

BREVE CURRICULUM Y ENTREVISTA -ENERO 2015-



Concha Bielza se licenció en Ciencias Matemáticas en la Universidad Complutense de Madrid en 1989 y se doctoró en Informática en la Universidad Politécnica de Madrid en 1996 con premio extraordinario. Desde julio-2010 es Catedrática de Universidad en la UPM en el Departamento de Inteligencia Artificial de la ETSI Informáticos. Codirige el [Computational Intelligence Group](#) reconocido por la UPM desde marzo-2010.

Su investigación se centra prioritariamente en las áreas de aprendizaje automático, modelos gráficos probabilísticos (especialmente redes bayesianas), análisis de decisiones y optimización con metaheurísticas. El punto de partida es habitualmente un conjunto de datos sobre el dominio a estudiar, del que hay que extraer todo el conocimiento posible. Es lo que hoy en día se denomina "data science". Sus desarrollos metodológicos se han aplicado en numerosos dominios reales como medicina, agricultura, mecanizado industrial, bibliometría, bioinformática y neurociencia. Ambas, las áreas de investigación básica y las de aplicación, se benefician mutuamente, pues en numerosas ocasiones los desarrollos metodológicos han surgido de preguntas planteadas en el área de aplicación.

Entre sus trabajos científicos se encuentran: 4 libros (uno con Premio UPM al mejor libro de texto), 2 libros editados, 102 artículos en revistas (93 de ellos en revistas JCR), 31 capítulos en libros, 18 informes técnicos y 127 contribuciones en congresos (3 plenarias, 18 invitadas, 106 aceptadas). El 56% de los 93 JCR son del primer cuartil. Su índice h es 19 según Google Scholar y 11 según Web of Science.

Ha obtenido varios premios por trabajos de investigación: "Premio 1996 Student Paper Competition in Decision Analysis" y Finalista en 2002 de "Decision Analysis Society Practice Award Competition" (los dos otorgados por INFORMS de EEUU), "Premio 2004 de X Edición de los Premios Nacionales de Informática y Salud", "Best paper award" de los congresos MICAI 2009, IEA-AIE 2010, Biomag 2010, ICANN-2011, AIME 2013 y GECCO-2013.

Ha sido miembro del Comité de Programa de 50 congresos internacionales. Ha organizado 12 congresos/eventos siendo la Presidenta del Comité Científico de la XV Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA-2013). Ha impartido más de 40 conferencias/seminarios invitados en prestigiosas universidades y centros de investigación así como 3 tutoriales en congresos. Ejerce una intensa actividad de revisión de trabajos científicos en revistas JCR (47 manuscritos), congresos (+200) y libros (2).

Ha dirigido 8 tesis doctorales (7 más están en progreso), 9 tesis fin de máster y 26 trabajos fin de carrera.

Ha participado en 44 proyectos de investigación subvencionados públicamente (National Institutes of Health, Programa Marco de la UE, European Science Foundation, Consolider-Ingenio, CENIT-Ingenio) y en 12 contratos de investigación con entidades privadas (Iberdola, Telefónica, Coopers & Lybrand, Arthur Andersen, Panda Security, Atos Origin, Produban, Abbott Products Operations AG). Es coinventora de una patente tecnológica con el CIEMAT sobre adenocarcinoma de pulmón.

Ha colaborado con distintas agencias de evaluación de proyectos de investigación: Agencia Nacional Francesa, ANEP y MINECO para el Plan Nacional, Agencia Andaluza del Conocimiento. Es codirectora y profesora de la escuela de verano de la UPM "Advanced Statistics and Data Mining Summer School", que lleva ya 9 ediciones.

-- ¿Qué supone para usted recibir el premio?

Como Licenciada en Ciencias Matemáticas, mi formación fue muy teórica y rigurosa. Al llegar a la UPM se me abrieron las puertas a un mundo interdisciplinar donde poder poner en práctica mis conocimientos y aportar nuevas metodologías que apoyaran la toma de decisiones en muy diversos problemas reales.

Este premio significa el reconocimiento al esfuerzo realizado durante los más de 23 años que llevo en la UPM y por ello estoy muy agradecida, honrada y satisfecha. Dedicarse a la investigación además de a la docencia y a la gestión supone un nivel de sacrificio nada despreciable, sin contar con el alto grado de burocracia al que estamos sometidos. La concesión de este premio supone una inyección de ánimo para seguir dedicándome con empeño y pasión a una investigación genuina, cuidada y de calidad, con la que siempre he disfrutado. Sin duda supone un aliciente para continuar por la misma senda, a pesar de los tiempos difíciles que atravesamos de falta de apoyo a la investigación. Celebro que la UPM apueste por la excelencia científica con este tipo de reconocimientos y felicito a mis compañeros premiados en las otras modalidades.

Hoy en día la investigación no puede entenderse si no es en grupo. El esfuerzo de estos años no es únicamente mío. Es sobre todo del grupo de jóvenes investigadores que codirijo junto al profesor Pedro Larrañaga, gran compañero y amigo, y apoyo continuo en mi desarrollo académico desde 2007. Generar vocaciones científicas desde muy jóvenes es nuestro leitmotiv. Entre los colegas de la UPM me gustaría mencionar a dos personas ya fallecidas con las que comencé mi carrera, Sixto Ríos Insua (Facultad de Informática) y Enrique Ballesteros (ETSI Agrónomos). Debo agradecer mi reciente inmersión en la neurociencia a Javier de Felipe (CSIC-UPM) y mi interés por la inteligencia computacional a mis colegas Prakash Shenoy (EEUU) y Peter Lucas (Holanda). Por supuesto no puedo olvidar a maestros como mi director de tesis David Ríos Insua, Cristina Núñez que me hizo mágicas las matemáticas, Agustín Alonso con su admiración por la estadística (siempre en femenino), mis compañeros de la Escuela y más en concreto del Departamento de Inteligencia Artificial, mis ex-doctorandos, mis familiares, especialmente mis padres y hermanas y mis amigos. Gracias a todos.

-- ¿En qué consiste su investigación? Aplicaciones prácticas

Solemos disponer de un conjunto de datos sobre el dominio a estudiar, con incertidumbre intrínseca, y a partir de ahí tratamos de construir modelos estadístico-computacionales que extraigan todo el conocimiento posible del problema. Se cuenta también con la aportación, si la hubiere, del experto en el dominio. Los modelos recogerán las características relevantes y permitirán realizar predicciones, segmentaciones, clasificaciones y razonamientos, contestando a preguntas de interés. Aunque trabajamos con muchos tipos de modelos, nuestros modelos favoritos son las redes bayesianas, por su transparencia y versatilidad.

La siguiente lista enumera ejemplos de aplicaciones prácticas realizadas en diferentes áreas y con diferentes instituciones públicas y privadas.

- **Medicina:** tratamiento de la ictericia neonatal y asistencia respiratoria extracorpórea (con Hospital Gregorio Marañón), dando lugar a los sistemas de ayuda a la decisión *IctNeo* y *AREC*, respectivamente. Se consiguió disminuir el número de exanguinotransfusiones realizadas (con alto riesgo) a los bebés. Tratamiento del linfoma gástrico no-Hodgkin (con Radboud University Nijmegen, Holanda). Predicción de la combinación de medicamentos antiretrovirales para el tratamiento del VIH (con Hospital Carlos III).
- **Neurociencia,** enfermedades neurodegenerativas: Alzheimer, Parkinson, epilepsia, glioblastoma (con Fundación CIEN, farmacéutica suiza Abbott, King's College Hospital, Hospital Santa Creu i Sant Pau, Hospital Universitario de la Princesa, Universidad de Salamanca).
- **Neurociencia,** ultranatomía: modelos 3D de la morfología de los árboles dendríticos, clasificación de espinas dendríticas, clasificación y consenso en la nomenclatura de interneuronas, discriminación de células interneuronas y piramidales, distribución espacial de sinapsis y astrocitos (con Instituto Cajal del CSIC, Columbia University, George Mason University, EPFL de Suiza, etc.).
- **Micología:** Establecimiento de los puntos de corte para la susceptibilidad por infección del hongo *Candida* (con Instituto de Salud Carlos III).
- **Bioinformática:** Selección de genes en microarrays de ADN, selección de picos en datos de espectrometría de masas (con CIEMAT).
- **Agricultura:** Minimización del daño en la fruta y simulación de líneas de clasificación de fruta (sistema *SimLin*®). Simulación espacio-temporal de gases en cámaras de frigo-conservación (sistema *SimGas*). Ambos con la ETSI Agrónomos (UPM).
- **Gestión de embalses:** dos embalses del sureste de África (con IMES - Strobl Gruppe, Austria y con Iberdrola), dando lugar al sistema *BayRes*.
- **Astrofísica:** Clasificación de estrellas según características de la curva de luz. Datos del satélite Hipparcos (con Observatorio Virtual Español del INTA).
- **Mecanizado industrial:** Predicción de la rugosidad superficial en mecanizado a alta velocidad (con Instituto de Automática Industrial del CSIC).
- **Bibliometría:** Minería de textos para predecir número de citas, índice *h*, etc. de los trabajos científicos.
- **Empresas de servicio:** Predicción de la calidad en el servicio (Airtel, Coopers & Lybrand y Fundación Gil Gayarre), predicción de precios del mercado eléctrico (Arthur Andersen), de ventas (Telefónica), de demanda (European Computing Consultants), de código malicioso o *malware* (Panda Security), de rendimiento de equipos (Prosuban).

-- ¿Cuáles son los retos a los que se enfrenta en próximas investigaciones?

Desde 2008 trabajo activamente en neurociencia, en el proyecto Cajal Blue Brain (iniciativa española del Blue Brain europeo), y desde 2013 en el [Human Brain Project](#), uno de los dos únicos FET-Flagship concedidos (VII Programa Marco de la UE) en el que intervienen más de 80 instituciones. Estos proyectos tienen grandes retos. Por ejemplo, nos interesa conocer los principios que gobiernan el diseño neuronal. Hemos constatado que los ángulos entre ramas dendríticas en los puntos de bifurcación van cerrándose según nos alejamos del soma en las neuronas piramidales de la corteza del ratón, lo que está en estrecha relación con las conexiones sinápticas, y hemos encontrado distribuciones probabilísticas circulares que modelizan dichos ángulos. Estudiar si esos ángulos son diferentes en enfermedades cerebrales como autismo o esquizofrenia podría suponer un avance en el conocimiento de las mismas. Otro ejemplo es resolver la falta de consenso existente en la comunidad neurocientífica de una nomenclatura aceptada para clasificar las interneuronas GABAérgicas, tal y como publicamos en la revista Nature Reviews Neuroscience en 2013. Este artículo ha generado más de 110 citas (Google Scholar) en menos de dos años y aborda un problema estudiado desde hace más de un siglo (desde Ramón y Cajal). Conseguir tal consenso mediante una herramienta automática es uno de nuestros próximos grandes retos.

Finalmente, desde nuestra posición de expertos en estadística y aprendizaje automático, nos gustaría extender el uso de estas técnicas al mundo de la neurociencia demostrando un potencial por ellos desconocido. Este reto lo afrontamos mediante la escritura de un libro en el (poco) tiempo libre del que disponemos. Además, nos gustaría trabajar en la dirección inversa, tal y como acaba de proponer la "Intelligence Advanced Research Projects Agency" (IARPA) del gobierno de EEUU con el programa "Machine Intelligence from Cortical Networks" ([MICrONS](#)). Es decir, emplear el conocimiento de cómo funciona el cerebro humano como procesador sensorial y de datos para idear nuevos algoritmos de aprendizaje automático más potentes que permitan complejas tareas y aplicaciones inconcebibles usando la tecnología de hoy en día. De nuevo, el beneficio mutuo que mencioné anteriormente.