

Carlos, arriba en el centro, junto a profesores del programa de doctorado en el B. Thomas Golisano College of Computing and Information Sciences.

“LA ‘HUMANIZACIÓN’ DE LOS DATOS: AHORA DESCRIBEN CÓMO SOMOS”



Carlos Rivero Osuna –Sevilla, 1982- es licenciado en Ingeniería informática por la Universidad de Sevilla y doctorado, en el año 2012, en el departamento de Lenguajes y Sistemas de la misma institución. Tras su paso por la Universidad de Idaho, Estados Unidos, en 2015 comienza su labor como docente e investigador en el departamento de Ciencias de la Computación del Instituto Tecnológico de Rochester, cerca de Nueva York, donde desarrolla, entre otras, diferentes técnicas destinadas a facilitar, tanto para el docente como el alumnado, la enseñanza y el aprendizaje de un área, hoy día en auge dentro de la Informática y otras disciplinas científicas, como es la programación.

Fuente: José T. del Pozo | Asesoría científica: Carlos Rivero Osuna

Trabaja como profesor e investigador en el departamento de Ciencias de la Computación del Instituto Tecnológico de Rochester –RIT-, ¿cuál ha sido su camino hasta llegar a esta institución?

Durante mi doctorado visité a los profesores Alberto Pan en la Universidad de A Coruña, Paolo Papotti, en la Universidad de Roma Tre, Italia, y Christian Bizer, en la Universidad Libre de Berlín, Alemania. Esta experiencia, unida a la finalización de mi doctorado, me llevó a trabajar, entre 2013 y 2015, como *postdoc* en la Universidad de Idaho. Cuando finalicé esta etapa, comencé a buscar nuevas opciones y tuve la oportunidad de encontrar mi sitio en el RIT. No obstante, antes de la Universidad, he participado como ingeniero de software en la creación de aplicaciones web en diferentes ámbitos, como la electricidad, los supermercados o la construcción para empresas como ‘Endesa’, ‘Carrefour’ y ‘Sando’.

Para las personas que leen esta entrevista y se plantean la búsqueda de oportunidades laborales en el mercado internacional, ¿puede describirnos un poco más su puesto y cómo es su evolución a lo largo del tiempo?

En concreto, hoy día, y durante los próximos seis años, trabajo como *Assistant profesor*, un rango académico empleado en algunos países como Estados Unidos o Canadá. Posteriormente, en agosto de 2021, tendré mi primera revisión para ascender a *Associate professor* –equivalente a profesor titular en España-, aunque el próximo curso, coincidiendo con el ecuador de mi primera etapa, me someteré a una observación intermedia donde mis supervisores me exponen qué tal voy: analizan investigación, docencia y servicios a la institución.

Sus comienzos estuvieron relacionados con la *deep web* o *web profunda*, hasta especializarse en la representación de programas mediante la teoría de grafos. Puede explicarnos ambos conceptos...

La *web profunda* o invisible podemos entenderla como la parte de Internet que contiene información y páginas web –como bases de datos o revistas académicas, por ejemplo- que no están indexadas en ninguno de los buscadores existentes hoy día, como son Google o Yahoo. De hecho, fíjese, la única manera de que esta zona ‘oculta’ de la red dejase de existir es que algún motor de búsqueda lograra indexar todas las web que coexisten en la actualidad.

Generalmente, ¿puede llegar a confundirse este concepto con la también conocida *dark net* o *web oscura*?

Efectivamente. La *dark net* es solo una pequeña parte de la *deep web* donde todo es anónimo y está cifrado, recibiendo su nombre por el tipo de información que alberga: pornografía o terrorismo, por ejemplo.

¿Y cómo llega a la teoría de grafos?

Simplificando, un grafo se puede representar como un conjunto de puntos, a los que denominamos nodos o vértices, unidos por líneas que son las aristas. De este modo, los mencionados grafos permiten estudiar las interrelaciones entre unidades que ya se encuentran interaccionando entre ellas. Por ejemplo, interpretados de la forma correcta, pueden proporcionar información relevante sobre mapas, circuitos electrónicos o mejorar la búsqueda de contenido en motores como Google o Yahoo, un ámbito, este último, que nos llamó mucho la atención y que decidimos investigar.

Tras su paso por la Universidad de Idaho decide embarcarse en un nuevo ‘mundo’ para usted, la aplicación de los grafos a la representación de programas, ¿en qué consiste?

La idea primaria partió de mi mentor en Idaho, ya que tenía un artículo donde describía que, mediante grafos, se podía mostrar la semántica de los programas informáticos. Una vez en Rochester, decidimos explotar esa idea, investigando y enviando una propuesta a la Fundación Nacional de Ciencia - *National Science Foundation*-, que fue rechazada, pero que obtuvo el sello de “competitiva”, una muy buena señal que nos invitó a seguir trabajando.

Los grafos proporcionan información relevante sobre mapas, circuitos electrónicos o mejoran la búsqueda de contenido en motores como Google o Yahoo.

Y en ello están...

Exacto. Ya nos han aceptado la participación en la, en mi opinión, tercera mejor conferencia sobre bases de datos (data management, en inglés) a nivel mundial como es la *IEEE International Conference on Data Engineering*, que celebra su vigésimo tercera edición el próximo mes de abril en San Diego, California.

¿Cuál es la principal innovación de esta línea de trabajo?

En Estados Unidos cada vez más estudiantes, de diferentes áreas de la ciencia, no solo ingeniería informática, desarrollan el curso de iniciación a la programación de forma *online* o presencial al comienzo de su carrera universitaria. De esta manera, el alumnado necesita una retroalimentación sobre las tareas

desarrolladas en la plataforma que hoy día no está recibiendo de forma puntual, ya que para los docentes es imposible resolver, al mismo tiempo, todas las dudas o exigencias que se les presentan. Tenga en cuenta que hablamos de cientos de miles de alumnos.

Y es aquí donde entra en juego, de nuevo, la teoría de grafos.

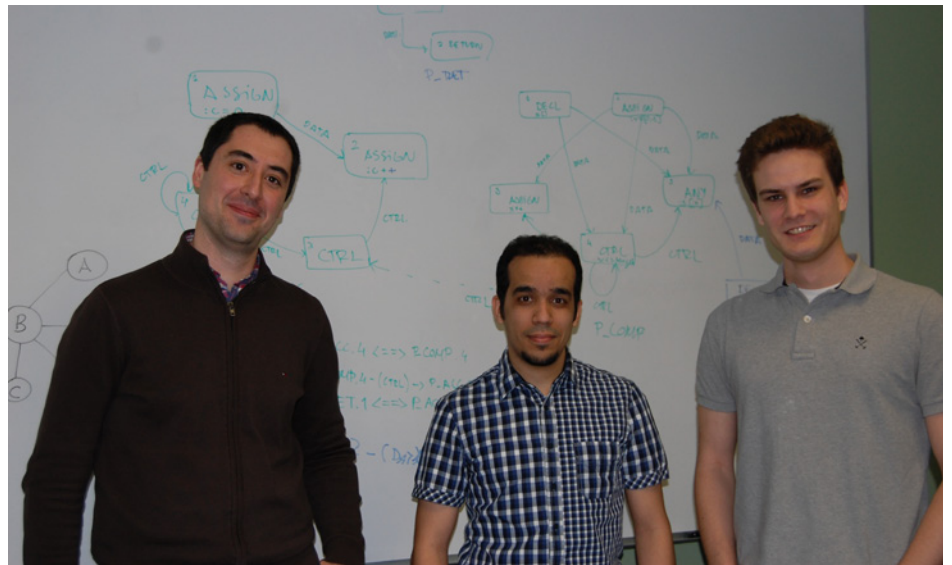
Eso es. Mediante grafos, estamos siendo capaces de dar el *feedback* correcto sobre casi diez millones de soluciones posibles a los estudiantes empleando, de media, aproximadamente 0.03 segundos.

Carlos Rivera trabaja en diferentes técnicas destinadas a facilitar, tanto para el docente como el alumnado, la enseñanza y el aprendizaje de la programación.

¿Cuáles son sus posibles aplicaciones prácticas?

El producto final, una vez detectados los patrones, es el desarrollo de un software que facilite la interacción del alumno con la plataforma –en la que se imparte el curso– y donde el profesor puede tener una visión completa de lo que está ocurriendo. Por tanto, el alumnado recibe, en tiempo real, información sobre sus errores o aciertos y el docente, un contenido útil sobre su propia metodología de enseñanza.

Igualmente, para la comunidad sorda, que aquí es muy extensa, se trata de una técnica especialmente útil ya que, hasta la fecha, aprenden el material vía intérprete. Sin embargo, mediante grafos, ellos tendrían directamente acceso al *feedback*, ayudando de este modo al interlocutor del proceso y facilitando el aprendizaje.



Carlos Rivera, a la izquierda, junto a Wilberto Núñez Pichardo y Victor Marín Aguilar.

Partiendo de las diferencias existentes, ¿son resultados o metodologías extrapolables al contexto español?

En principio, sí, aunque interfieren, como señala, circunstancias culturales que habría que estudiar en profundidad. Nosotros trabajamos en el aspecto tecnológico y si éste funciona habría que analizar cómo hacerlo, ya que el lenguaje, la metodología y, en definitiva, los patrones, son muy distintos.

Carlos Rivera: “El análisis de ingentes cantidades de datos describen los perfiles de la sociedad, cómo somos”.

¿Por dónde pasa el futuro de la informática?

La inteligencia artificial traducida en ordenadores inteligentes – más que los humanos–, algo que, según los expertos, puede ocurrir en aproximadamente diez o quince años. Esto dará lugar a un escenario donde parece obligatorio que las computadoras portasen, en su interior, algún tipo de razonamiento ético.

Por último: ¿qué tipo de información nos aporta la ciencia de los datos?

Desde hace algunos años, empresas como Facebook o Google perciben que sus necesidades han cambiado. Un ejemplo: ¿sabe qué es lo que ocurre durante un minuto de Internet en todo el mundo? Netflix y Youtube ocupan el aproximadamente el 50 por ciento del tiempo. Dentro de estos porcentajes existe mucha heterogeneidad –tipos de usuarios, ámbitos de acción, webs o vídeos, por ejemplo–, es decir, ingentes cantidades de datos que describen cómo somos –se han humanizado– donde el reto, para los especialistas, es extraer información útil: hablamos de detectar patrones de fraude, tendencias personalizadas de compra, de películas o sugerencias sobre productos o viajes a países que, muy probablemente, le gustará visitar. |