más frecuentes ante las que se encuentra un usuario a la hora de mantener a salvo su privacidad y seguridad navegando por la web, aportando ideas y soluciones prácticas y accesibles a cualquier usuario medio.

Una vez que accedes a la página, se encuentra un listado con distintas preguntas enunciadas de manera práctica, con un enlace al contenido donde se da respuesta a la pregunta en forma de ficha práctica con ejemplos.

Las fichas se presentan en formato PDF descargable, muy visual y conteniendo lo básico, de forma que el usuario pueda, en un vistazo, tener una primera idea de las cuestiones más importantes a tener en cuenta

Más información en:

<http://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/Guias_online/PrivacidadSeguridad-ides-idphp.php>.

Tema: *La AEPD inicia actuaciones de investigación por la comunicación de datos entre Whatsapp y Facebook*

La Agencia Española de Protección de Datos (AEPD) ha iniciado una investigación de oficio a Whatsapp y a Facebook, debido al uso y cesiones de datos que se realizan entre ellas.

De esta forma, va a revisar, de manera prioritaria, qué tipo de cesiones o comunicaciones de datos se llevan a cabo, la finalidad de éstas, qué datos se comunican y qué uso se hace de los mismos por parte del cesionario (que en este caso sería, por ejemplo, Facebook)

Atenderá también a qué plazos se prevén en cuanto a conservación de los datos de los usuarios, y si se ofrece de manera conforme a la legislación vigente la posibilidad del ejercicio de los denominados derechos ARCO (derechos de acceso, rectificación, cancelación y oposición) a todos los usuarios. Una vez concluidas las investigaciones pertinentes, la AEPD procederá a decidir qué acciones toma.

La AEPD está coordinada con las diferentes autoridades de control de privacidad europeas, como pueden ser la alemana o italiana, que llevan la misma línea, si bien cada una actúa conforme a las potestades y obligaciones que el ordenamiento interno de cada país les otorga.

Más información en:

<http://www.agpd.es/portalwebAGPD/revista_prensa/revista_prensa/2016/notas_prensa/news/2016_10_05-ides-idphp.php>.

Sección Técnica “Entorno Digital Personal”
(Diego Gachet Páez, Andrés Marín López)

Tema: *Nuevas tecnologías de comunicación para el Internet de las Cosas*

Existen en la actualidad algunas alternativas de conectividad para dispositivos que forman parte del Internet de las Cosas. De este modo, son muy conocidas las soluciones basadas en redes tipo zigbee o bien wifi.

Últimamente han aparecido en el mercado dos soluciones de conectividad como son Sigfox y LoRa, la primera, especialmente diseñada para *comunicar objetos conectados a larga distancia*, con un coste muy ajustado y reduciendo enormemente el consumo energético de los dispositivos, permitiendo a éstos transmitir sus datos a mucha distancia sin depender de otras redes o conexiones. Se trata de una solución propietaria.

Por otra parte, en el último año y en varios foros dedicados al Internet de las Cosas, se está hablando mucho de LoRa, una *red de comunicaciones basada en el protocolo LoRaWAN (Long Range Wide-area network)*, y que al igual que Sigfox, es una red de *baja tasa de transferencia, bajo coste de despliegue* y que ofrece comunicación bidireccional y está enfocada a “cosas” alimentadas por baterías.

Esta red está soportada por la *LoRa Alliance*, una asociación sin ánimo de lucro integrada por diversas compañías que colaboran en el desarrollo del protocolo común.

LoRaWAN es una buena opción con un alcance y un consumo energético excelentes, además de ofrecer acceso libre al código fuente. La red LoRa garantiza las comunicaciones a gran distancia a través de servidores, sin necesidad de usar otras redes como Wifi o conexiones a Internet.

Desde un punto de vista más técnico, LoRa se basa en la utilización de espectro ensanchado de frecuencia modulada pulsada “*chirping*”, la cual utiliza pulsos de frecuencia lineal de banda ancha modulados para codificar la información. La arquitectura de red inicial de largo alcance tiene un presupuesto de enlace de 5-10 dB adicionales a los sistemas celulares e, incluso, imita la arquitectura celular existente. Esto permite el uso de las estaciones base celulares elevadas existentes para ser equipadas con puertas de enlace propias.

Actualmente, existen redes LoRa abiertas al usuario final en Francia, Bélgica, Suiza, Países Bajos y Sudáfrica. Esas redes LoRa para el Internet de las Cosas son desplegadas por grandes operadores de telecomunicaciones, como Orange o Bouygues Telecom en el caso de Francia, sin que hasta ahora nos conste que exista actualmente una oferta en España.

Sección Técnica “Gobierno corporativo de las TI”
(Manuel Palao García-Suelto, Miguel García Menéndez)

Tema: *¡Abra los ojos!*

Nada hacía presagiar que aquel lunes fuese a ser muy diferente a los de las semanas anteriores. Los camiones habían descargado en los muelles de entrada durante la tarde del domingo, como era costumbre, y los operarios del turno de noche habían llenado los silos que alimentaban las mezcladoras, en las últimas horas, antes del alba.

¡Nada hacía presagiar que aquel iba a ser el último lunes! Sin embargo, a las 08:07 de la mañana, apenas dos horas después del comienzo del primer turno en el taller y cuando casi todo el personal de las oficinas aún estaba colgando sus abrigos en las perchas de sus despachos, comenzaron a producirse las primeras anomalías, que, inicialmente, estuvieron a punto de pasar inadvertidas.

Algunas impresoras dejaron de imprimir, tanto en las oficinas, como en la planta; y, en ésta última, las consolas de operación de las secciones de hornos y de envasado comenzaron a bloquearse intermi-

tentemente, haciendo caso omiso de las consignas que introducían los operadores. El problema no duró más de media hora y a eso de las 09:00 horas, desde el Departamento de Informática se comunicó que el incidente ya estaba resuelto, con lo que el resto de la mañana continuó transcurriendo en la más absoluta (y aparente) normalidad.

Durante el turno de tarde, las cosas se pusieron bastante peor. A penas quedaba, ya, personal en las oficinas; pero la zona de taller mostraba una actividad frenética. Bajo la supervisión del director de producción, que aún tenía estacionado su BMW en el parking "A", situado frente al edificio de recepción, los operarios estaban tratando de recuperar el tiempo perdido en el turno anterior. De repente, a las 17:23, la línea de enfriamiento dejó de funcionar, lo que provocó que la máquina de empaquetado recibiera un producto semiterminado, imposible de envasar.

Las consecuencias del incidente de la enfriadora y, más aún, las del deterioro de la empaquetadora, provocado por el vertido del producto licuado, habrían de verse en los próximos días, con una inesperada parada de la producción, un elevado porcentaje del personal ocioso a la espera de la reparación de los equipos y la maquinaria, y un canal de distribución desatendido, que dejaría un atractivo espacio de juego a la competencia.

La investigación, iniciada aquella misma noche, permitió señalar rápidamente a un posible sabotaje, sufrido en los sistemas de control de la fábrica. Las perspectivas de crecimiento de la que hasta fechas recientes había sido una empresa de corte claramente familiar, que apuntaban a colocarla a la cabeza de su mercado, no eran ningún secreto en el sector. No obstante, nunca antes se había contemplado, por parte del equipo directivo, la acción de unos "hackers" como amenaza para los planes de expansión de la sociedad.

Tras una velada de intenso trabajo, analizando escrupulosamente lo sucedido, la hipótesis del ciberataque se consolidó. Diversas pistas encontradas en los equipos informáticos que controlaban la operativa del taller corroboraron la sospecha inicial. Sin embargo, el alcance del golpe recibido no llegó a adivinarse, ni con las suposiciones más pesimistas. ¡El mayor mazazo estaba, aún, por llegar!

A las 10:19 de la mañana del martes se recibió la notificación de las autoridades sanitarias que instaban al cierre preventivo de la planta (estaba claro que en el Ministerio de Sanidad no conocían las tensas últimas veinticuatro horas que había vivido la instalación). La razón del cierre: las intoxicaciones de las que se habían venido haciendo eco los noticiarios durante la semana anterior (varios cientos, ya, en este momento, repartidos por hospitales a lo largo de toda la geografía nacional) se habían relacionado, finalmente, con el consumo de algunos de los productos que la fábrica servía a su canal.

Los ciberdelincuentes, sin duda, habían hecho bien su trabajo. Lo vivido el día anterior no había sido más que el colofón a una actuación, iniciada semanas atrás. Durante ese tiempo habían tenido ocasión de profundizar en su conocimiento de los procesos llevados a cabo en la instalación, hasta llegar a modificar la proporción de los componentes que formaban parte del producto final.

Las consecuencias de dicha manipulación, ahora, estaban a la vista de todos.

La historia de ficción que acaba de leer, basada en una idea original de nuestro colega, José Valiente, Director del Centro de Ciberseguridad Industrial, no pretende, a pesar de su verosimilitud y gravedad, provocar el miedo en Ud.; más al contrario, únicamente intenta ayudarle a que abra los ojos.

Dicho esto, permítanos aclararle (por si llegados a este punto, el

asunto aún requiriera de aclaración alguna) que la razón por la que le presentamos este mini-armagedón en esta sección dedicada al gobierno corporativo de las tecnologías de la información reside en nuestro convencimiento de que los temas de ciberseguridad (como los de seguridad, en general, y todos los que puedan acarrear consecuencias [negativas] derivadas del [mal] uso de la tecnología en el seno de una organización) no son temas que pertenezcan sólo a los ámbitos de la gestión y/o la operación de la citada tecnología, sino que son una responsabilidad ineludible del órgano de gobierno (consejo de administración o equivalente) de la entidad.

En palabras de Tom Farley, presidente del Mercado de Valores de Nueva York, "ningún problema ha creado, hoy, más preocupación entre consejeros y ejecutivos que el riesgo relacionado con la ciberseguridad"¹. Como directivo, intente que éste sea también su caso. Ello debería ayudarle a tomar las debidas cautelas antes de que, como en nuestra historia, sea demasiado tarde.

¹ **Matt Rosenquist (ed.)**. "Navigating the Digital Age: The Definitive Cybersecurity Guide for Directors and Officers" (Navegando por la Era Digital: La Guía Definitiva sobre Ciberseguridad para Consejeros y Ejecutivos). Caxton Business & Legal, Inc. Octubre de 2015. Prólogo (introducción) de Tom Farley, Presidente, Mercado de Valores de Nueva York.

<https://www.securityroundtable.org/wp-content/uploads/2015/09/Cybersecurity-9780996498203-no_marks.pdf>. Último acceso: 23-09-2016.

Sección Técnica "Ingeniería del Software" (Luis Fernández Sanz, Daniel Rodríguez García)

Tema: Libro

Rob de Bie, Bryan Bakker, Rene van den Eertwegh. *Finally... Reliable Software!: A practical approach to design for reliability*. ISBN: 9781499226669, 1499226667, 2015. Editorial: Createspace, USA.

El libro tiene una extensión de 135 páginas dividido en dos partes y dos apéndices.

Se trata de un libro práctico y sencillo que introduce rápidamente al lector en el concepto de fiabilidad (*reliability*) del software.

El texto huye de la descripción de fórmulas complicadas y de conceptos teóricos. En la primera parte ("*Fundamentals*"), los autores describen mediante ejemplos prácticos el concepto de fiabilidad del software y cómo aplicarlo a éste. Los autores proponen que el diseño del software esté orientado a la fiabilidad. Definen cuatro pasos a seguir para llegar a ese objetivo.

En la segunda parte, los autores describen un caso práctico de diseño de software con el objetivo de fiabilidad. Se trata de un sistema de vigilancia y seguridad. A lo largo de 40 páginas describen la aplicación de los conceptos de fiabilidad a ese caso de estudio.

Los apéndices describen una manera de analizar los fallos y una clasificación de defectos ortogonal.

En resumen, nos encontramos ante un libro introductorio de fácil lectura, de carácter práctico y de utilidad para quienes están preocupados por la fiabilidad del software.

Sección Técnica: "Lenguajes de Programación" (Oscar Belmonte Fernández, Inmaculada Coma Tatay)

Tema: Curso de introducción a la programación

Si eres asiduo lector de **Novática** probablemente eres programador, o al menos entiendes de programación, y como tal sabrás que la programación es vista, por los legos en la materia, como una actividad complicada, oscura y ajena a toda satisfacción.

Si eres programador también sabrás que esta visión poco tiene que ver con la realidad. La programación es una tarea tremendamente gratificante, que exige un alto grado de creatividad y que, afortunadamente, puede ser enseñada a otros.

Y este es el objetivo del curso “*Introducción a la programación. Descubre el lenguaje de la era digital*”, un curso gratuito ofrecido a través de la plataforma MiriadaX, que iniciará en breve su segunda edición.

Una excelente oportunidad para todos aquellos que quieran aproximarse al mundo de la programación sin necesidad de ningún nivel previo. <<https://miriadax.net/web/introduccion-a-la-programacion-descubre-el-lenguaje-de-la-era-digital>>.

Tema: Libro

Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia. *Learning Spark Lightning: Fast Big Data Analysis*. O’Reilly Media, Edición: 1 (27 de febrero de 2015). ISBN-10: 1449358624.

Vivimos en la era de los datos. Los datos son producidos por humanos, a través, por ejemplo, de las redes sociales, los dispositivos electrónicos, y las redes de sensores. Y están a nuestra disposición, en mayor o menor medida, a través de Internet.

Pero todos sabemos que hay una gran distancia entre el frío dato y la información, y que transformar uno en la otra suele ser tarea costosa no exenta de errores. Más aún, el increíble volumen de datos a tratar para extraer información es a veces disuasorio por falta de recursos materiales, tiempo o profesionales cualificados.

Apache Spark <<http://spark.apache.org/>> es un *framework* para el procesamiento de grandes volúmenes de datos. Una de sus grandes diferencias frente a otras soluciones, como el conocido Hadoop <<http://hadoop.apache.org/>>, es que puede trabajar, si el volumen de datos lo permite, directamente en memoria, siendo por ello mucho más rápido.

Apache Spark está programado en Scala y es posible acceder a su API desde este lenguaje de programación, pero también desde Java y Python. Si te quieres iniciar en la ciencia de los datos, este libro es una detallada y práctica introducción a este mundo de la mano de Apache Spark.

Tema: Libro

Brian W. Fitzpatrick, Ben Collins-Sussman. *Debugging Teams: Better Productivity through Collaboration*. O’Reilly Media, Edición: 1 (26 de octubre de 2015). ISBN-10: 1491932058.

Una buena organización en los grupos de trabajo es esencial para alcanzar el éxito en cualquier tipo de proyecto, y los proyectos informáticos no son una excepción.

En el libro aquí reseñado, los autores nos cuentan las claves que han encontrado para la buena gestión de los grupos de trabajo a través de su propia experiencia.

El libro se inicia con un capítulo titulado “*The myth of the genius programmer*”, donde se desarticula la falsa creencia del valor de los programadores geniales, que trabajan de una manera autónoma, en favor de los grupos de trabajo altamente cohesionados. En los siguientes capítulos se sigue desgranando las claves en la gestión de grupos de trabajo.

Este es un libro de lectura muy recomendable tanto para gestores de grupos como para los integrantes de los mismos.

Sección Técnica: “Modelado de software” (Jesús García Molina, Gustavo Rossi)

Tema: Un libro sobre la historia del modelado del software

En esta columna hemos comentado muchos libros relacionados con la tecnología orientada a objetos y el modelado de software. En esta ocasión, sin embargo vamos a llamar la atención sobre la necesidad de que se escriba un libro que repase la historia del modelado en el desarrollo de software.

Creemos que relatar e interpretar cuál ha sido la evolución en el uso de modelos para la creación de software es un tema de gran interés para la comunidad de la Ingeniería del Software. Aquí nos limitaremos a esbozar de forma superficial las tres etapas o hitos que debería cubrir ese libro.

Los modelos se han utilizado desde los primeros años del desarrollo de software. En 2014, Juha Pekka Tolvanen presentó un interesante dibujo que ilustraba la evolución de los lenguajes de modelado, mostrando las dependencias entre ellos, como se puede ver en [1].

Diagramas de flujo, redes de Petri, máquinas de estado y escenarios de interacción son quizá los modelos más extendidos y han sido de gran utilidad a lo largo de la historia de la programación. Durante la década de los ochenta surgieron un montón de métodos estructurados con sus correspondientes notaciones pero tuvieron un impacto muy limitado. En aquellos años también se definieron un buen número de lenguajes de especificación formal aunque la mayoría no tuvieron el éxito esperado dado que realmente no servían al propósito de ayudar a escribir o a generar código.

En [2] se puede encontrar un excelente análisis de Bran Selic sobre la “Ingeniería del software basada en modelos” (*Model-Driven Engineering*, MDE) que analiza la evolución del modelado y comenta las limitaciones de los lenguajes de especificación formal, lenguajes de cuarta generación y herramientas CASE.

En los 90, con el éxito de la orientación a objetos y su adopción industrial, se produjo la aparición de un montón de métodos orientados a objetos, lo cual dio lugar a una gran confusión y a lo que Ivar Jacobson denominó “Guerra de Métodos”.

Fue entonces cuando emergió UML (*Lenguaje Unificado de Modelo*) como una notación estándar para el modelado basado en los conceptos de la orientación a objetos.

Es bien conocido el éxito de UML y menos conocida su escasa adopción industrial. Nuestra experiencia coincide con la conclusión del estudio empírico realizado por Marian Petre en 2013 [3]. UML sólo es utilizado por un porcentaje reducido de desarrolladores que sólo utilizan una pequeña parte de la notación para tareas muy concretas dentro del ciclo de vida: diagramas de clases para razonar sobre las clases del dominio, diagramas de interacción para documentar y máquinas de estado cuando hay comportamiento dependiente del estado.

Cabe destacar que Grady Booch en una conferencia que impartió en ICSE'2002 señaló que se abría una nueva etapa para UML como lenguaje para el desarrollo dirigido por modelos ([4]. Sin embargo, UML ha fracasado en ese intento y se puede pensar, ahí están los comentarios del propio Grady Booch (por ejemplo “*When we began with the UML, we never intended it to become a programming language*” [5]), que uno de los motivos de este fracaso ha sido el intento de OMG de convertirlo en un lenguaje de modelado de propósito general, capaz de cubrir las necesidades de MDE, debido a que este objetivo ha propiciado que se convierta en un lenguaje grande y complejo.

Cabe resaltar que Bran Selic defiende que UML no es tan complejo como se piensa [5]. Si analizamos el metamodelo de UML parece que uno debería inclinarse más por consideraciones como las que realiza Grady Booch “*The UML metamodel became grossly bloated, because of the drive to model driven development. I think that complexity was an unnecessary mistake*”.

Y como tercer hito en la evolución del modelado del software, hay que considerar la emergencia de MDE como una nueva disciplina, a raíz del lanzamiento de MDA por parte de OMG.

En la última década, la creación de DSLs ha aumentado considerablemente como se señala en [6], así como del número de herramientas para automatizar su creación, como se ha discutido en varias de nuestras columnas.

Por otra parte, pensamos que, de forma gradual, está aumentando el número de entornos que proporcionan lenguajes de modelado para automatizar el desarrollo de código, aunque MDE esté lejos de ser adoptado por la industria del software.

Jordi Cabot en un reciente artículo en su blog “*Modeling Languages*” comenta un reciente informe de Forrester que señala como una tendencia actual las plataformas “*low-cost*” para la producción de software, y en su opinión “*low-cost platforms*” es una nueva palabra de moda para hablar de generadores de código basados en modelos de cualquier naturaleza [7].

En definitiva, quién acepte el reto de escribir este libro tiene por delante una desafiante y atractiva tarea: contar la historia de los lenguajes de modelado, analizar el papel que ha jugado y jugará UML, y analizar el futuro del desarrollo basado en modelos, entre otras cuestiones.

Referencias

- [1] **Jordi Cabot**. *History of modeling languages in one picture* (by J-P Tolvanen). <<http://modeling-languages.com/history-modeling-languages-one-picture-j-p-tolvanen/>>, 16 de abril de 2014.
- [2] **Bran Selic**. “Personal reflections on automation, programming culture, and model-based software engineering”. *Autom. Softw. Eng.* 15(3-4): pp. 379-391, 2008.
- [3] **Marian Petre**. “UML in practice”. *ICSE 2013*: pp. 722-731.
- [4] **Grady Booch**. “Growing the UML”. *Software and System Modeling* 1(2): pp. 157-160, 2002.
- [5] **Jordi Cabot**. *What do their creators think about UML now?* <<http://modeling-languages.com/uml-opinions-creators/>>, 5 de agosto de 2016.
- [6] **Jon Whittle, John Edward Hutchinson, Mark Rouncefield**. “The State of Practice in Model-Driven Engineering”. *IEEE Software* 31(3): pp. 79-85, 2014.
- [7] **Jordi Cabot**. *Low-code platforms, the new buzzword*. <<http://modeling-languages.com/low-code-platforms-new-buzzword/>>, 1 de septiembre de 2016.

Sección Técnica “Seguridad”

(Javier Areitio Bertolín, Javier López Muñoz)

Tema: Libros

- **S. Li, L. Xu**. *Securing the Internet of Things*. Syngress. ISBN: 0128044586, 2017.
- **M.E. Whitman, H.J. Mattord**. *Principles of Information Security*, 6th Edition. Cengage Learning. ISBN: 1337102067, 2017.
- **I. Traore, A. Awad, I. Woungang**. *Information Security Practices: Emerging Threats and Perspectives*. Springer. ISBN: 3319489461, 2017.
- **J. Bone**. *Cognitive Hack: The New Battleground in Cybersecurity . . . the Human Mind*. Auerbach Publications. ISBN: 149874981X, 2017.
- **M. Goodrich, R. Tamassia**. *Introduction to Computer Security*, 2nd Edition. Addison-Wesley. ISBN: 0133575470, 2017.
- **F. El Khoury**. *Iris Biometric Model for Secured Network Access*. CRC Press. ISBN: 1138198722, 2017.
- **H. Abie**. *Adaptive Security for the Internet of Things: Research, Standards and Practices*. Syngress. ISBN: 0128016272, 2017.
- **R.M. Lee**. *Active Cyber Defense Cycle*. Wiley. ISBN: 1119215196, 2017.

Tema: Congresos y conferencias

- **International Conference on Cyber-Physical Systems**. Del 13 al 14 de enero de 2017. Zurich, Suiza.
- **ICS Cyber Security**. Del 26 al 27 de abril de 2017. Londres, UK.
- **Threat Intelligence Summit**. Del 25 de enero al 1 de febrero de 2017. Arlington, Virginia, USA.
- **CDANS Cyber Defence and Network Security**. Del 25 al 26 de enero de 2017. Londres, UK.
- **Black Hat USA**. Del 22 al 27 de julio de 2017. Mandalay Bay, Las Vegas, Nevada, USA.

Sección Técnica “Software Libre”

(Jesús M. González Barahona, Fernando Tricas García)

Tema: XI edición del Concurso Universitario de Software Libre

Como viene siendo habitual desde el año 2016, por estas fechas (mes de octubre) se ha abierto el periodo de inscripción para el Concurso Universitario de Software Libre que alcanza ya con la de este año su 11ª edición.

Se trata de un concurso de desarrollo de software, hardware y documentación técnica libre en el que pueden participar estudiantes universitarios de primer, segundo y tercer ciclo; así como estudiantes de bachillerato, ciclos de grado medio y superior; y universitarios (incluido grado, máster y doctorado) del ámbito estatal español.

Durante el transcurso del mismo, los participantes dispondrán de un blog donde contarán su experiencia en el desarrollo durante el curso académico, además de emplear un repositorio, como GitHub o LaunchPad, para alojar el código fuente.

El objetivo principal del Concurso es fomentar la creación y contribuir a la consolidación de la comunidad del Software Libre en la Universidad.

El periodo de inscripción se extiende este año entre el 15 de octubre de 2016 y el 15 de febrero de 2017, estando prevista la fase de evaluación de los proyectos para el mes de abril de 2017 y el evento final, con la entrega de premios a los ganadores, para mayo de 2017.

El Concurso Universitario de Software Libre cuenta con una amplia base de patrocinadores entre los que destacan este año, como

patrocinadores principales, las empresas informáticas Zevenet y Solutia. Al igual que en las 10 ediciones anteriores, *Novática* consta como entidad colaboradora entre los medios oficiales de difusión del evento.

Quisiéramos destacar una vez más el mérito de los organizadores de este Concurso y la enorme utilidad de sus esfuerzos con vistas a completar la formación de los estudiantes universitarios informáticos proveyéndoles de los recursos necesarios para la realización de proyectos reales en un marco a la vez colaborativo y competitivo en el área del desarrollo de software libre.

Reiteramos desde aquí nuestras más sinceras felicitaciones por esta magnífica iniciativa.

Para ampliar información, he aquí los enlaces principales:

Página principal:

<<https://www.concursosoftwarelibre.org/1617/portada>>.

Calendario:

<<https://www.concursosoftwarelibre.org/1617/calendario>>.

Patrocinadores y colaboradores:

<<https://www.concursosoftwarelibre.org/1617/node/14>>.

Sección Técnica: “TIC y Turismo”

(Andrés Aguayo Maldonado, Antonio Guevara Plaza)

Tema: *Libros*

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) ha publicado el libro *OECD Tourism Trends and Policies 2016* (Tendencias y políticas de turismo de la OCDE 2016), en el que se recoge un análisis de la situación actual del turismo en los principales países del mundo y se hacen una serie de recomendaciones sobre las actuaciones públicas y privadas que pueden llevarse a cabo para adaptarse a una situación siempre cambiante.

El libro se articula en dos secciones. En la primera parte, que consta de tres capítulos, se presenta un resumen de las principales tendencias en turismo y las políticas y desarrollos asociados.

En el segundo capítulo, se examina la relación de interdependencia que existe entre el transporte y el turismo, la importancia de las conexiones integradas en centros de transporte, y los factores clave a mejorar la experiencia del visitante.

El tercero es el más interesante desde el punto de vista de las tecnologías de la información y las comunicaciones. Este capítulo trata sobre lo que significa el crecimiento exponencial de la economía para el intercambio en el sector turístico y evalúa las principales actuaciones políticas para adaptarse a esta economía compartida.

El rápido crecimiento del *peer-to-peer* y las plataformas de uso compartido ha afectado de forma importante a áreas tan diversas como el transporte, el alojamiento, los viajes e incluso la restauración. Estos desarrollos presentan nuevas oportunidades comerciales, pero también grandes retos a los responsables de las políticas de turismo, que afectan a cuestiones como los impuestos, la regulación o los derechos del consumidor. También se destaca la necesidad de reforzar los mecanismos de recopilación de datos y la investigación aplicada.

En la segunda parte del libro se presenta la información relacionada de cada uno de los países miembros de la OCDE (38) y de un grupo de países asociados. De cada país se describen la estructura general de las administraciones encargadas de tomar decisiones políticas

en materia de turismo y las principales iniciativas llevadas a cabo. También se muestran en forma de tabla, desglosadas por país, la evolución de las principales cifras asociadas al turismo.

Estas tablas y sus gráficos asociados están disponibles para su descarga, en formato Excel, desde la web de la librería de la OCDE: <<http://www.oecd-ilibrary.org/>>.

El otro libro, titulado *Tourism Management, Marketing, and Development: Performance, Strategies, and Sustainability* (2016), tiene un enfoque más académico y es el resultado del esfuerzo intelectual colectivo de una serie de expertos internacionales, que cultivan la investigación original sobre la gestión del turismo desde una variedad de perspectivas teóricas.

Está coordinado, entre otros, por **Dimitrios Buhalis**, uno de los principales referentes internacionales en el campo de las TIC aplicadas al turismo.

Julio Javier Castillo, Diego Javier Serrano, Marina Elizabeth Cárdenas

Laboratorio de Investigación de Software MsLabs, Dpto. Ing. en Sistemas de Información, Facultad Regional Córdoba - Universidad Tecnológica Nacional (Argentina)

<jotacastillo@gmail.com>, <diegojserrano@gmail.com>, <ing.marinacardenas@gmail.com>

El problema del robot de exploración de Marte

Este es el enunciado del problema 5 que fue planteado en la Sexta Competencia de Programación de la Facultad Regional de Córdoba (Universidad Tecnológica Nacional, Argentina) UTN-FRC celebrada en octubre de 2014.

Nivel del problema: Medio

La Agencia Espacial Europea (ESA) es una organización internacional dedicada a la exploración espacial y está desarrollando un proyecto para colocar un robot de exploración sobre la superficie de Marte. El robot debe ser capaz de realizar exploraciones de forma remota controlado por un operador desde la Tierra ayudado por un programa que le permita realizar movimientos válidos. Dado que un mal movimiento del robot puede suponer el fracaso de la misión, el módulo de IA debe comportarse de forma conservadora y notificar permanentemente al robot acerca de posiciones peligrosas.

El robot parte siempre desde una posición base donde aterrizó la nave de transporte, que considera totalmente segura. A partir de allí, realiza movimientos de exploración siguiendo las órdenes del operador remoto. Cuando el módulo de IA determina que la posición actual es peligrosa, el robot realiza movimientos de regreso hacia la posición anterior para continuar desde ella su exploración por otros caminos.

El robot será capaz de realizar cuatro tipos de movimientos básicos: desplazamientos hacia adelante, atrás, derecha e izquierda. Para realizar un desplazamiento es necesario especificar la distancia (número entero que se representa en centímetros).

Como parte del equipo de ingenieros de la ESA, se le ha encargado realizar un programa que permita registrar los movimientos del robot de tal forma que éste sepa volver en todo momento a la posición base deshaciendo movimientos básicos hechos previamente. De esta manera, el operador del robot podrá ser capaz de conocer la ruta segura que debe realizar el robot para volver a la nave y la distancia total recorrida por el mismo.

Para poder identificar los obstáculos, la zona de navegación del robot está representada con una matriz dada, donde cada celda representa 1 centímetro cuadrado. El robot parte de la posición determinada en el pun-

to medio de la matriz donde se encuentra la nave (por ejemplo para una matriz de 10 x 10 se considera el punto medio en la celda 4 x 4, para una matriz de 15 x 15 se considera el punto medio la celda 7 x 7) y procede con la navegación por la grilla según las indicaciones del operador. Previamente se ha realizado un reconocimiento del terreno por medio de imágenes satelitales y se han cargado los datos de los obstáculos en una matriz. Si en el desplazamiento del robot, se detecta algún obstáculo, el mismo retrocede a la posición anterior segura conocida, es decir a la posición anterior a la realización del desplazamiento actual. Cuando el operador termina la navegación el sistema debe mostrar el camino que el robot debe realizar hasta la nave y la distancia total recorrida.

Entrada

Se indicará con un número las dimensiones del área de exploración definida por una matriz cuadrada binaria. Las siguientes líneas indicarán cómo está completa el área de exploración, donde los obstáculos estarán indicados por el número "1" y los espacios sin obstáculos estarán indicados por el número "0".

En la siguiente línea se indicará la cantidad de desplazamientos que debe realizar el robot desde la posición inicial, identificando cada uno de los movimientos con un número (0: Arriba, 1: Abajo, 2: Derecha, 3: Izquierda) y, separada por un espacio, la distancia recorrida. Los movimientos que forman parte de un desplazamiento estarán separados entre sí por una coma. Por ejemplo: 0 10, 2 15, 1 25, 3 10 (suponiendo una matriz cuadrada de, al menos, 30 x 30) lo que se traduce a 10cm a hacia arriba, 15cm a la derecha, 25cm hacia abajo y 10cm a la izquierda.

Salida

Por cada caso de prueba (cada desplazamiento), se deberá imprimir la distancia total recorrida si los movimientos desde el origen hacia el destino están libres de obstáculos, y "NO SEGURO" en el caso que

se detecte un obstáculo en alguna parte del desplazamiento. En caso que la distancia recorrida por un determinado movimiento supere las dimensiones de la matriz, se considerará la presencia de un obstáculo.

Ejemplo de entrada

```
10
1 1 1 0 0 0 0 1 1 1
1 1 1 0 0 0 0 0 0 0
0 0 0 0 0 0 1 1 0 0
0 0 1 0 0 0 0 1 0 0
0 0 1 0 0 0 0 0 0 1
1 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 1
0 0 0 0 1 1 0 0 0 1
0 0 0 0 0 1 0 0 0 0
0 0 0 0 0 0 0 0 0 0
```

```
5
0 10, 2 15, 1 25, 3 10
3 1, 1 5, 2 5, 0 7
3 1, 1 5, 2 5, 0 7, 3 1
1 5, 3 4, 2 5, 0 5
1 2, 2 4, 0 5, 3 5, 1 8
```

Ejemplo de salida

```
NO SEGURO
18
NO SEGURO
NO SEGURO
24
```

Julio Javier Castillo, Diego Javier Serrano, Marina Elizabeth Cárdenas

Laboratorio de Investigación de Software MsLabs, Dpto. Ing. en Sistemas de Información, Facultad Regional Córdoba - Universidad Tecnológica Nacional (Argentina)

<jotacastillo@gmail.com>, <diegojserrano@gmail.com>, <ing.marinacardenas@gmail.com>

En este problema se nos solicita programar un algoritmo que permita calcular el espacio de almacenamiento que ocupan los directorios de determinados discos duros, e informar la cantidad de kilobytes que se ahorrarían si se eliminan las copias de archivos repetidos.

Para resolver este problema se propone el uso de tablas de dispersión (*hash*) que nos permitan almacenar (sin repeticiones) los nombres de los directorios de un disco duro, y la cantidad de bytes que ocupan todos los archivos de ese directorio. Asimismo, se propone el uso de otra tabla de dispersión cuyo valor de clave esté representado por el nombre del archivo y su tamaño, y el contenido con un valor igual al tamaño del archivo.

De esta manera, el problema puede resolverse realizando una tabla *hash* de acumulación, la cual se encuentra representada por un `HashMap` de “directorios” en la solución propuesta. Para ello se analizarán cada uno de los archivos del disco duro y se acumulará el tamaño de los mismos.

Evidentemente, este problema puede ser resuelto mediante el empleo de otras estructuras de datos como vectores o listas, pero se ha optado por el uso de tablas de dispersión dado que presentan un costo computacional menor respecto de las otras estructuras mencionadas. El costo de acceso y de acumulación en el interior de la tabla es constante.

La solución propuesta se codifica en el lenguaje de programación Java y utiliza las clases `HashMap` que son las estructuras de tablas de dispersión con las que cuenta Java en su paquete `java.util`.

La primera parte del programa lee los casos de prueba que en el problema representan diferentes discos duros. Seguidamente, se leen los archivos junto con los tamaños que los mismos ocupan. Se emplea el método `lastIndexOf()` de la clase `String` para encontrar la última posición de las “\”, los cuales sirven para delimitar el nombre del archivo y su ruta. Esta información se utiliza para poblar las tablas de dispersión.

Debemos notar que la tabla denominada “archivos” utiliza como clave al valor “nombre_archivo+tamaño” ya que es una manera de identificar las repeticiones, dado que en el planteo del problema solo se consideran archivos repetidos a aquellos que tengan el mismo nombre y tamaño.

Discos duros

El enunciado de este problema apareció en el número 236 de *Novática* (abril-junio 2016, p. 79).

La visualización de los resultados en consola se lleva a cabo realizando una iteración (mediante la estructura repetitiva denominada `for-each` de Java) sobre la tabla de dispersión. Nótese que los valores de bytes son divididos por 1.024 para obtener el correspondiente valor en kilobytes.

Finalmente, se provee como salida el valor del ahorro al “borrar archivos innecesarios” indicado en cantidad de kilobytes.

A continuación se presenta la solución del problema en el lenguaje de programación Java:

```
import java.util.HashMap;
import java.util.Map;
import java.util.Scanner;

public class DiscosDuros {

    public static void main(String[] args) {
        Scanner sc = new Scanner(System.in);

        String linea = sc.nextLine();
        int C = Integer.parseInt(linea);

        for (int caso = 1; caso <= C; caso++) {
            linea = sc.nextLine();
            int N = Integer.parseInt(linea);

            HashMap<String,Long> archivos = new HashMap<>();
            HashMap<String,Long> directorios = new HashMap<>();
            long ahorro = 0;

            for (int i = 0; i < N; i++) {
                linea = sc.nextLine();
                int posEsp = linea.lastIndexOf(" ");

                long tamaño = Long.parseLong(linea.substring(posEsp+1));
                int posUBI = linea.lastIndexOf("\\");
                int posPBI = linea.indexOf("\\\\",2);

                String nombreDir = linea.substring(1,posPBI);
                String nombreArchivo = linea.substring(posUBI+1,posEsp-1);
                String nombreCompleto = nombreArchivo + " " + tamaño;

                if (!archivos.containsKey(nombreCompleto))
                    archivos.put(nombreCompleto, tamaño);
                else
                    ahorro += tamaño;

                if (!directorios.containsKey(nombreDir))
                    directorios.put(nombreDir, tamaño);
                else
                {
                    Long tamActual = directorios.get(nombreDir);
                    directorios.put(nombreDir, tamActual + tamaño);
                }
            }

            System.out.println("Disco duro " + caso + ":");

            for (Map.Entry<String, Long> dir: directorios.entrySet()) {
                Long tamañoKB = (long)Math.ceil(dir.getValue() / 1024f);
                System.out.println(dir.getKey() + " " + tamañoKB);
            }
            long ahorroKB = (long)Math.ceil(ahorro / 1024f);
            System.out.println(ahorroKB);
        }
    }
}
```

Nombramiento de Representantes de ATI en IFIP, CEPIS y CLEI

En la reunión de Junta Directiva General celebrada en noviembre, se acordó nombrar a **Carlos Juiz García** representante de ATI en IFIP (*International Federation for Information Processing*), a Luis Fernández Sanz en CEPIS (*Council of European Professional Informatics Societies*), y en CLEI (*Centro Latinoamericano de Estudios en Informática*) a **Dídac López Viñas**. A continuación, publicamos una breve nota biográfica de cada uno de ellos.

Carlos Juiz García es Profesor Titular de Universidad, y diplomado, licenciado y doctor en Informática por la Universidad de las Islas Baleares (UIB). Es el investigador principal del grupo de investigación ACSIC <<http://acsic.uib.es>> y su interés en la investigación se centra principalmente en ingeniería de rendimiento, web semántica y gobierno de TI. Es miembro sénior del IEEE y también sénior de la ACM. También es experto invitado de la Unión Internacional de Comunicaciones (ITU) y Abogado Académico de ISACA. Ha sido también nombrado miembro del Comité de dominio sobre el Cloud Computing de IFIP. Actualmente es miembro de la junta directiva del *clúster* TURISTEC y también del *clúster* Balears.t. Es, además, el coordinador del subgrupo de Gobierno de TI en AENOR, la entidad española en ISO.

Luis Fernández Sanz es licenciado y doctor en informática con premio extraordinario de doctorado. Profesor titular de la Universidad de Alcalá. Ha sido Vicepresidente de CEPIS (Consejo Europeo de Sociedades Profesionales de Informática) entre 2011 y 2013. También ha sido vicepresidente de ATI entre 2011 y 2016 donde impulsó la participación de la asociación en dos proyectos europeos: Appskil <<http://appskil.eu/es/>> y Coding for Young People <<http://codingforyoungpeople.eu/>>. Ha sido CEO de una PYME de servicios informáticos durante 4 años y ha sido profesor en la Universidad Politécnica de Madrid y en la Universidad Europea de Madrid donde también fue director de departamento y de titulaciones. Es el único experto español participante en la nueva clasificación laboral ESCO para el sector TIC. Actualmente, dirige varios proyectos europeos <<https://www.m-commerce.enterprises/>>, <<http://www.iceberg-sqa.eu/>>, o <<http://openqass.itstudy.hu/>> con algunos de ellos trabajando sobre competencias y cualificaciones TIC: <<http://www.eskillsmatch.eu/es/>> y <<http://www.ecfalliance.org/>>.

Dídac López Viñas, socio 5561 de ATI, es Jefe del Servicio Informático de la Universidad de Girona, Director de TIC del Parque Científico y Tecnológico de la UdG, y Consultor de la UOC. Es Licenciado en Informática por la UPC, Postgrado en Dirección de Informática por el ICT (UPC), postgrado en Gestión de la Información en las Organizaciones (infonomía, UPF) y MBA por Les Heures (UB). Fue presidente de ATI hasta febrero de 2016 y durante 5 años, y previamente Presidente del Capítulo de Cataluña.

Programación de Novática

Por acuerdo del Consejo Editorial de *Novática*, el tema y editor invitado de la primera monografía de 2017 serán, salvo causas de fuerza mayor o imprevistos, los siguientes:

Nº 238 (noviembre 2016-febrero 2017): "Seguridad digital"

Editor invitado: **Miguel García-Menéndez** (Responsable de Gobierno Corporativo y Estrategia del Centro de Ciberseguridad Industrial).

Socios institucionales de ati

Según los Estatutos de ATI, pueden ser socios institucionales de nuestra asociación "*las personas jurídicas, públicas y privadas, que lo soliciten a la Junta Directiva General y sean aceptados como tales por la misma*".

Mediante esta figura asociativa, todos los profesionales y directivos informáticos de los socios institucionales pueden gozar de los beneficios de participar en las actividades de ATI, en especial congresos, jornadas, cursos, conferencias, charlas, etc. Asimismo los socios institucionales pueden acceder en condiciones especiales a servicios ofrecidos por la asociación tales como Bolsa de Trabajo, cursos a medida, *mailings*, publicidad en *Novática*, servicio ATInet, etc.

Para más información dirigirse a <info@ati.es> o a cualquiera de las sedes de ATI. En la actualidad son socios institucionales de ATI las siguientes empresas y entidades:

AGROSEGURO, S.A.
COSTAISA, S.A.
ESTEVE QUÍMICA, S.A.
FUNDACIÓ CATALANA DE L'ESPLAI
FUNDACIÓ PRIVADA ESCOLES UNIVERSITÀRIES GIMBERNAT
FUNDACIÓN TALENTO MATEMÁTICO
INSTITUT D'ESTUDIS CATALANS
INSTITUT MUNICIPAL D'INFORMÀTICA
INSTITUTO MADRILEÑO DE FORMACIÓN
KRITER SOFTWARE, S.L.
MASTER. D MASTER DISTANCIA, S.A.
NEXTRET, S.L.
ONDATA INTERNATIONAL, S.L.
PROGRAMAMOS
S Q S, S.A.
SE ACEPTAN IDEAS - COWORKING TIC
SERIFORM MELILLA
SOCIEDAD DE REDES ELECTRICAS
UNIVERSIDAD DE ALCALÁ DE HENARES (UAH)
UNIVERSIDAD EUROPEA DE MADRID
UNIVERSITAT DE GIRONA
UNIVERSITAT OBERTA DE CATALUNYA

Todos los datos son obligatorios a menos que se indique otra cosa / All the data must filled in unless otherwise stated

Una vez cumplimentada esta hoja, se ruega enviarla a / Please fill in this form and send it to:
e-mail novatica.suscripciones@atinet.es o/or ATI, C/ Ávila 50, 3a planta, local 9 - 08005 Barcelona, España / Spain

Nota importante / Important Notice: Novática es una revista que se publica solamente en formato digital, de aparición trimestral, es decir cuatro números al año¹ / Novática is a digital-only publication that appears quarterly, i.e. four issues per year¹.

► **Cuota anual: 50 Euros** (IVA incluido – este impuesto se aplica solamente a residentes en España) / **Annual fee: 50 Euros** (VAT applicable only to subscribers that reside in Spain)

- El suscriptor es una empresa o entidad ___ o una persona física ___ (marcar con X lo que corresponda) /
- The subscriber is an organization (business, university, government, etc) ___ or a person ___ (mark your option with X)

- Datos del suscriptor empresa o entidad / Data of organizational subscriber

Empresa o entidad / Organization	Sector / Business
Dirección / Address	
Localidad / City	Cód. Postal / Post Code
Provincia / Country	
Datos de la persona de contacto / Data of contact person	
Nombre y apellidos / Full name	
Correo electrónico / E-mail address ¹	Teléfono / Phone

- Datos del suscriptor persona física / Data of personal subscriber ²

Apellidos / Last name	
Nombre / First name	
Localidad / City	Cód. Postal / Post Code
Provincia / Country	Teléfono / Phone
Correo electrónico / E-mail address ¹	

- Datos bancarios para domiciliación del pago / Bank account data for payment (si desea pagar por otro método contacte por favor con novatica.suscripciones@atinet.es / if you want your payment to be made using a different method please contact novatica.suscripciones@atinet.es)

Nombre de la entidad bancaria / Name of the Bank (if the bank is not located in Spain please provide also BIC Code)

IBAN:

Cód. país/Country Code	Cód. Banco/Bank Code	Cód. oficina / Branch Code	DC/CD	Núm. Cuenta / Account number

- NIF para su factura / Tax ID for invoice

Firma / Signature

Fecha / Date

Mediante su firma la persona que ha cumplimentado este impreso declara que todos los datos contenidos en el mismo son ciertos y acepta todos los términos y condiciones del servicio de suscripción a Novática / Along with his/her signature the person filling in this form declares that all the data provided are true and accepts all the terms and conditions of the Novática subscription service

Nota sobre protección de datos de carácter personal / Data Protection Notice: De conformidad con la LO 15/99 de Protección de Datos de Carácter Personal, le informamos de que los datos que usted nos facilite serán incorporados a un fichero propiedad de Asociación de Técnicos de Informática (ATI) para poder disfrutar de los servicios que su condición de suscriptor de Novática socio le confiere, así como para enviarle información acerca de nuevos servicios y ofertas que ATI ofrezca en relación con sus publicaciones. Si usted desea acceder, rectificar, cancelar u oponerse al tratamiento de sus datos puede dirigirse por escrito a secregen@ati.es. / ATI is fully compliant with the Spain Data Protection Law (LO 15/99). You can enact your rights to access, cancellation or opposition writing to secregen@ati.es.

¹ Una vez validados por el servicio de suscripciones de Novática los datos de este formulario, Vd. recibirá la información sobre el procedimiento para acceder a la zona de la Intranet de ATI donde se almacenan los números publicados por nuestra revista / Once the data in this form have been validated by the Novática subscription staff you will receive the information about the procedure required to access the ATI Intranet area where the issues edited by our journal are stored.

² Si Vd. es profesional informático o estudiante de Informática, o simplemente una persona interesada por la Informática, debe tener en cuenta que la revista Novática es solamente uno de los diferentes servicios que los socios de ATI reciben como contrapartida de su cuota anual, de forma que, muy probablemente, le será más beneficioso hacerse socio que suscribirse únicamente a la revista. Por ello le recomendamos que se informe sobre qué es ATI y sobre los servicios que ofrece en <http://www.ati.es/> o en info@ati.es.



Hoja de solicitud de inscripción en ATI (2016)

(Asociación de Técnicos de Informática)

La solicitud puede hacerse también mediante una hoja online disponible en <http://www.ati.es/sersocio>

Todos los datos son obligatorios a menos que se indique otra cosa
Una vez cumplimentada esta hoja, se ruega enviarla por correo electrónico a secregen@ati.es,
o por correo postal a ATI, Calle Ávila nº 50, 3ª Planta, local 9 - 08005 Barcelona

- Solicito inscribirme como: **Socio de número** (88€)* / **Socio junior** (28€)* / **Socio jubilado** (28€)* / **Socio adherido** (60€)*
(Para inscribirse como **socio estudiante** se ruega utilizar la hoja de inscripción específica disponible en <http://www.ati.es/estudiantes>
- ver en la siguiente página información detallada sobre ATI y los diferentes tipos de socios)

* **Nota importante:** la cuota cubre el año natural, de 1 de enero a 31 de diciembre. Las inscripciones a socios de número realizadas de 1 de julio a 31 de octubre tienen una reducción de cuota del 50% y todas las cuotas son gratuitas si se realizan del 1 de noviembre al 31 de diciembre. En este último caso, si se desea acceder a descuentos en servicios ofrecidos por terceros no se aplicarán reducciones a la cuota anual de asociado, que deberá abonarse en su totalidad.

- Datos personales del solicitante

Apellidos		
Nombre		
Domicilio	Nº	Piso
Localidad	Código Postal	
Provincia	Teléfono	
Dirección de correo electrónico ¹		
Fecha de nacimiento	DNI	

- Datos de la empresa o entidad donde trabaja (si es autónomo indíquelo en el campo "Empresa o entidad")

Empresa o entidad	Sector
Puesto actual	Depto.
Dirección	Nº
Localidad	Código Postal
Provincia	Teléfono

- Domiciliación de la cuota anual (ATI se encarga de su envío al banco o caja)

Nombre de la entidad bancaria: _____

IBAN:

Cód. país	Cód. Banco	Cód. oficina	D.C.	Núm. Cuenta

- Datos complementarios (si necesita más espacio para estos datos continúe en otra hoja)

Títulos superiores o medios que posee y centros otorgantes:

Resumen de experiencias profesionales:

Número de años de experiencia profesional informática:

- Presentado por los Socios de número (**)

(**) Esta información no es necesaria para solicitar inscribirse como socio junior, estudiante o adherido; para inscribirse como socio de número o jubilado, si el solicitante no conoce a ningún socio de número que pueda presentarle, la Secretaría General de ATI le contactará para determinar otra forma fehaciente de acreditar su profesionalidad.

1) Apellidos y Nombre Nº de socio Fecha .../.../..... Firma

2) Apellidos y Nombre Nº de socio Fecha .../.../..... Firma

Firma del solicitante

Fecha _____

Mediante su firma el solicitante declara que todos los datos incluidos en esta solicitud son ciertos.

Nota sobre protección de datos de carácter personal: De conformidad con la LO 15/99 de Protección de Datos de Carácter Personal, le informamos de que los datos que usted nos facilite serán incorporados a un fichero propiedad de Asociación de Técnicos de Informática (ATI) para poder disfrutar de los servicios que su condición de socio le confiere, así como para enviarle información acerca de nuevos servicios, ofertas y cursos que ATI ofrezca y puedan resultar de su interés. Sus datos podrán ser comunicados a aquellas instituciones, sociedades u organismos, con los que ATI mantenga acuerdos de colaboración, relacionados con el sector de los seguros, la banca y la formación para el envío de información comercial. Si usted desea acceder, rectificar, cancelar u oponerse al tratamiento de sus datos puede dirigirse por escrito a secregen@ati.es.

- No deseo recibir información comercial de ATI ni de terceras entidades colaboradoras de ATI.
- No deseo recibir información comercial de terceras entidades colaboradoras de ATI.
- No autorizo la comunicación de mis datos a terceras entidades colaboradoras de ATI.

¹ Una vez validados por la Secretaría de ATI la hoja de inscripción y los documentos requeridos, y aceptada su solicitud, Vd. recibirá en esta dirección de correo la información sobre el procedimiento para poder utilizar todos los servicios de la red ATINET (ver reverso).



Una asociación abierta a todos los informáticos

Una asociación útil a sus socios, útil a la Sociedad

Creada en 1967, **ATI (Asociación de Técnicos de Informática)** es la asociación profesional más numerosa, activa y antigua de las existentes en el Sector Informático español, con sedes en Barcelona (sede general), Madrid y Valencia. Cuenta con más de 3.000 socios, que ejercen sus funciones como profesionales informáticos en empresas, universidades y Administraciones Públicas, o como autónomos.

ATI, que está abierta a todos profesionales informáticos independientemente de su titulación, representa oficialmente a los informáticos de nuestro país en Europa (a través de CEPIS, entidad que coordina a asociaciones que representan a más de 400.000 profesionales informáticos de 32 países europeos) y en todo el mundo (a través de IFIP, entidad promovida por la UNESCO para coordinar trabajos de Universidades y Centros de Investigación), y pertenece a la CLEI (Centro Latinoamericano de Estudios en Informática). ATI tiene también un acuerdo de colaboración con ACM (*Association for Computing Machinery*).

En el plano interno tiene establecidos acuerdos de colaboración o vinculación con Ada Spain, ASTIC (Asociación Profesional del Cuerpo Superior de Sistemas y Tecnologías de la Información de la Administración del Estado), Hispalinux, AI2 (Asociación de Ingenieros en Informática), Colegios de Ingenierías Informáticas de Cataluña y con RITSI (Reunión de Estudiantes de Ingenierías Técnicas y Superiores de Informática).

Tipos de socio

✓ **Socios de número:** deben acreditar un mínimo de tres años de experiencia profesional informática (o dos años si se posee un título de grado superior o medio), *o bien* poseer un título de grado superior o medio relacionado con las Tecnologías de Información, *o bien* haber desarrollado estudios, trabajos, o investigaciones relevantes sobre dichas tecnologías

✓ **Socios estudiantes:** deben acreditar estar matriculados en un centro docente cuya titulación dé acceso a la condición de Socio de Número (la hoja específica de inscripción para socios estudiantes está disponible en <http://www.ati.es/estudiantes>)

✓ **Socios junior:** profesionales informáticos con una edad máxima de 30 años y que no sean estudiantes

✓ **Socios jubilados (Aula de Experiencia):** socios de ATI que, al jubilarse y cesar su actividad laboral, deciden continuar perteneciendo a ATI colaborando con su experiencia con la asociación

✓ **Socios adheridos:** profesionales informáticos que no cumplan las condiciones para ser Socios de Número o también personas que, no siendo profesionales informáticos, quieran participar en las actividades de ATI

✓ **Socios institucionales:** personas jurídicas, de carácter público o privado, que quieran participar en las actividades de ATI (para más información sobre esta modalidad se ruega ponerse en contacto con la sede general de ATI)

¿Qué servicios ofrece ATI a sus socios?

Mediante el pago de una cuota anual, los socios de ATI pueden disfrutar de la siguiente gama de servicios:

✓ **Formación Permanente**

- Cursos, Jornadas Técnicas, Mesas Redondas, Seminarios,
- Conferencias, Congresos
- Secciones Técnicas y Grupos de Trabajo sobre diversos temas
- Intercambios con Asociaciones Profesionales de todo el mundo

✓ **Servicios de información**

- Revista trimestral **Novática** (decano de la prensa informática española)
- Red asociativa **ATInet** (IntrATInet, acceso básico gratuito a Internet, correo electrónico con dirección permanente, listas de distribución generales y especializadas, foros, blogs, página personal, ...)
- Servidor web <http://www.ati.es>, pionero de los webs asociativos españoles.

✓ **Servicios profesionales**

- Asesoramiento profesional y legal
- Peritajes, diagnósticos y certificaciones
- Bolsa de Trabajo
- Emisión en España del certificado profesional europeo EUCIP (*European Certification of Informatics Professionals*)
- Emisión en España del certificado ECDL (*European Computer Driving License*) para usuarios

✓ **Servicios personales**

- Los que ofrece la Mutua de los Ingenieros (Seguros, Fondo de pensiones, Servicios Médicos)
- Los que ofrece la Caja de Ingenieros (gozar de las ventajas de ser socio de esta caja cooperativa)
- Promociones y ofertas comerciales

¿Dónde está ATI?

✓ **Sede General y Capítulo de Catalunya** – Calle Avila 50, 3a planta, local 9, 08005 Barcelona - Tlfn. 93 4125235; <secregen@ati.es>

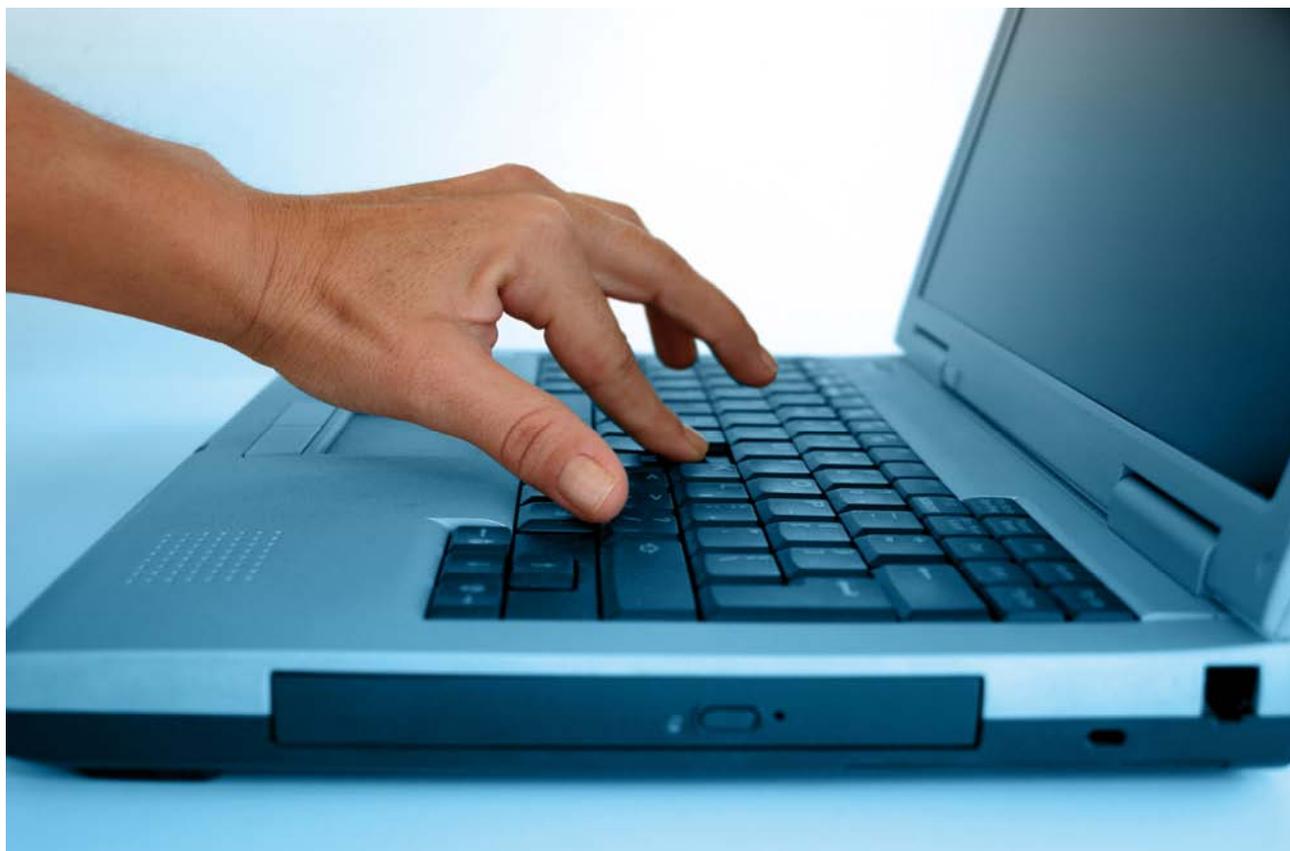
✓ **Capítulo de Andalucía** - <secreand@ati.es>

✓ **Capítulo de Galicia** - <secregal@ati.es>

✓ **Capítulo de Madrid** – Gutierre de Cetina, 24 28017 Madrid - Tlfn. 91 4029391; <secremdr@ati.es>

✓ **Capítulo de Valencia y Murcia** – Universidad Politécnica de Valencia. Asociación de Técnicos de Informática. Edificio 1H – ETSINF. Camino de Vera, s/n 46022 Valencia / <secreval@ati.es>

✓ **Revista Novática** – Gutierre de Cetina, 24 28017 Madrid – Tlfn. 91 4029391 / <novatica@ati.es>



Acreditación Europea de habilidades informáticas

Líder internacional en certificación de competencias TIC

11.409.855 Candidatos ECDL / ICDL

41 Idiomas

148 Países

24.000 Centros autorizados

45 millones de exámenes