

Luciana Benotti: “El extractivismo de datos no nos permite ser dueños de nuestra propia información”

Por: Nadia Luna. 10/05/2022

En la década del 80, eran muy pocos los y las niñas argentinas que tenían la suerte de tener una computadora en sus hogares. Luciana Benotti era una de ellas. En 1985, su papá llevó una computadora a la casa pero los sistemas operativos no permitían hacer muchas cosas. “No era nada divertido, la verdad. No había cosas pre-armadas. Entonces, lo que yo hacía era copiar y modificar códigos para poder programar mis propios jueguitos”, recuerda hoy Benotti, doctora en Informática e investigadora del CONICET en la Universidad Nacional de Córdoba (UNC).

Sin embargo, ese entusiasmo por la computación que adquirió de chica se disipó un poco cuando empezó a cursar la carrera. “No me sentía muy incluida en el área, las mujeres siempre éramos minoría. Y la diferencia se notó más cuando empecé a trabajar porque era un entorno muy masculinizado y estresante”, cuenta. Pero la motivación resurgió cuando le salió una beca para hacer una maestría en Italia y España. Ahí conoció el área de procesamiento de lenguaje natural, dentro de lo que se denomina inteligencia artificial (IA), y le interesó mucho ya que reunía otra pasión que tenía de chica: aprender nuevos lenguajes. “Enseñarle a ‘hablar’ a las computadoras es algo que me pareció muy desafiante”, dice.

Hace unos días, Benotti se convirtió en la primera latinoamericana en obtener el cargo de *chair* (cargo principal) de la Asociación Norteamericana de Lingüística Computacional.

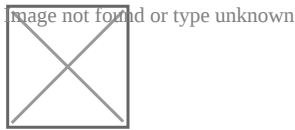
Dentro del área de lingüística computacional, ¿en qué se especializó?

Mi área de investigación durante el doctorado abordó los sistemas de diálogo, que son conocidos también como chatbots. Yo trabajé con sistemas situados en un entorno físico simulado. Esto es, como si fuera un chatbot que está dentro de un robot y, para interactuar, necesita saber qué cosas tiene alrededor. Esto combina procesamiento de lenguaje y de imágenes. Los sistemas en los que trabajábamos eran sistemas con los que un humano podía hablar, por ejemplo, con personajes de un entorno similar a un videojuego. También trabajé esa línea en Estados Unidos.

Yo estaba en un proyecto de realidad aumentada destinado a tratar el estrés postraumático en personas del ejército norteamericano. Este es un tipo de realidad virtual que también produce estímulos físicos.

¿Sigue trabajando en esas líneas?

Sí, pero con menos recursos porque acá tenemos el problema de que somos pocos los que investigamos sobre esto. Me presenté a muchas convocatorias y cuesta que salgan porque la línea de corte en las distintas áreas se define por la cantidad de proyectos presentados. Como en computación somos pocos, los que obtienen financiamiento son menos todavía. De todos modos, sigo trabajando en la aplicación de técnicas de procesamiento de lenguaje y de imágenes, no tanto con simulaciones porque eso requiere más poder computacional, sino con imágenes estáticas. Por ejemplo, un estudiante de doctorado que dirijo está en un proyecto que busca procesar fotos y preguntas de una persona no vidente. Preguntas como “esta remera que me voy a poner, ¿de qué color es?”. Saca una foto, la manda a una aplicación y obtiene la respuesta. En algunos países, hay gente detrás de esas aplicaciones. Por eso, hay un área de investigación que busca contestar algunas de esas preguntas de forma automática.



“En los 70, las mujeres eran mayoría, pero el área de computación no estaba muy desarrollada. Incluso, antes de que estuvieran las computadoras, los cálculos se hacían a mano y eran hechos por mujeres porque no eran tareas agradables ni representaban poder”, dice Benotti.

¿En qué aspectos de la vida cotidiana podemos encontrar sistemas de inteligencia artificial?

Lo que se conoce como inteligencia artificial es el área de aprendizaje automatizado por computadora. Básicamente, es la construcción de algoritmos que permiten que el comportamiento de un programa no dependa de un programador, sino de los datos que le fueron dados. Las aplicaciones de esto son muy variadas. Por ejemplo, cuando vas al cajero y te dice que tenés un crédito preaprobado por un monto determinado. Es muy probable que, para determinar el monto, haya un algoritmo de aprendizaje supervisado que se encarga de estimar cuál es la probabilidad de que la

persona devuelva una cierta cantidad de plata. Otro ejemplo es la búsqueda de personas en las fotos que sacás con el celular, por reconocimiento facial. Esto usa aprendizaje automático, que forma parte del procesamiento de imágenes. La parte de procesamiento de lenguaje natural también usa aprendizaje automático. Cuando empezás a escribir algo en un buscador y te sugiere opciones, el sistema usa una técnica que se llama modelos de lenguaje para hacer las sugerencias. Estos modelos se usan también para reconocimiento del habla y han mejorado mucho en los últimos años.

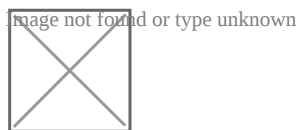
Imagino que todavía deben quedar grandes desafíos en el área a la hora de intentar reproducir, de alguna manera, el pensamiento humano.

Sí, hay algoritmos que simplifican muchas tareas pero que también cometen errores que para nosotros serían muy “tontos”. Por ejemplo, con la pandemia, surgieron muchas aplicaciones que trataban de usar datos para hacer predicciones. Hubo un proyecto que se realizó en hospitales de Estados Unidos donde un sistema de IA, cruzando distintos datos del paciente como historia clínica y estudios, trataba de predecir quiénes de los que daban positivo para COVID necesitaban ser internados. Resultó que uno de los parámetros que usaba el sistema para predecir era el tipo de letra de la historia clínica (si era times new roman, arial u otra), un criterio que no tiene sentido para una persona. Estos sistemas tienen un problema grave para darse cuenta qué es relevante y qué no, porque no entienden de relaciones causales. Sí entienden de correlaciones: si ven que algo pasa frecuentemente junto con otra cosa, entienden que hay una relación.

Algo que se discute mucho en el área y que tiene que ver con lo que usted menciona son los dilemas éticos que pueden producirse en torno a estos sistemas. Por ejemplo, si cometen un error grave que genera daños a una persona, ¿sería culpa del desarrollador? ¿Pueden los sistemas llegar a tener en cuenta consideraciones éticas en sus decisiones?

Es algo que recién ahora se está reconociendo como un problema. El año pasado, se agregó un requisito por el que todos los artículos científicos en esta área deben incluir una reflexión sobre el impacto social y los riesgos relacionados con las cuestiones éticas que puede tener una determinada tecnología. Son cosas que se están empezando a discutir ahora y todavía hay bastante resistencia de algunas partes de la comunidad, que consideran que eso no es su problema. Como diciendo, “yo desarrollé esto y después dependerá de cómo lo usen”, y eso es lavarse las

manos.



Benotti fue designada chair de la Asociación Norteamericana de Lingüística Computacional (NAACL) y es la primera latinoamericana en obtener este puesto de dirección. “Para mí, significa poder visibilizar que hacen falta más minorías involucradas en el desarrollo de estas tecnologías”, dice.

¿Cómo está el área de investigación en la Argentina?

Yo volví al país hace 10 años, como científica repatriada por el programa Raíces. En ese entonces, éramos pocos los que investigábamos sobre estos temas pero cada vez somos menos. Esto tiene que ver con que las grandes empresas se chupan a los estudiantes porque pagan mejor, algunos ni llegan a terminar la carrera. También muchos eligen ir a trabajar al exterior por el mismo motivo. Es muy difícil para nosotros retener estudiantes para que se queden a hacer doctorados porque el incentivo económico que hay desde la industria o desde áreas como las criptomonedas es mucho más grande. Y no solo se van estudiantes, también se van profesores. Nosotros perdimos a un tercio de nuestros profesores durante la pandemia.

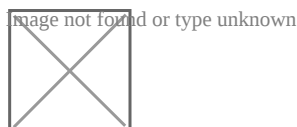
Hace unos días, fue designada chair de la Asociación Norteamericana de Lingüística Computacional (NAACL, por sus siglas en inglés) y es la primera latinoamericana en obtener este puesto de dirección. ¿Qué significa para usted este reconocimiento y qué objetivos busca alcanzar desde este lugar?

Para mí, significa poder visibilizar que hacen falta más minorías involucradas en el desarrollo de estas tecnologías. Todo el dilema de las cuestiones éticas es cada vez más fuerte porque cada vez se usan más y las ciencias que tienen un impacto social tan grande necesitan representación de las distintas partes que componen la sociedad para entender cómo afecta a cada una. Yo había participado de esta asociación hace un tiempo, desde mi lugar de estudiante europea privilegiada a la que me mandaban a todas las conferencias. Después dejé de hacerlo por una cuestión económica y, cuando empezó la virtualidad, pude volver a participar. Ahí vi que éramos muy pocos los latinoamericanos que participábamos, pero fuimos

armando una comunidad y tratamos de atraer más gente. Incluso, de otras minorías, como africanos o de la comunidad queer, con la idea de una mayor diversidad en la gente que participa de los debates y publicaciones en el área. No es sencillo, sobre todo porque en estas asociaciones científicas hay mucha influencia de las grandes empresas tecnológicas, como Google y Facebook. Así que básicamente es, desde mi pequeño lugar, tratar de contrarrestar un poco esa intención de las grandes empresas de pasar por encima de todo, sin medir el impacto social. Por suerte, vamos consiguiendo algunas cosas, como la declaración de impacto social que comentaba antes y que a unos cuantos les molesta bastante.

Hablando de minorías, la informática es una de las disciplinas más masculinizadas, pero tiene la particularidad de que no siempre fue así. Un [informe de la Fundación Sadosky](#) señala que, en los 70, las mujeres eran mayoría. ¿Cómo está la situación actualmente?

Sigue estando muy masculinizada. En los 70, las mujeres eran mayoría, pero el área de computación no estaba muy desarrollada. Incluso, antes de que estuvieran las computadoras, los cálculos se hacían a mano y eran hechos por mujeres porque no eran tareas agradables ni representaban poder. Como eso fue cambiando y el área pasó a tener otra relevancia, se masculinizó. Hace poco, con unas compañeras de la UNC, hicimos un pequeño relevamiento, con una muestra de 300 estudiantes, donde vimos que las mujeres que ingresan ahora (que siguen siendo menos del 20%) son más grandes. Lo que nos decían es que recién se enteraban de la existencia de la carrera después de haber empezado a estudiar otra cosa. Por otro lado, dentro del área hay algunos matices. En lingüística computacional, por ejemplo, hay un poco más de mujeres que en el área de computación en general. Tendrá un 30% aproximadamente, porque hay estudiantes que vienen de la lingüística, donde hay más mujeres.



Según Benotti, se ha encontrado que enseñar programación a temprana edad puede tener un impacto positivo en reducir la brecha de género en el área.

En los últimos años, se empezó a abordar más la importancia de enseñar programación en las escuelas. ¿Esto contribuye a que las niñas empiecen a familiarizarse más con la computación desde chicas?

Sí, de hecho, yo dirigí una tesis de doctorado en esta área y tuvimos algunos datos. Mi tesista, Marcos Gómez, siguió a un grupo primaria durante tres años y una de las cosas que vio es que cuando las nenas aprendían programación de chicas, si después querías enseñarles otras cosas de computación, no mostraban diferencias con sus compañeros varones en cómo aprendían los nuevos temas. En cambio, en otra escuela en la que no enseñaban programación surgía una brecha en el aprendizaje entre varones y nenas. Hay distintas hipótesis de por qué sucede esto, una tiene que ver con el tipo de juegos a los que juegan, pero se veía que enseñar programación a temprana edad tenía un impacto positivo que reducía la brecha.

¿Y cómo se trasladan los sesgos de género al producir tecnología?

Esto se ve de diversas maneras. Por ejemplo, en los sistemas de reconocimiento del habla. Por mucho tiempo, el reconocimiento andaba mejor con voces masculinas que con femeninas. Hace poco también salió la noticia de que el reconocimiento facial en mujeres negras era bastante peor que el de hombres blancos. En general, cuando se trata de grupos de menos poder o minorías, siempre encontrás una tasa de error más alta.

¿Cuáles son sus principales desafíos para los próximos años en materia de investigación?

Tengo varios, pero uno que me representa un gran desafío es un proyecto que estoy haciendo con un grupo de personas que forman el Programa PRODA y asesoran en huertas urbanas, en Neuquén. Hay una cuestión muy familiar también porque mi hermana está en ese grupo y nosotras hemos hecho huerta durante muchos años. Con la pandemia, aumentó tanto el interés de la comunidad en el tema que superó su capacidad para responder a consultas. Entonces estoy trabajando en un procesamiento automático de imágenes combinado con lenguaje que pueda responder algunas consultas de forma automática. Es un desafío lindo porque tiene una aplicación directa. La transferencia de tecnología sigue siendo una deuda grande del sistema científico en estas áreas.

¿Por qué es importante que la Argentina invierta recursos en investigación en el área de ciencias de la computación y en IA?

Para mí, es muy importante que tengamos soberanía tecnológica. Cuando una persona dice “tengo las cosas en la nube” no se pone a pensar que las nubes son computadoras que están en Estados Unidos. La pérdida de privacidad que tenemos en este momento en nuestro país es grande, incluso hay información oficial guardada en “la nube” de Estados Unidos y esto es algo que me preocupa. La tecnología es un derecho al que todos deberíamos acceder, pero existe lo que nosotros llamamos extractivismo de datos, que nos nos permite ser dueños de nuestra propia información. Si un día esas empresas a las que confiamos los datos nos quieren cobrar para acceder, lo pueden hacer. La verdad es que si uno no tiene ciertos conocimientos en el área, no se pone a pensar en estas cosas pero la dependencia tecnológica que tenemos hoy es enorme.

[LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ](#)

Fotografía: Hamartia

Fecha de creación

2022/05/10