

MEMORIA LIBRE: "gpyVozControl", Ingeniería Trascendental y Libre

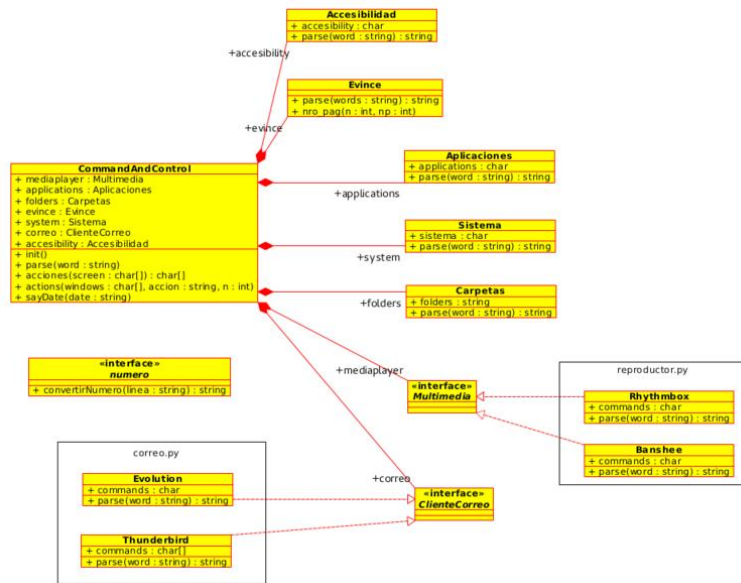


Categoría: [Memoria Libre](#)

Publicado el Viernes, 29 Enero 2016 11:03

Escrito por Jacinto Dávila

Visto: 246



Jorge Enrique Ortega está, por estos días, recibiendo su título de Ingeniero de Sistemas de la Universidad de Los Andes en la República Bolivariana de Venezuela. Lleva ya varios años trabajando como analista mayor para el Distrito Socialista Tecnológico, DST, de PDVSA en Mérida. Con el apoyo de la Empresa, atendió su carrera y completó una tesis en la que desarrolló una pieza de software para controlar cualquier computador personal estándar con comandos de voz. A sus 37 años, Jorge es un joven computista brillante, emprendedor, proactivo, con un refinado sentido del humor y un don especial para comunicar y demostrar herramientas de tecnologías de la información. Un solo detalle merece consideración respetuosa: Jorge no puede ver. Perdió la vista, primero en un ojo, luego en el otro, a causa de una enfermedad sistémica y mientras realizaba estudios de pregrado en Ingeniería Química en esa misma Universidad.

La historia de resiliencia¹ y serendipia² que siguió a esos eventos es un relato extraordinario que debe ser contado por su actor/autor principal. Permitánnos acá contar, solamente, el mas reciente episodio: el Proyecto de Grado en Ingeniería. gpyVozControl, nombre técnico del software desarrollado en el proyecto, es un pequeño módulo de software que capta sonidos a través del micrófono conectado a la máquina, reconoce palabras y frases, las procesa y, cuando entiende que son órdenes de su listado predefinido, las ejecuta (es decir, hace que el computador realice la acción). Por esta vía, un usuario o usuaria puede "abrir" el editor de textos, el reproductor de música o una presentación y operar sobre esas aplicaciones usando solamente su voz, si así lo desea. Más aún, puede "manipular" el ambiente de ventanas gráficas (GNOME, para quienes pregunten) que, por cierto, es invisible para Jorge, (es decir, puede abrir, maximizar, minimizar y cerrar ventanas) sin tocar el bendito ratón que suele ser tan incómodo para algunos adultos. Pero la interacción no es exclusiva. Es también posible seguir interactuando con el computador en las formas habituales. Si no pueden creer que eso sea posible, les invito a ver la presentación del proyecto que es, además y al mismo tiempo, una demostración del sistema gpyVozControl³

gpyVozControl, con todo y su extraordinaria utilidad no es, sin embargo, el mayor logro de Jorge en esta tesis. El mayor logro, para él y para todas y todos nosotros, es su experiencia y su sentido trascendental. Jorge se propuso ese desarrollo como parte de un ejercicio de investigación sobre plataformas para procesamiento de voz. Para ello, trató de estudiar los desarrollos previos y constató que su única oportunidad práctica para avanzar era aprovechar desarrollos libres (en software libre), que le permitieron a él (y le permitirán a cualquier otro) conocer en detalle el funcionamiento del sistema y adecuarlo en la medida en que lo necesite con total libertad. Como dice en su tesis "En este punto nos encontramos con dos aspectos de esa realidad. Por un lado, muy pocos desarrolladores o arquitectos de software incorporan en sus sistemas y aplicaciones alguna plataforma de accesibilidad que las personas con alguna discapacidad puedan usar libremente. Por el otro lado se tiene el alto costo de las licencias privativas de algunas aplicaciones con este tipo de ayudas tecnológicas. Es importante resaltar que cuando se hace referencia a 'libre', no quiere decir que es 'gratis', sino que los fuentes puedan estar disponibles para su revisión, distribución, uso y modificación, así puedan ser adaptados de acuerdo a cada circunstancia. Esto las hace particularmente valiosas para atender discapacidades específicas y en lugares de mundo con menos recursos tecnológicos."(Ortega, J. 2016). Establecido eso, Jorge debió estudiar cuidadosamente sendas APIs (ref)⁴ de proyectos libres de reconocimiento de voz que han venido evolucionando lentamente, en algunos aspectos con claros intereses comerciales, a partir de una comprensión teórica profunda. Jorge no solo programó su software, sino que entrenó dos máquinas de aprendizaje automático diferentes (Julius HTK y CMUSphinx), ambas basadas en modelos ocultos de Markov (un formalismo matemático para describir máquinas estocásticas auto-adaptables), para generar modelos acústicos y

lingüísticos alternativos del lenguaje que hablamos por acá, aplicado al control del computador. Con la ayuda de una red de amigas y amigos que participaron en la campaña Dona tu Voz⁵, Jorge compiló muestras de voces (con las lecturas de entre 125 y 250 frases en Español cuidadosamente seleccionadas, cada una) de 34 personas, 20 hombres y 14 mujeres. Con ese pequeño corpus de audio, de 5400 muestras, Jorge generó sendos motores de reconocimiento del habla que, en una primer evaluación sistemática, alcanzan a reconocer hasta un 80% de las palabras que "escucha" la máquina.

¿Cómo es posible que un joven invidente pueda hacer todas estas cosas?. Uno, viejo ingeniero, está tentado a apuntar a un empoderante duro: el computador. La máquina lee para Jorge (quien se ha resistido a aprender Braille) y con la creciente disponibilidad de programas y servicios "accesibles" (es decir, que consideran como servir a usuarios y usuarias con alguna discapacidad) será cada vez más fácil para él y para nosotros percibir por medio de otros sentidos. Jorge entiende perfectamente el concepto de accesibilidad y es un activo promotor de esa clase de andamiaje⁶. De hecho, este proyecto es un ejemplo de un esfuerzo por hacer que la máquina sea "más accesible". Noten, sin embargo, que Jorge no estaba pensando en un problema de accesibilidad visual. El puede operar la máquina fantásticamente con el teclado y los audífonos. Ha querido hacer una contribución para otros tipos de discapacidades. Es un gesto generoso que, disculpen Uds si subo el tono, tenemos que destacar y agradecer emocionados y conmovidos. Es un ejemplo de Ingeniería Trascendental. Así que la preferencia hacia el empoderante duro debe hacer lugar a un empoderante blando: Ha sido la generosidad el motor poderoso de esta iniciativa. Primero, el amor de la madre, la Señora Velquiz Josefina, quien lo ha acompañado en todo y a quien no podremos hacerle justicia en un escrito; Pero también la generosidad de toda la familia, con un hermano dispuesto a prestar su ojos y sus manos cuando el terco aprendiz lo requiere y el amor de una pareja que lo llena de dulzura; la generosidad y la confianza de una empresa pública que, hasta donde sé, es la primera que se atreve a contratar a un programador que no puede ver (Gracias Nerissa Aguilera y al proyecto Mundo Accesible, por abrir esas posibilidades). Sin hacer de esto una lista exhaustiva (disculpen los faltantes), permitan señalar la generosidad de una comunidad universitaria. Una Universidad que, sumergida en patrañas y contradicciones de toda índole, recibió a este joven y cuidó de él aún cuando esto significó, para una treintena de profesores, adaptar sobre la marcha sus cursos y hacerlos un tanto más accesibles. Es una muestra preciosa de todo lo que podemos hacer. Y si creen que nos hará falta la providencia y la buena fortuna, permitan agradecerle lo que ya ha hecho por Jorge y por nosotros: Jorge cuenta que haber perdido la vista después de aprender el cálculo infinitesimal: límites, derivadas e integrales (un formalismo matemático original de los maestros Newton y Leibniz y que se ha convertido en el núcleo básico de la matemática de las ciencias y las ingenierías) fue fundamental para avanzar en su carrera. La forma en la que se presenta esa matemática en el medio electrónico suele incluir imágenes (con "dibujos" de las fórmulas) que son inaccesibles (la máquina no las puede leer). Aprenderla sin poder ver es mucho más difícil. Pero es algo que podemos corregir y lo haremos.

Remembranza

Jorge ha tenido varios gestos de respeto hacia sus profesores y profesoras y compañeras y compañeros en su alma mater. Permitan honrar solo dos de ellos. En la tesis, a pesar de su consciencia de la accesibilidad (y de las recomendaciones de uno de sus tutores) insistió en insertar una imagen no accesible, la misma que, con su permiso, acompaña a este texto. Lo hizo para cumplir con lo que suele ser una forma tradicional de mostrar la arquitectura del sistema desarrollado. Jorge no puede ver esa imagen (hay formas que le permitirían escucharla y que todavía requieren mayor elaboración) pero, doy fe responsable, la entiende mejor que nadie. Que quede como un recordatorio de lo que tenemos pendiente para empoderar a estos espíritus libres. El otro gesto es algo que dijo al final de la defensa de su proyecto de grado⁷ "le agradezco a los profesores [y profesoras] porque no me subestimaron, ... tuve que sudar para poder pasar ... porque presentaba el mismo parcial que todos [los demás] ... fueron estrictos ... al momento de explicarme algo aparte me lo explicaban, [pero] al momento de decirme que estaba malo, me decían que estaba malo.. se nota que aprendí realmente lo que querían enseñarme ... y eso lo valoro muchísimo". Que esto nos quede como un recordatorio sutil del origen del valor. Gracias @joenco_

Referencias:

- 1.- https://es.wikipedia.org/wiki/Resiliencia_%28psicolog%C3%ADa%29
- 2.- <https://es.wikipedia.org/wiki/Serendipia>
- 3.- <https://youtu.be/MzteJI9djm4> Presentación de gpyVozControl (con Jorge Ortega) en YouTube
- 4.- <http://blog.mavenlink.com/why-leveraging-an-api-matters-for-your-business> Por qué importan la APIs
- 5.- <http://jacinto-davila.blogspot.com/2015/11/campana-dona-tu-voz.html> Dona Tu Voz
- 6.- <http://www.rebellion.org/noticia.php?id=84583> Empoderamiento, Andamiaje y Libertad
- 7.- http://nux.ula.ve/abiertos/jorge/presentacion-gpyVozControl-de-Jorge-Ortega-160111_001.mp3 Audio

Ortega, J. (2016). Sistema de órdenes de voz en software libre. Escuela de Ingeniería de Sistemas. Universidad de Los Andes. Por publicar.