

Fecha del CVA	Enero, 2025
---------------	-------------

Parte A. DATOS PERSONALES

Nombre	Patricia		
Apellidos	Jiménez		
Sexo (*)	Mujer	Fecha de nacimiento (dd/mm/yyyy)	[REDACTED]
Dirección email	patriciajimenez@us.es		
Open Researcher and Contributor ID (ORCID) (*)	0000-0001-5070-5904		

* datos obligatorios

A.1. Situación profesional actual

Puesto	Profesor Titular de Universidad
Fecha inicio	04/07/2023
Organismo/ Institución	Universidad de Sevilla
Departamento/ Centro	Departamento de Lenguajes y Sistemas Informáticos
País	España
Palabras clave	Extracción de Información, Integración de aplicaciones empresariales, aprendizaje no supervisado, métodos de ranking automáticos, validación y evaluación de datos

A.2. Situación profesional anterior (incluye interrupciones en la carrera investigadora, de acuerdo con el Art. 45. 2 b) de la convocatoria)

Periodo	Puesto/ Institución/ País / Motivo interrupción
04/07/2023 –	Profesor Titular Universidad / Universidad de Sevilla / España
28/08/2018 – 03/07/2023	Profesor Contratado Doctor / Universidad de Sevilla / España
28/07/2017 – 27/07/2018	Profesor Ayudante Doctor / Universidad de Sevilla / España
19/09/2016 – 27/07/2017	Profesor Sustituto Interino / Universidad de Sevilla / España
06/02/2016 – 05/05/2017	Beca de Ampliación de Estudios / Fundación Fidertia / España
17/02/2016 – 18/09/2016	Profesor Sustituto Interino / Universidad de Sevilla / España
06/12/2015 – 06/02/2016	Titulado Superior / Universidad de Sevilla / España
01/09/2011 – 05/12/2015	Becaria FPI / Universidad de Sevilla / España
25/02/2009 – 30/08/2011	PSI y Técnico Apoyo Investigación / Universidad de Huelva / España

(Incorporar todas las filas que sean necesarias)

A.3. Formación Académica

Grado/Master/Tesis	Universidad/País	Año
Tesis. Enterprise Application Integration: New Approaches to Web Information Extraction	Universidad de Sevilla / España	2015
Máster. Máster Universitario en Tecnologías Informáticas Avanzadas. Universidad de	Universidad de Huelva / España	2009

Huelva		
Titulado superior. Ingeniería Informática	Universidad de Huelva / España	2009
Ingeniería Técnica en Informática de Gestión.	Universidad de Huelva / España	2005

(Incorporar todas las filas que sean necesarias)

Parte B. RESUMEN DEL CV (máx. 5000 caracteres, incluyendo espacios):

Patricia Jiménez inició su carrera universitaria en el curso 2008-09 en el Departamento de Tecnologías de la Información de la Universidad de Huelva, donde fue profesora sustituta interina. Desde 2009 hasta 2011 fue técnica de apoyo en el contexto del proyecto de excelencia de la Junta de Andalucía denominado IntegraWeb Wrapping (TIC-4100), donde su trabajo se centró en la implementación de varios prototipos de sistemas de extracción de información usando técnicas de inferencia de patrones. En 2011 consiguió una beca de Formación de Personal Investigador en el proyecto denominado Integración de Aplicaciones y Datos en la Web (TIN2010-21744-C02-01). Ha sido en el contexto de este último proyecto en el que ha desarrollado su tesis doctoral, cuya defensa tuvo lugar en octubre de 2015. En su tesis, Patricia ha desarrollado un sistema de aprendizaje lógico inductivo de reglas de extracción de información para documentos web semi-estructurados llamado Tango; a partir de esta base y de las conclusiones obtenidas ha desarrollado un sistema de aprendizaje proposicio-relacional llamado Roller, que mejora notablemente la eficiencia del primero; también ha desarrollado un método automatizado para realizar rankings de extractores de información. Posteriormente, ha participado en proyectos de transferencia tecnológica y ha codirigido una tesis doctoral defendida en diciembre de 2020, centrada en la extracción de información de tablas HTML mediante técnicas de Deep Learning y heurísticas.

En cuanto a investigación, cuenta con 13 publicaciones en revistas de impacto en Q1 y Q2 y una decena de artículos publicados en conferencias y ha participado en varios proyectos competitivos a nivel nacional y autonómico; también cuenta con varias estancias tanto predoctorales como postdoctorales, en las Universidades de Bristol (3 meses), Southern California (6 meses) y Queensland University of Technology (4 meses). En todas ellas trabajó en propuestas para extracción de información de documentos web y validación de los datos extraídos. En cuanto a docencia, lleva impartiendo clases durante los últimos 10 años en las Universidades de Huelva y Sevilla, en una gran variedad de asignaturas, cursos y titulaciones.

En la actualidad, Patricia Jiménez se encuentra desarrollando sus actividades de docencia e investigación en la Universidad de Sevilla, con la figura de profesor titular de universidad desde julio de 2023. Sus labores de investigación se centran en la extracción de información y en el desarrollo de herramientas de evaluación y validación de datos extraídos. Sus labores de docencia se centran principalmente la enseñanza de lenguajes de programación y en el diseño y desarrollo de sistemas de información web.

Parte C. LISTADO DE APORTACIONES MÁS RELEVANTES (últimos 10 años)- Pueden incluir publicaciones, datos, software, contratos o productos industriales, desarrollos clínicos, publicaciones en conferencias, etc. Si estas aportaciones tienen DOI, por favor inclúyalo.

C.1. Publicaciones más importantes en libros y revistas con “peer review” y en conferencias (ver instrucciones).

1. **On the Design of an Advanced Business Rule Engine.** Jiménez, Patricia, Corchuelo, Rafael. 2022. Journal of Software: Practice and Experience. JCR=2.028, SJR = 0.437. Q2. (In press).
2. **On Validating Web Information Extraction Proposals.** Jiménez, Patricia, Corchuelo, Rafael. 2022. Experts Systems With Applications. JCR = 6,954, SJR = 1.368. Q1. DOI: <https://doi.org/10.1016/j.eswa.2022.116700>.

3. **An Experimental Study of Neural Approaches to Multi-Hop Inference in Question Answering.** Jiménez, Patricia, Corchuelo, Rafael. 2022. International Journal on Neural Systems (IJNS). JCR: 5.866, SJR: 1.376, Q1. <https://doi.org/10.1142/S0129065722500113>.
4. **A hybrid quantum approach to leveraging data from HTML tables.** Jiménez, Patricia, Roldán, Juan C., Corchuelo, Rafael. 2022. Knowledge and Information Systems. JCR: 2,822, SJR: 0.634, Q2. DOI: <https://doi.org/10.1007/s10115-021-01636-7>.
5. **On Exploring Data Lakes by Finding Compact, Isolated Clusters.** Jiménez, Patricia, Roldán, Juan C., Corchuelo, Rafael. 2022. Information Sciences. JCR=6.795, SJR = 1.524. Q1. DOI: <https://doi.org/10.1016/j.ins.2021.12.045>.
6. **A coral-reef approach to extract information from HTML tables.** Jiménez, Patricia, Roldán, Juan C., Corchuelo, Rafael. 2022. Applied Soft Computing. JCR = 6,725, SJR = 1.290, Q1. DOI: <https://doi.org/10.1016/j.asoc.2021.107980>.
7. **A clustering approach to extract data from HTML tables.** Jiménez, Patricia, Roldán, Juan C., Corchuelo, Rafael. 2021. Information Processing & Management. JCR=6,222, SJR = 1.061. Q1. DOI: <https://doi.org/10.1016/j.ipm.2021.102683>.
8. **TOMATE: A heuristic-based approach to extract data from HTML tables.** Roldán, Juan C., Jiménez, Patricia; Corchuelo, Rafael. 2021. Information Sciences. JCR=6,795, SJR = 1.524. Q1. DOI: <https://doi.org/10.1016/j.ins.2021.04.087>.
9. **On Extracting Data from Tables that are Encoded using HTML.** Roldán, Juan C., Jiménez, Patricia; Corchuelo, Rafael. 2020. Knowledge-Based Systems. JCR=8,038, SJR = 1.587. Q1. DOI: <https://doi.org/10.1016/j.knsys.2019.105157>.
10. **On the synthesis of metadata tags for HTML files.** Jiménez, Patricia, Roldán, Juan C., O. Gallego, Fernando, Corchuelo, Rafael. 2020. Journal of Software: Practice and Experience. JCR=2.028, SJR = 0.437. Q2. DOI: <https://doi.org/10.1002/spe.2886>.
11. **ARIEX: Automated ranking of information extractors.** Jiménez, Patricia; Corchuelo, Rafael; Sleiman, Hassan A. 2016. Knowledge-Based Systems. 93: 84-108. JCR=4,529, SJR = 1.709. Q1. DOI: <https://doi.org/10.1016/j.knsys.2015.11.004>.
12. **On learning Web Information Extraction Rules with TANGO.** Jiménez, Patricia; Corchuelo, Rafael. 2016. Information Systems. JCR=2,777, SJR=0.896. Q2. DOI: <https://doi.org/10.1016/j.is.2016.05.003>.
13. **Roller: A novel approach to web information extraction.** Jiménez, Patricia; Corchuelo, Rafael. 2016. Knowledge and Information Systems. 49, 197–241. JCR=2,004, SJR = 0.733. Q2. DOI: <https://doi.org/10.1007/s10115-016-0921-4>.

C.2. Congresos, indicando la modalidad de su participación (conferencia invitada, presentación oral, póster):

1. **Extracting web information using representation patterns.** Juan C. Roldán, Patricia Jiménez, Rafael Corchuelo. HotWeb '17: Proceedings of the fifth ACM/IEEE Workshop on Hot Topics in Web Systems and Technologies (HOTWEB'17). GGS: Work in Progress; CORE B.
2. **Feeding Software Agents with Web Information.** Patricia Jiménez, Hassan A. Sleiman, Rafael Corchuelo. International Conference on Practical Applications of Agents and Multi-Agents. (PAAMS'15). GGS: Work in Progress; CORE B.
3. **On Extracting Information from Semi-Structured Deep Web Documents.** Patricia Jiménez, Rafael Corchuelo. International Conference on Business Information Systems (BIS'15). GGS: Work in Progress; CORE B.
4. **A Novel Approach to Web Information Extraction.** Antonia M. Reina Quintero, Patricia Jiménez, Rafael Corchuelo. International Conference on Business Information Systems (BIS'15). GGS: Work in Progress; CORE B.

5. **On Member Labelling in Social Networks.** Rafael Corchuelo, Antonia M. Reina Quintero, and Patricia Jiménez. International Web Conference On Artificial Neural Networks (IWANN'15). GGS: Work in Progress; CORE B.

C.3. Proyectos o líneas de investigación en los que ha participado, indicando su contribución personal. En el caso de investigadores jóvenes, indicar líneas de investigación de las que hayan sido responsables.

1. **TED2021-130355B-C32.** “Mejora de la ciberseguridad y sus sostenibilidad en Beneficio de la sociedad y de las personas”. Proyecto de Transición Ecológica y Transición Digital. Ángel Jesús Varela Vaca, María Teresa Gómez López (Universidad de Sevilla). 2022 – 2025). 223.330,00 €. Equipo de Investigación.
2. **PID2020-112540RB-C44.** “Aether-US: Una aproximación holística de smart data para el análisis de datos guiado por el contexto: procesos de negocio inteligentes”. Plan Estatal 2017-2020 Retos. María Teresa Gómez, Ángel Varela (Universidad de Sevilla). 2021 – 2025. 221.672,00 €. Equipo de investigación.
3. **US-1381375.** “METAMORFOSIS: Marco de Transformación digital Mediante Customización software sobre gestión de datos, procesos de negocio y gobierno de la Seguridad”. Proyectos I+D+i FEDER Andalucía 2014-2020. David Benavides Cuevas / María Teresa Gómez López. 2022. 90.000,00. Equipo de investigación.
4. **P18-RT-1060.** “MARTITA: Ingeniería de Datos Aplicada a la Extracción, Semantización, Refinamiento y Explotación de Grafos de Conocimiento a Escala Web”. PAIDI: Proyectos I+D+i. Rafael Corchuelo, David Ruiz (Universidad de Sevilla). 2020-2022. 116.042,00. Equipo de Investigación.
5. **H2020-645751.** “Propelling Business Process Management by Research and Innovation Staff Exchange”. Horizonte 2020. Manuel Resinas, Antonio Ruiz (Universidad de Sevilla). 2015-2019. 1.273.500 € (99.000 € Universidad de Sevilla). Equipo de investigación.
6. **TIN2016-75394-R.** “Herramientas para la Ciencia de los Datos de la Web”. Plan Estatal 2013-2016 retos - Proyectos I+D+i. David Ruiz, Rafael Corchuelo, (Universidad de Sevilla). 2016-2019. 92.202 €. Equipo de investigación.
7. **FA8750-14-C-0240.** “DIG: Domain-specific Insight Graphs”. DARPA. Pedro Szekely, Craig Knoblock (Universidad de Southern California). 2014-2017. 7.685.884,00 \$. link . Phd Student.
8. **TIN2013-40848-R.** “Semantización y Publicación de datos Abiertos para la Integración de Servicios Electrónicos”. Plan Estatal 2013-2016 retos - Proyectos I+D+i. David Ruiz, Rafael Corchuelo (Universidad de Sevilla). 2014-2016. 142.375,86 €. Equipo de investigación.
9. **TIN2010-21744-C02-01.** “Integración de Aplicaciones y Datos en la Web”. Plan Nacional del 2010. Rafael Corchuelo (Universidad de Sevilla). 2011-2013. 66.800,00 €. PIF.

C.4. Participación en actividades de transferencia de tecnología/conocimiento y explotación de resultados *Incluya las patentes y otras actividades de propiedad industrial o intelectual (contratos, licencias, acuerdos, etc.) en los que haya colaborado. Indique: a) el orden de firma de autores; b) referencia; c) título; d) países prioritarios; e) fecha; f) entidad y empresas que explotan la patente o información similar, en su caso.*

1. Proyecto de transferencia. P030-15/E10. “SMART RETAIL OMNICHANNELUX: Sistema Inteligente de Gestión de Empresas del Sector Retail”, Rafael Corchuelo (IP, Universidad de Sevilla), Tier1, S.A., 2015-2016. 67.500 €. Investigador.
2. Patente: RTA-526-15 (201599902656710). ROLLER: A novel approach to web information extraction, Patricia Jiménez, 2015.
3. Patente: RTA-525-15 (201599902656528). TANGO: An ILP system for Web Information Extraction, Patricia Jiménez, 2015.