

El software utiliza un modelo estadístico

Detectan plagio en traducciones

► Participa mexicano en método automático desarrollado en Valencia

Erika P. Bucio

Un equipo de la Universidad Politécnica de Valencia logró desarrollar un método capaz de detectar el plagio en traducciones después de realizar experimentos con multitud de textos, entre ellos artículos de la Wikipedia.

“Si detectar casos de plagio que vayan más allá de un simple copiar-pegar de textos en la misma lengua ya es difícil, en el caso de plagio traducido lo es todavía más”, explica Paolo Rosso, responsable del Laboratorio de Ingeniería en Lenguaje Natural (Lab NLE) de la UPV.

El método automático de detección de plagio ha sido desarrollado en el marco de la tesis doctoral del mexicano Alberto Barrón-Cedeño, egresado de la UNAM.

Sirve para casos en que la fuente original está, por ejemplo, en inglés y el documento sospechoso de plagio en español.

Se trata de un modelo estadístico que aprende cómo una frase en un idioma, por ejemplo el inglés, puede traducirse al español.

Una vez aprendidas las posibles traducciones puede medir la similitud entre dos fragmentos de texto, describe Rosso.

Sin embargo, ataja, la detección automática no pretende sustituir al ojo humano sino ser una herramienta al servicio de los expertos.

“De esta manera el experto estará en condiciones de poder analizar manualmente el presunto caso de plagio y así tomar la decisión al respecto”, precisa Rosso.

Barrón-Cedeño explica que para desarrollar el modelo de detección automática se hicieron pruebas con artículos de la Wikipedia, entre otro tipo de documentos.



► El mexicano Alberto Barrón-Cedeño junto a Paolo Rosso del LabNLE.

Policía textual

El modelo aprende cómo una frase en un idioma puede traducirse, y a partir de esa estadística mide la similitud entre dos fragmentos.

Qué es

► Un detector automático de plagio en traducciones.

Cómo funciona

► Mide la similitud entre la fuente original y el documento sospechoso escrito en una lengua distinta.

Quién lo creó

► Alberto Barrón-Cedeño, estudiante de doctorado, y Paolo Rosso, responsable del Laboratorio de Ingeniería en Lenguaje Natural (Lab NLE) de la Universidad Politécnica de Valencia.

“Nos hemos dado cuenta de que entre Wikipedias hay muchísimos artículos que se generan primero en una lengua y después se utilizan como base para generar otros, y eso parece ser muy común desde el inglés hacia otros idiomas.

“Es decir, las Wikipedias en alemán, en español o en chino tienen muchos artículos cuyo contenido se generó originalmente en inglés”, expone el estudiante de doctorado.

Ya existen herramientas comerciales para la detección de plagio en un mismo idioma como Turnitin y DOC Cop, pero no había nada, agrega, para el caso de las traducciones, con excepción

de un método desarrollado por la Universidad de Weimar —que se limita a los artículos de la Wikipedia en los idiomas en que se quiere trabajar.

El modelo estadístico desarrollado en la Universidad Politécnica de Valencia funciona no sólo con textos traducidos del inglés al español sino a otras lenguas como alemán, francés, italiano, polaco y holandés.

Pero podría trabajar con todos los idiomas siempre que se entrene con pares de traducciones en las lenguas en cuestión, añade Rosso.

De momento, el método para la detección automática de plagio traducido no está disponible en el

Impulsan campaña

Los actuales sistemas de detección de plagio fracasan cuando van más allá del sencillo “copiar-pegar” y se recurre a la paráfrasis para decir lo mismo con palabras distintas, un problema que el equipo de Laboratorio de Ingeniería en Lenguaje Natural (Lab NLE) se propone enfrentar.

Esa es la razón, dice el profesor Paolo Rosso, por la cual desde hace tres años organiza junto con la Universidad de Weimar en Alemania una competencia internacional, patrocinada por la empresa estadounidense Yahoo! Research.

El año pasado participaron en el concurso 18 equipos pero ninguno mexicano. En 2011, agrega Rosso, está prevista la participación del Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE) en Puebla.

La competencia se celebrará durante la Conferencia on Multilingual and Multimodal Information Access Evaluation (CLEF).

sitio de internet del Lab NLE.

En la página se pueden encontrar otros prototipos como el Stylus, de análisis de estilo, desarrollado por griegos y alemanes, que puede revelar si en un texto hay párrafos que no fueron escritos por el mismo autor.

El Laboratorio anima otra competencia dirigida a desarrollar sistemas de detección de vandalismo en la Wikipedia.

Aunque hay programas automáticos y personas que se dedican a depurar la información, se estima que un 7 por ciento de los artículos (en inglés) corresponden a actos vandálicos, precisa Rosso.

Ese vandalismo se refiere a “cualquier adición, supresión o cambio en el contenido con la intención deliberada de comprometer la integridad de Wikipedia”.

La competencia tendrá lugar del 19 al 22 de septiembre, en el marco de la Conferencia on Multilingual and Multimodal Information Access Evaluation (CLEF).



► Las obras que Roberto Cortázar expone en el Olmedo pertenecen a la Colección Pago en Especie de la SHCP.

Compilan arte de Cortázar

► El artista chiapaneco expone dos décadas de quehacer plástico en el Dolores Olmedo

Oscar Cid de León

Entre cuerpos concretos, pero de pronto deconstruidos por trazos que devienen, en cierta forma, en una suerte de intrusos, se encuentra el pintor y escultor Roberto Cortázar (Chiapas, 1962).

Bajo esa contexto, retine 20 años de trabajo a través de *Corpus et anima*, que inaugura hoy a las 11:30 horas en el Museo Dolores Olmedo.

En total se integran 22 piezas, entre óleos y esculturas, girando alrededor de la figura humana.

“Trabajo con la representación de la figura humana como centro de mi modo de componer”, explica en entrevista. “Parto de una pintura clásica, con una técnica clásica, óleo sobre tela, para después llegar a algo así como un clasicismo contemporáneo que hace que la representación rigurosa, o de un grado de dificultad intenso, derive en modos de expresividad más amplios”.

Celebra a Pago en Especie

Roberto Cortázar celebra la existencia del Programa Pago en Especie. “Ha ido de menos a más”, señala quien forma parte de él desde hace casi dos décadas.

Lo que resulta plausible, considera, no es sólo su éxito, sino la extensión, a través de él, de las labores de la SHCP a los terrenos del arte. “Lo cual es muy interesante, porque ése no es el trabajo de Hacienda”.

El artista observa que en los tiempos actuales impera la crítica hacia las acciones gubernamentales, pero hay virtudes del Estado que deben reconocerse. Pago en Especie es una de ellas, dice.

Después de una pausa, como buscando traducirse, añade: “Si un espectador pone atención, por ejemplo, de pronto verá figuras muy concretas y de repente rayas tiradas a lo largo de un espacio cuadro o de una escultura, pero en realidad esas rayas simplifican o rehacen el modo de representar una figura, y esto es porque su expresividad se va ampliando”.

Las piezas que dan cuenta de su quehacer, a través de *Corpus et anima*, pertenecen a la Colección Pago en Especie y Acervo Patrimonial que custodia la Secretaría de Hacienda y Crédito Público (SHCP), conformada por la dependencia a través del Programa Pago Especie, que permite a los creadores saldar sus impuestos con obra.

Cortázar está inscrito a él desde hace casi dos décadas.

“Es decir que en la muestra se verá una selección de mis pagos a lo largo de unos 20 o 15 años”.

Bajo la curaduría de Rafael Pérez y Pérez, Subdirector del Museo de la SHCP, Antiguo Palacio del Arzobispado, *Corpus et anima* despliega una revisión que muestra el desarrollo de su obra.

Al referirse a su quehacer, quien egresó de La Esmeralda, pondera su carácter clásico, su técnica forjada a la manera de la “vieja escuela”.

Lamenta que las escuelas de hoy no fomenten el aprendizaje de la plástica de forma integral, partiendo de cómo elaborar, por ejemplo, una pintura con los métodos del Siglo 15, hasta desembocar en las técnicas más actuales.

Eso, según dice, representa una deficiencia.

“Si estudias matemáticas, vas a estudiar desde el origen hasta el presente en una forma lineal y progresiva indubitable. Lo mismo si estudias astrofísica o cualquier materia lógica o de ciencias exactas, incluso de humanidades, pero en el arte no pasa así, sólo estudias, en términos realistas, el periodo que corresponde al momento que estás estudiando”.

Corpus et anima se presenta en la Galería Pago en Especie del Dolores Olmedo (Av. México 5843, La Noria, Xochimilco), destinada de forma exclusiva a la difusión del acervo conformado a través del programa.

Presentación de libro año académico 2011

El Colegio Nacional

LEOPOLDO GARCÍA-COLÍN
Miembro de El Colegio Nacional

LEONARDO DAGDUG
MICHAEL PICQUART
EDGAR VÁZQUEZ

Coordinadores

Mayo
Jueves 12

Tomo 1
Alejandro Fernández
Facultad de Medicina, UNAM

Tomo 2
Herminia Pasantes
División de Neurociencias
Instituto de Fisiología Celular, UNAM

Tomo 3
Juan Carlos Echeverría
Universidad Autónoma Metropolitana,
Iztapalapa

El evento se llevará a cabo en **El Aula Mayor de El Colegio Nacional**, Donceles 104, Centro Histórico, México, D.F.

10:30 horas
Transmisión: videoconferencias e internet

ENTRADA LIBRE

www.colegionacional.org.mx e-mail: contacto@colegionacional.org.mx
Tel. 5789-4330 Fax. 5702-1779

SUDOKU

Completa la cuadrícula de manera que cada línea, columna y caja de 3x3 tenga un dígito del 1 al 9 sin repetirse.

NIVEL BÁSICO

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3 | | 2 | | | 5 | 9 | 7 | 1 |
| | | | 4 | | | | | |
| | 6 | | | 3 | | 8 | 5 | 4 |
| 1 | 7 | | 3 | | | 5 | | |
| | | 9 | 6 | | 4 | 7 | | |
| | | 4 | | | 9 | | 3 | 6 |
| 6 | 9 | 3 | | 7 | | | 1 | |
| | | | | | 6 | | | |
| 8 | 1 | 7 | 9 | | | 2 | | 5 |

NIVEL MEDIO

| | | | | | | | | |
|---|---|---|---|--|---|---|---|---|
| | | 2 | 9 | | 4 | | 8 | |
| 9 | 3 | | 1 | | 8 | 6 | | 7 |
| | | | | | | | 9 | 4 |
| | | 9 | 4 | | | | | 8 |
| | | | 5 | | 9 | | | |
| 5 | | | | | 6 | 2 | | |
| 4 | 1 | | | | | | | |
| 6 | | 3 | 7 | | 2 | | 4 | 5 |
| | 9 | | 6 | | 3 | 8 | | |