

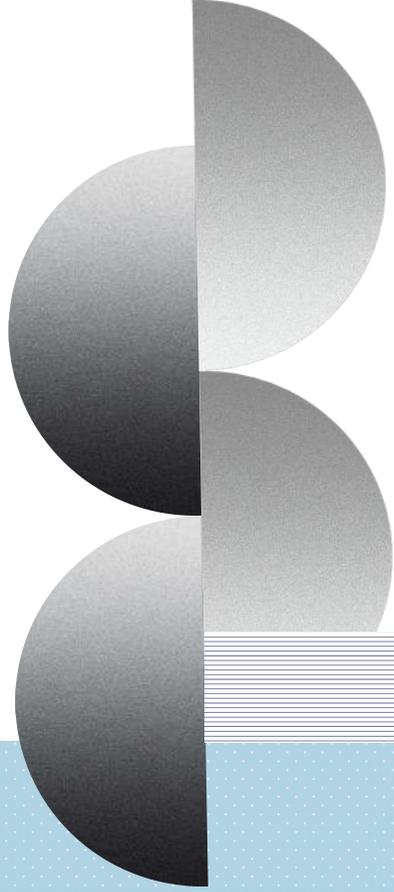


IMPULSO A LOS ESPACIOS DE DATOS

tecnal:a

MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE

JULIO/2021



tecnalia

MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE



IMPULSO A LOS ESPACIOS DE DATOS

Editado por TECNALIA Research & Innovation

JULIO/2021



Esta obra está bajo una licencia
Creative Commons Reconocimiento 4.0 Internacional

Para cualquier uso de fotos y otros materiales que no pertenezcan a TECNALIA Research & Innovation,
el permiso deberá obtenerse directamente del propietario de los derechos.



“ La gestión de los datos y su procesamiento inteligente es uno de los grandes desafíos por los que TECNALIA apuesta en los próximos años.

Queremos contribuir ideando y diseñando soluciones innovadoras y tecnológicas para Espacios de Datos confiables, interoperables y soberanos que desarrollen todo el potencial de la Economía del Dato para la sostenibilidad social, económica y medioambiental.

Aportamos talento especializado, la extensa capilaridad sectorial y nuestra capacidad de orquestación orientada a la construcción y operación de estos Espacios de Datos. ”



Joseba Laka

Director de Transformación Digital
TECNALIA



Presentación

“Los datos son el “nuevo petróleo”; son nuevos activos valiosos para los productos y servicios emergentes de los próximos años. Quienes dominen su procesamiento inteligente ejercerán un enorme control sobre el mercado y marcarán el futuro en un contexto cada vez más interconectado, automatizado y complejo.

Actualmente, unas pocas grandes plataformas tecnológicas (GAFAM - Google, Amazon Facebook, Apple y Microsoft) ostentan un enorme control sobre la gran mayoría los datos existentes, en general, de naturaleza B2C. Existe sin embargo una gran cantidad de datos de carácter industrial y no personal, que seguirá creciendo con el despliegue de redes y tecnologías IoT. Estos datos serán determinantes tanto para la mejora en productividad y competitividad de las empresas como para la resolución de los grandes desafíos relacionados con el bienestar o el medioambiente, retos que no podrán superarse sin una gestión inteligente de los datos.

Las empresas y organizaciones de todo tipo podrían beneficiarse de este potencial de mejora a través de un mercado abierto, seguro y confiable de datos, de una economía de datos, a través de espacios que faciliten el acceso, uso, interoperabilidad y control de los datos para los participantes.”¹

Nota (1): extraído del documento Posicionamiento sobre los Espacios de Datos, PLANETIC, enero 2021.

Los Espacios de Datos son estructuras que proporcionan confianza y seguridad para la compartición voluntaria de datos entre diversos agentes de manera homogénea a través de mecanismos combinados de gobernanza, organizativos, legales y técnicos. Estos espacios facilitan la interoperabilidad para acceder a los datos o transferirlos y posibilitan su reutilización eficiente y legítima en un contexto de soberanía y control para las partes sobre sus propios datos.

El acceso y compartición de datos útiles, en las condiciones de privacidad y seguridad adecuadas, resulta por tanto vital en un nuevo contexto de la economía de los datos. Porque el valor de los servicios asociados a estos datos muchas veces se incrementa agregando significado a partir de su combinación desde diferentes fuentes. La interoperabilidad y la reducción de los costes de transacción de datos de calidad en contextos voluntarios, seguros y confiables, en los que las partes tengan soberanía y control sobre ellos, resultan por tanto elementos claves.

Los Espacios de Datos proporcionan:

- ▶ Confianza para el acceso a datos (eventualmente, tanto datos personales como no personales, privados y abiertos) en posesión de distintas entidades y que voluntariamente quieran proporcionar.
- ▶ Protección para que el acceso y utilización de los datos se realice en las condiciones de privacidad y seguridad precisas.
- ▶ Garantías para que quienes ostenten control total o parcial sobre los datos puedan ejercer su soberanía.
- ▶ Mecanismos estandarizados para funcionar dentro de un sistema abierto e interoperable.
- ▶ Fórmulas para, eventualmente, monetizar esta operativa.

Según la *Estrategia Europea para Datos*, las cifras previstas para 2025 con relación a los datos son:

- ▶ 530% de incremento del volumen global de datos respecto a 2018.
- ▶ 829.000 millones de euros de valor de la economía de los datos en la UE27 respecto a 301.000 millones de euros (2,4% del PIB) en 2018.
- ▶ 10,9 millones de profesionales de datos en la UE27 respecto a los 5,7 millones en 2018.
- ▶ 65% de la población de la UE con competencias digitales básicas respecto al 57% en 2018.

Esta visión se produce en un contexto caracterizado por:

- ▶ La explosión de los datos en los próximos años, especialmente los de carácter industrial.
- ▶ La adopción limitada de servicios de datos y de nube por parte de empresas y administraciones, y una demanda latente de este tipo de servicios no explotada.
- ▶ La migración de paradigmas desde la nube al nuevo *cloud/edge continuum*.
- ▶ Las nuevas expectativas de los usuarios con relación a los datos.
- ▶ Las limitadas capacidades competenciales tanto en la empresa como en la ciudadanía con relación a estos temas.



Índice

Páginas
6 – 9

01

Contexto y
motivación

Páginas
10 – 19

02

Iniciativas
singulares

Páginas
20 – 37

03

Aportación
tecnológica y
proyectos

Páginas
38 – 48

04

Casos de uso

Páginas
49 – 54

05

Equipo
TECNALIA

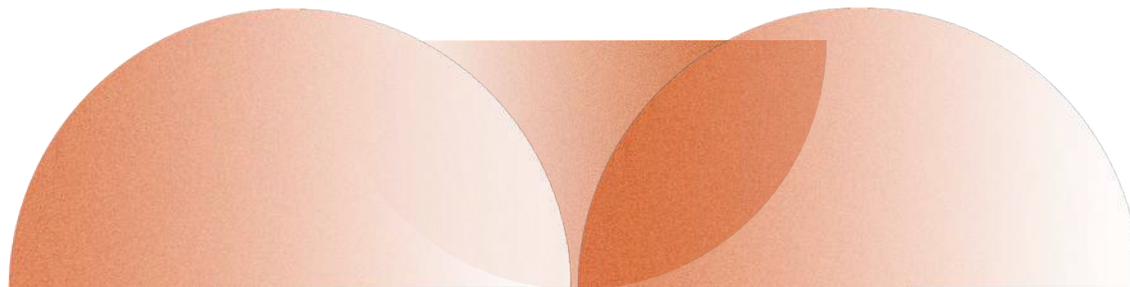
Páginas
55 – 56

06

Anexo



01 Contexto y motivación



Páginas
6 – 9

Páginas
10 – 19

Páginas
20 – 37

Páginas
38 – 48

Páginas
49 – 54

Páginas
55 – 56

01

02

Iniciativas
singulares

03

Aportación
tecnológica y
proyectos

04

Casos de uso

05

Equipo
TECNALIA

06

Anexo

Contexto y
motivación

- Contexto y antecedentes
- Motivación para una apuesta por los Espacios de Datos



01 Contexto y motivación

Contexto y antecedentes

A finales de 2019, con la presentación del Pacto Verde Europeo, la Comisión Europea establecía los pilares de *digitalización* y *sostenibilidad ambiental* como claves en la hoja de ruta a seguir e identificaba la necesidad, entre otras cuestiones, de impulsar los datos accesibles e interoperables que, combinados con nuevas infraestructuras digitales y las soluciones de inteligencia artificial, facilitan la toma de decisiones informada.

También se identificaba la necesidad de impulsar la *soberanía digital europea* mediante el desarrollo de infraestructuras digitales, seguras y sostenibles y el impulso de la aparición de operadores europeos en infraestructuras y servicios de la nube así como otras acciones para evitar posibles riesgos relativos a las interferencias de países terceros en los datos europeos.

Estas cuestiones se recogen en la *Estrategia Europea para Datos* publicada en febrero de 2020 a la que se acompaña posteriormente con una considerable inversión, a partir del *Plan de Recuperación para Europa*. La Comisión propone que cada plan de recuperación y resiliencia de los estados miembros incluya como mínimo un 20% de inversión en digitalización.

La estrategia tiene como objetivo la creación de un espacio único europeo de datos que sirva para aumentar el uso y la demanda de datos, y de servicios y productos basados en datos en todo el mercado único, garantizando su acceso y su uso responsable. Mediante las normas y mecanismos asociados se pretende garantizar que:

- ▶ Los datos puedan fluir;
- ▶ Se respeten plenamente las normas y valores europeos;
- ▶ Las normas de acceso a los datos y su uso sean justas, prácticas y claras.

Cloud rulebook
 European Alliance on Industrial data, edge and cloud
 Common data spaces
 EU Cloud federation



Inversión total esperada: 4.000-6.000 millones euros entre 2021 y 2027

La ambición de la Estrategia es que la UE se convierta en "la economía ágil en el manejo de los datos más atractiva, más segura y más dinámica del mundo, y empoderar así a Europa con datos para mejorar las decisiones que se toman y la vida de todos sus ciudadanos".



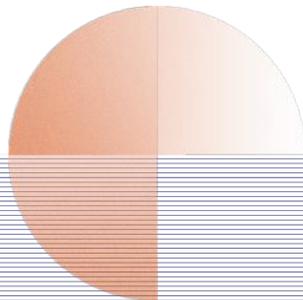
01 Contexto y motivación

Contexto y antecedentes

En paralelo a la *Estrategia Europea para Datos*, se han y se están desarrollando muchas **políticas relacionadas** e instrumentos para favorecer la economía del dato en Europa.

Los Espacios de Datos son elementos singulares de estas políticas e instrumentos. Estos espacios de datos, centralizados o distribuidos, servirán para fomentar ecosistemas que creen nuevos productos y servicios basados en datos más accesibles. Los problemas a los que se dirige la estrategia son:

- ▶ Disponibilidad de datos: el valor de los datos reside en su uso y reutilización, sea G2B, B2B, B2G y G2G.
- ▶ Desequilibrios en el poder de mercado.
- ▶ Interoperabilidad de datos.
- ▶ Gobernanza de datos.
- ▶ Infraestructuras y tecnologías de datos.
- ▶ Empoderamiento de las personas para que ejerzan sus derechos.
- ▶ Cualificación y alfabetización en materia de datos.
- ▶ Ciberseguridad.



Hacia un espacio común europeo de datos	Una Estrategia Europea para Datos	Estrategia Anual de Crecimiento Sostenible 2021	Guidance for member states recovery and resilience plans	Data Governance Act	Towards a European strategy on B2G data sharing for the public interest
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0232&from=EN	https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0066&from=ES	https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0057&from=EN	https://ec.europa.eu/info/sites/info/files/3_en_document_tray_all_service_part1_v3_en_0.pdf	https://ec.europa.eu/info/law/liter-regulation/have-your-say/initiatives/12491-legislative-framework-for-the-governance-of-common-european-data-spaces	https://www.euractiv.com/document/12491-legislative-framework-for-the-governance-of-common-european-data-spaces
The Digital Economy and Society Index (DESI) 2020 - Spain	Digital Europe Programme	España Digital 2025	España Puede	Estrategia Nacional en Inteligencia Artificial	Horizon Europe Cluster 4 Digital, Industry and Space
https://ec.europa.eu/digital-single-market/en/scoreboard/spain	https://ec.europa.eu/digital-single-market/en/europe-investing-digital-digital-europe-programme	https://portal.mineco.gob.es/RecursosArticulo/mineco/rensa/ficheros/noticias/2018/Anenda_Digital_2025.pdf	https://www.lamocloa.gob.es/residente/actividades/Documents/2020/07/02020_PlanRecuperacion.pdf	https://portal.mineco.gob.es/RecursosNoticia/mineco/rensa/noticias/2020/201202_nn_eniay.pdf	https://ec.europa.eu/research/horizon-europe/annex-4.pdf



01 Contexto y motivación

Motivación para una apuesta por los Espacios de Datos

La Estrategia propone el desarrollo de espacios de datos en nueve ámbitos estratégicos y en ámbitos de interés público entre los que se citan fabricación, Pacto Verde, movilidad, salud, financiero, energía, agrario, administraciones públicas y cualificaciones. Esto debe conducir a la disponibilidad de grandes repositorios de datos en tales sectores y ámbitos, junto con las herramientas y las infraestructuras técnicas necesarias para utilizar e intercambiar datos, así como unos mecanismos de gobernanza adecuados. Estos espacios deben complementarse con políticas que estimulen el uso de datos y la demanda de servicios enriquecidos con datos.

La Comisión propone que cada plan de recuperación y resiliencia de los estados miembros incluya un nivel mínimo del 20% de gasto relacionado con la tecnología digital. Se invita a tomar como referencia, siempre que sea posible, indicadores existentes como los incluidos en el *Índice de Economía y Sociedad Digitales (DESI)*. También, aportan algunas categorías para comprobar los objetivos de las inversiones de los estados miembros entre las que figuran:

- ▶ Digitalización de las empresas: acelerar la toma de decisiones y la ejecución con la automatización basada en la inteligencia artificial, rediseñar las cadenas de suministro para optimizar la resiliencia y la velocidad en base a espacios de datos intersectoriales y ejecutar en infraestructuras europeas de borde (*edge*) y de nube innovadoras, seguras y energéticamente eficientes.
- ▶ Inversión en capacidades digitales y despliegue de tecnologías avanzadas: como espacios de datos, computación de borde (*edge*), computación de alto rendimiento, ciberseguridad, inteligencia artificial, infraestructuras de computación cuántica, infraestructura en la nube, cadena de suministro para Internet de las cosas, componentes y sistemas electrónicos y microelectrónica y promover la diversificación de la oferta.

El *Plan España Digital 2025* de julio de 2020 es la nueva Agenda Digital orientada a impulsar la Transformación Digital de España como una de las palancas fundamentales para relanzar el crecimiento económico, la reducción de la desigualdad, el aumento de la productividad y el aprovechamiento de todas las oportunidades que brindan estas nuevas tecnologías. El Plan recoge un conjunto de medidas, reformas e inversiones, articuladas en diez ejes estratégicos, alineados con las políticas digitales marcadas por la Comisión Europea para el nuevo periodo. En el noveno de estos diez ejes se indica: "Favorecer el tránsito hacia una economía del dato, garantizando la seguridad y privacidad y aprovechando las oportunidades que ofrece la Inteligencia Artificial". La medida 44 dice:

ESTRATEGIA CLOUD: ESPACIOS COMPARTIDOS EUROPEOS DEL DATO. España desempeñará un rol activo para formar parte de los espacios compartidos de la *European Cloud Federation* y para potenciar un espacio ibérico, junto a Portugal, que impulse el desarrollo de tecnologías avanzadas de computación de datos: *HPC, Quantum computing o Edge computing*, entre otras, con el objetivo de convertirse en un hub de conectividad y, consecuentemente, un potencial punto de concentración de infraestructuras de datos. Es irrenunciable apostar por la disponibilidad de datos y la innovación digital a nivel europeo, y desde España se trabajará para desarrollar alianzas en este sentido, impulsando las inversiones privadas en *Data Centers* y situando al país como un hub de negocios alrededor del espacio cloud europeo.

La nueva *Estrategia Nacional de Inteligencia Artificial* de diciembre de 2020, en su eje 3, "desarrollar plataformas de datos e infraestructuras tecnológicas para dar soporte a la IA" propone:

La regulación deberá reforzar los compromisos para la protección, la gestión integral, la apertura y el acceso a la información de la que dispone el sector público y para que, de manera coordinada con el sector privado, se facilite la creación de nuevos servicios de valor añadido basados en datos. En esta dirección, y en línea con lo propuesto en la nueva Estrategia Europea de Datos que fomenta la creación de data pools sectoriales, es necesario investigar y desarrollar un marco de acción para instrumentar modelos de compartición de datos para incentivar la innovación y experimentación a distintos niveles:

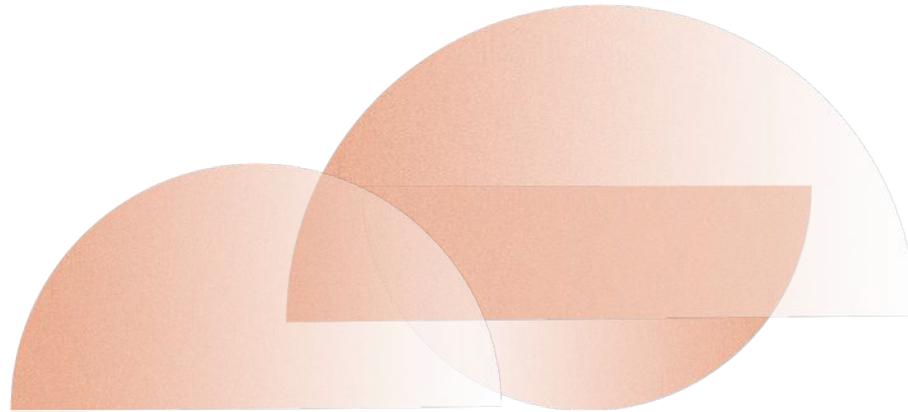
- *G2B (Government to Business):* compartir datos entre el gobierno y las empresas.
- *G2G (Government to Government):* compartir datos entre administraciones públicas.
- *G2C (Government to Citizen):* compartir datos entre el gobierno y la ciudadanía.

Y se establece como una de las medidas, dentro de la Línea de Actuación "Impulsar las plataformas de datos":

Medida 13. Creación de espacios compartidos de datos sectoriales e industriales y repositorios descentralizados y accesibles, que faciliten la creación de servicios de valor añadido basados en infraestructuras de datos y la construcción de aplicaciones de IA seguras e innovadoras en un entorno de confianza.



02 Iniciativas singulares



Páginas
6 – 9

Páginas
10 – 19

Páginas
20 – 37

Páginas
38 – 48

Páginas
49 – 54

Páginas
55 – 56

01

Contexto y
motivación

02

Iniciativas
singulares

03

Aportación
tecnológica y
proyectos

04

Casos de uso

05

Equipo
TECNALIA

06

Anexo

- Mapa de iniciativas
- Iniciativas y Programas Relacionados
- Programas de I+D+i Relacionados



02 Iniciativas singulares

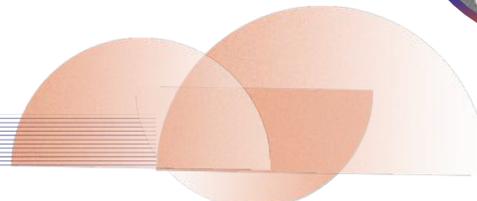
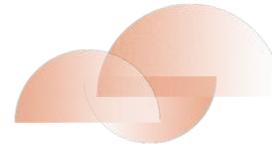
Mapa de iniciativas

La Comisión Europea y, en algunos casos, determinados estados miembros, han venido promoviendo medidas e iniciativas orientadas al impulso de la soberanía digital europea mediante el desarrollo de infraestructuras digitales, seguras y sostenibles, y el fomento de la aparición de operadores europeos en infraestructuras y servicios en la nube, así como otras acciones para evitar posibles riesgos relativos a las interferencias de países terceros en los datos europeos. Entre estas medidas e iniciativas se pueden citar las siguientes:

- ▶ GAIA-X.
- ▶ Proyecto Importante de Interés Común Europeo en la Nueva Generación de Infraestructuras y Servicios en la Nube/Borde – IPCEI CIS.
- ▶ *International Data Space Association - IDSA.*
- ▶ *European Union Agency for Cybersecurity (ENISA) – Cybersecurity certification scheme on cloud services y Cloud Service Providers Certification Working Group.*
- ▶ *European Alliance on Industrial Data, Edge and Cloud.*

Además, existen iniciativas nacionales y regionales complementarias que están promoviendo actividades para facilitar el despliegue de casos de uso sobre estas infraestructuras como:

- ▶ Consorcio Nacional Industrial para el Impulso de la Economía del Dato y de la Inteligencia Artificial – INDESIA.
- ▶ *Basque Artificial Intelligence Center – BAIC.*
- ▶ Plataforma Tecnológica Española para la Promoción y Difusión de las Tecnologías Digitales – PLANETIC.
- ▶ Cluster Big Data Madrid.





02 Iniciativas singulares

Iniciativas y programas Relacionados



GAIA-X

GAIA-X es una iniciativa, originariamente Franco-Alemana, lanzada en 2020 para el desarrollo de una infraestructura de datos eficiente y competitiva, segura y fiable para la Unión Europea. La iniciativa, que persigue la soberanía tecnológica en la nube, se ha ido extendiendo a otros estados miembros de la UE. El ecosistema asociado se constituye como una Asociación sin ánimo de lucro.

Su concepto se basa en crear un sistema federado, seguro e interoperable que cumpla con los estándares de soberanía digital y permita dar soporte a la innovación empresarial. Su implementación se soporta en una serie de servicios: Servicios de Identidad y Confianza, Catálogo Federado; Servicios de Intercambio de Datos Soberanos y Servicios de Cumplimiento. Estos servicios sirven para vincular los ecosistemas de datos e infraestructuras en la nube existentes para que las empresas conserven la soberanía y el control sobre sus datos y tengan más transparencia y capacidad de elección sobre las plataformas y los recursos en la nube que utilizan, evitando al mismo tiempo la dependencia de los proveedores.

GAIA-X conecta infraestructuras centralizadas y descentralizadas para convertirlas en un sistema homogéneo y utilizable en el que los participantes puedan utilizar datos y servicios de forma soberana en espacios de datos específicos del sector.

Los principios que rigen GAIA-X son la apertura-transparencia, interoperabilidad, federación, autenticidad y confianza.

Soberanía de datos Reducir la dependencia de proveedores no europeos Conseguir mayor adopción para los servicios en la nube Crear infraestructuras de datos que generen innovación

HUB ESPAÑA GAIA-X

GAIA-X se apoya en un conjunto de hubs nacionales como mecanismos de articulación de los ecosistemas en cada estado miembro que, a su vez, se encuentran interconectados para facilitar la alineación hacia los intereses comunes. Los hubs buscan desarrollar el ecosistema local, agrupar las iniciativas nacionales, impulsar iniciativas de abajo a arriba y proporcionar un punto central de contacto para las partes interesadas en sus respectivos países. El hub español busca dar forma a un ecosistema sólido e integrador sobre compartición de datos a nivel nacional, integrando al sector público y privado, a los centros de investigación y desarrollo tecnológico, a grandes empresas así como PYMEs, microempresas y startups. Trabjará en el despliegue de la gobernanza de datos, en el desarrollo y asimilación de arquitecturas de referencia para los diferentes sectores, y en la integración de las PYMEs y microPYMEs en los espacios de datos, acompañándolas en su formación y en la implementación de casos de uso basados en datos e IA. En este hub, la industria y los centros de investigación tecnológica contribuirán a construir e implementar las infraestructuras de cloud/edge continuum, los servicios de computación y la arquitectura de compartición de datos que sirvan para potenciar el desarrollo de aplicaciones de datos e IA para la industria sectorial.

ROL DE TECNALIA

TECNALIA arranca trabajos en GAIA-X en su etapa inicial en marzo de 2020, y entra a formar parte en la asociación GAIA-X AISBL con su constitución en 2021. Ha estado participando en diferentes comités y grupos de trabajo. Ha estado involucrado también en el equipo impulsor del Hub España GAIA-X.

A partir de un portafolio de proyectos en marcha, fundamentalmente en el contexto Horizonte Europa, está desarrollando componentes y activos relacionados con servicios GAIA-X: políticas y control de uso, acreditación y certificación, catálogos federados, conectores de datos, sistemas de monitorización, prototipos y pruebas de concepto en determinados escenarios.



02 Iniciativas singulares

Iniciativas y programas Relacionados



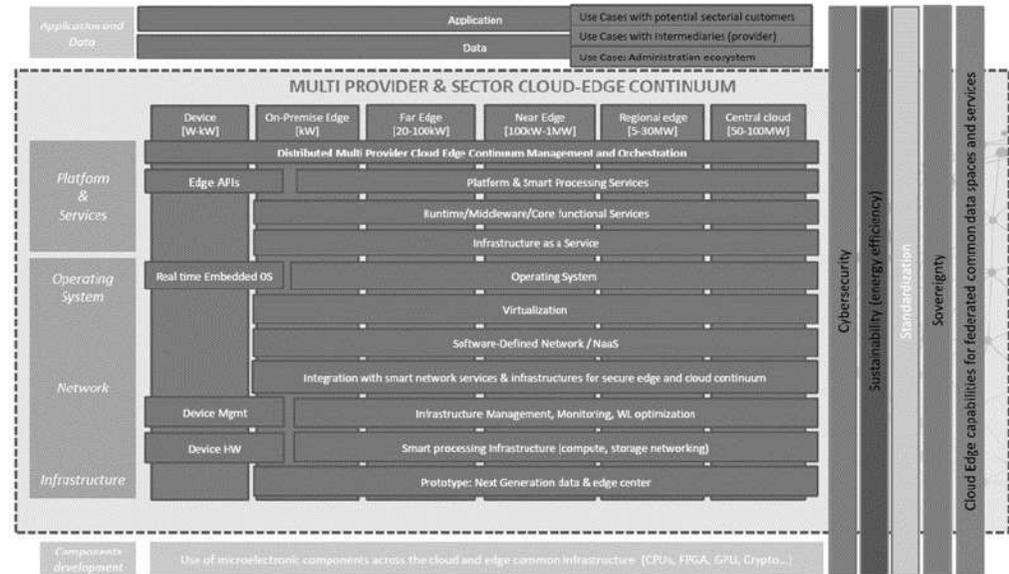
IPCEI – CIS: IMPORTANT PROJECT OF COMMON EUROPEAN INTEREST – NEXT GENERATION CLOUD INFRASTRUCTURES AND SERVICES

Varios estados miembros están promoviendo el desarrollo de un Proyecto Importante de Interés Común Europeo (*Important Project of Common European Interest o IPCEI*) sobre la próxima generación de servicios e infraestructuras en la nube / borde – IPCEI CIS- para proporcionar un acceso competitivo y justo a estas capacidades desde cualquier lugar de la UE.

Los IPCEI son proyectos en colaboración público – privada llevados a cabo conjuntamente por varios estados miembros con el fin de emprender proyectos transnacionales a gran escala, de importancia estratégica para la UE y para la consecución de los objetivos europeos comunes. Los IPCEI buscan reunir conocimientos, experiencia, recursos financieros y agentes económicos en toda la Unión, para superar importantes fallos del mercado y retos sociales que no podrían abordarse de otro modo. Los IPCEI se establecen para generar externalidades positivas más allá de los estados miembros, empresas o sectores participantes. Este tipo de proyectos cuentan con diferentes ventajas de financiación: sobre la base de un amplio conjunto de costes subvencionables, incluyendo los costes de I+D+i y del primer despliegue industrial (es decir, entre la línea piloto/demo y antes del inicio de la producción en masa), un IPCEI puede recibir una ayuda de hasta el 100% del déficit de financiación.

Participan actualmente en la definición del proyecto Alemania, Bélgica, Eslovenia, España, Francia, Hungría, Italia, Letonia, Luxemburgo, Países Bajos y Polonia.

El proyecto está orientado a la construcción de la próxima generación de infraestructuras de procesamiento de datos paneuropea soberanas, interoperables, multiusos y neutrales respecto a los proveedores asociados, que estén interconectadas, distribuidas y sean escalables. Se incluye en su concepción el desarrollo y despliegue de capacidades de computación de borde (*edge computing*) en tiempo real (muy baja latencia); el diseño de plataformas *middleware* seguras, de bajo consumo e interoperables para múltiples usos sectoriales; y el desarrollo e implementación de la nube inteligente y los servicios de borde de carácter ultraseguro, en tiempo real y de bajo consumo de energía. Se refieren a estas infraestructuras como “*Distributed Multi Provider Cloud-Edge Continuum*”.



ROL DE TECNALIA

Los IPCEI van dirigidos a la comunidad empresarial, por lo que el rol de TECNALIA tiene que ser de soporte y acompañamiento a empresas. Puede aportar como activos para este acompañamiento: desarrollos relacionados con infraestructuras y servicios (similares a los de GAIA-X), capacidad de orquestar Casos de Uso donde validar despliegues industriales y una extensa red de colaboradores internacionales que sirva para dar forma al proyecto integrado.



02 Iniciativas singulares

Iniciativas y programas Relacionados



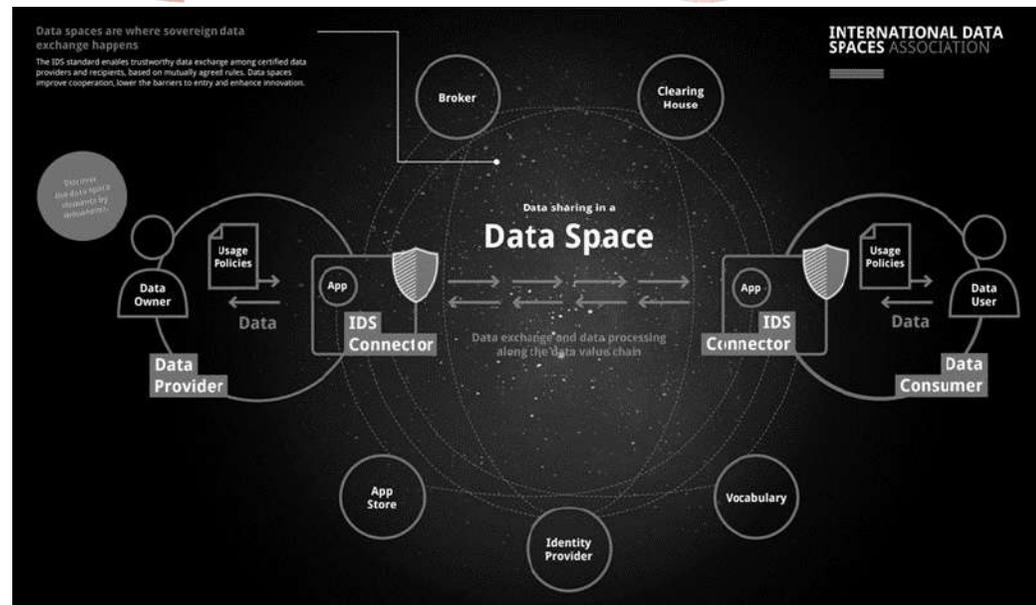
IDS - INTERNATIONAL DATA SPACE ASSOCIATION

La Iniciativa promovida por IDSA (*International Data Spaces Association*) tiene como objetivo contribuir a la economía global digital a partir de mecanismos estandarizados y seguros de intercambio y compartición de datos en entornos confiables, garantizando a mismo tiempo la soberanía sobre estos datos para sus creadores o proveedores, lo que se constituye en el elemento central de la aproximación. Estos mecanismos forman la base para el desarrollo de nuevos servicios inteligentes y facilitan procesos innovadores trans-empresariales.

IDSA está construyendo un modelo de referencia que forma la base para un mercado de datos basado en los valores europeos, es decir, la privacidad y seguridad de los datos, la igualdad de oportunidades a través de un diseño federado, la garantía de la soberanía de los datos para el creador de los datos y la generación de confianza entre los participantes.

El modelo se compone de cinco capas:

- ▶ La Capa de Negocio específica y categoriza los diferentes roles que pueden asumir los participantes de los Espacios de Datos Internacionales y especifica las principales actividades e interacciones conectado con cada uno de estos roles.
- ▶ La capa funcional define los requisitos funcionales de los espacios de datos internacionales, además de las características concretas que se derivan de ellos.
- ▶ La capa de proceso especifica las interacciones que tienen lugar entre los diferentes componentes de los espacios de datos internacionales.
- ▶ La capa de información define un modelo conceptual que hace uso de los principios de datos vinculados (*linked data*) para describir tanto los aspectos estáticos como dinámicos de los componentes del Espacio de Datos.
- ▶ La capa de sistema se ocupa de la descomposición de los componentes lógicos del software, considerando aspectos como la integración, configuración, implementación y extensibilidad de estos componentes.



ROL DE TECNALIA

TECNALIA se incorpora a IDSA en 2018; desde entonces está participando en sus Grupos de Trabajo y en el desarrollo de activos y componentes vinculados con la arquitectura IDS-RAM (Reference Architecture Model): componentes para la gestión de políticas y control de uso, conectores de datos, sistemas de monitorización, prototipos y pruebas de concepto en determinados escenarios (energía, industria y comunidades inteligentes). Recientemente, TECNALIA ha sido oficialmente reconocido como un "IDS Competence Center" y también como "Implementation Partner". Participamos en el "IDSA- Acceleration programme" para el soporte a organizaciones que desean adoptar IDS.



02 Iniciativas singulares

Iniciativas y programas Relacionados



INDESIA: CONSORCIO NACIONAL INDUSTRIAL PARA EL IMPULSO DE LA ECONOMÍA DEL DATO Y DE LA INTELIGENCIA ARTIFICIAL

IndesIA es un consorcio impulsado por seis empresas -Repsol, Gestamp, Navantia, Técnicas Reunidas, Telefónica y Microsoft- para promover el uso de los datos y la inteligencia artificial en el ámbito industrial e impulsar el desarrollo de una nueva economía que genere crecimiento económico en el país. El consorcio cuenta con el apoyo y experiencia de un amplio número de empresas y organismos de referencia en este ámbito, como el *Basque Artificial Intelligence Center* (BAIC). El consorcio prevé la formación de un ecosistema de empresas, startups, centros tecnológicos y universidades especializadas en la aplicación de la Inteligencia Artificial en el ámbito industrial. Esta red de colaboración permitirá extender rápidamente los conocimientos y prácticas más eficientes, además de adaptarlos a las necesidades particulares de cada sector.

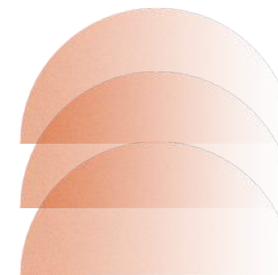
La iniciativa se articula en torno a los siguientes ejes:

- ▶ Identificar casos de uso en el ámbito industrial que puedan ser resueltos mediante el dato y la Inteligencia Artificial, demostrando el valor que pueden aportar estas tecnologías y su transversalidad.
- ▶ Crear unos mecanismos aceleradores que agilicen el proceso de desarrollo de soluciones basadas en *Big Data* e Inteligencia Artificial, facilitando el acceso a los recursos técnicos y económicos necesarios para implementarlos.
- ▶ Generar un ecosistema de *startups*, centros tecnológicos y universidades especialistas en la investigación y el desarrollo de soluciones de Inteligencia Artificial aplicadas al ámbito industrial que permitirá compartir y aplicar los conocimientos y las soluciones más eficientes.

- ▶ Impulsar la creación de una gran plataforma interoperable de datos industriales que promueva el desarrollo y consumo de soluciones de Inteligencia Artificial.
- ▶ Alcanzar acuerdos para facilitar el acceso a tecnologías de vanguardia (IoT, 5G, cloud, super-computación, quantum, *edge computing*...) que sirven para habilitar el desarrollo de los casos.
- ▶ Crear una Escuela de Datos & Inteligencia Artificial para poder involucrar y capacitar a profesionales del sector industrial en el uso y la analítica del dato mediante itinerarios de formación adecuados, que además se focalicen en el fomento de la diversidad, la igualdad de género y la apuesta por los perfiles STEM.

En sus primeras etapas se habían identificado más de 60 casos de uso basados en Inteligencia Artificial y analítica de datos para ayudar a traccionar las cadenas de valor de cinco grandes ámbitos industriales y su tejido empresarial: energía, automoción, naval, telecomunicaciones e ingeniería. IndesIA contempla la generación de una biblioteca de casos industriales transversales y funcionales, debidamente documentados y con acceso a los datos que han permitido resolverlos, que reducirá las barreras de entrada a la tecnología de Inteligencia Artificial para las empresas y para las más de 100 pequeñas y medianas empresas del consorcio.

El proyecto tractor también contempla impulsar la creación de una gran plataforma interoperable de datos industriales que facilite el desarrollo y el consumo de soluciones de Inteligencia Artificial y analítica de datos. La plataforma busca acelerar la ingesta de datos desarrollando conectores con los principales proveedores de software y hardware industriales, garantizando la captación de datos en tiempo real de la actividad de las distintas empresas. Con ellos se busca generar *data lakes* abiertos con datos agregados, fiables y listos para ser utilizados transversalmente en el desarrollo de sus soluciones de Inteligencia Artificial.



Para la gestión de los datos se contemplan los principios europeos de protección y soberanía del dato, proporcionando mecanismos para que el propietario del dato controle dónde se almacena, quién accede a él y qué tipo de procesamiento se realiza sobre el mismo, garantizando un tratamiento seguro de esos datos (anonimización, etc.). La seguridad y privacidad del dato industrial y personal será un principio de diseño básico para la tecnología, las plataformas y los casos de uso que desarrolle y promocióne el consorcio.

El consorcio contempla la colaboración con universidades públicas y privadas para potenciar la empleabilidad mediante el *upskilling* (perfeccionamiento) y *reskilling* (recapacitación) de los empleados en STEM, con foco en la Inteligencia Artificial. Para ello, se pretende trabajar en itinerarios de formación para cubrir tanto los conocimientos generalizados que deben obtener los empleados del sector industrial para entender mejor cómo estas soluciones pueden ayudar en su día a día, como los más especializados para hacer *reskilling* y formar internamente a nuevos perfiles como *data scientists* y *data engineers*.

ROL DE TECNALIA

Participa indirectamente como miembro del Basque Artificial Intelligence Center para incrementar la competitividad a través de los datos y de la IA para la mejora industrial, la sostenibilidad, la eficiencia energética, el uso de nuevos materiales y la defensa la economía circular.



02 Iniciativas singulares

Iniciativas y programas Relacionados



CLOUD SERVICE PROVIDER CERTIFICATION WORKING GROUP

La Ley de Ciberseguridad de la Unión Europea (EUCA) sienta las bases para establecer un marco de la UE para la certificación de ciberseguridad de los servicios, productos y procesos de TI, incluidos los servicios prestados por los proveedores de servicios de nube (CSP).

El Grupo de Trabajo de Certificación de Proveedores de Servicios en la Nube (*Cloud Service Provider Certification Working Group*), CSPCERT, se crea en diciembre de 2017 en respuesta a esta Ley con el objetivo sentar las bases para un marco europeo para la certificación de ciberseguridad de los CSP.

ROL DE TECNALIA

TECNALIA participó inicialmente como miembro contribuyendo a las discusiones técnicas y liderando el análisis de esquemas de certificación *cloud* y definición de controles de seguridad. Fue editora del documento final entregado a la Comisión Europea y ENISA con las recomendaciones; es actualmente *co-chair* del grupo, como experta neutral.



EUROPEAN UNION AGENCY FOR CYBERSECURITY (ENISA) – CLOUD SERVICES

El Grupo de Trabajo 2 de ENISA, Cloud Services, ha estado trabajando en la implementación del *cybersecurity certification scheme on cloud services* alineado a la Regulación (EU) 2019/881, Cybersecurity Act.

Las tareas clave de este grupo de trabajo ad hoc han sido la revisión de los esquemas de certificación de seguridad en la nube existentes, el establecimiento de un alcance preciso del futuro esquema candidato, la precalificación de los elementos que ENISA necesita incluir en un esquema de certificación de ciberseguridad, el apoyo a la redacción de un esquema de conformidad con las disposiciones del Reglamento y, de manera general, el apoyo a ENISA en el desempeño de sus tareas en relación al esquema de certificación de la ciberseguridad para servicios en la nube.

ROL DE TECNALIA

Aportación experta para el Grupo de Trabajo, liderando el grupo temático 2 encargado de definir los controles y requisitos de seguridad *cloud* en los tres niveles, básico, medio y alto. Asimismo, ha participado en el grupo temático que ha definido el alcance de cada nivel en sus diferentes dimensiones (p.ej., rigor, alcance, profundidad, etc.)



Leire Orue-Echevarria es designada copresidenta técnica neutral para promover el trabajo de CSPCERT Working Group, septiembre 2019.

Leire Orue-Echevarria lidera en TECNALIA la línea de trabajo en Tecnologías Cloud. Es doctora en Humanidades y Ciencias Sociales (Economía) por la Universitat Abat Oliba CEU y graduada en informática por la Universidad de Deusto. Es evaluadora independiente experta, editora y copresidenta del Grupo de Trabajo CSPCERT y miembro del SCCG (Stakeholder Cybersecurity Certification Group); ha sido editora del estudio SMART – Certification Schemes for Cloud Computing (2016/029, <https://op.europa.eu/es/publication-detail/-/publication/6a83608d-0fdb-11e9-81b4-01aa75ed71a1>), miembro del subcomité CTN 71/SC 38, servicios y plataformas para aplicaciones distribuidas, y participa en los Grupos de Trabajo de GAIA-X.

Ha publicado más de 30 artículos en temas como ingeniería de software para (multi) *cloud*, privacidad en la nube, modelos de negocio y la aplicación de tecnologías *cloud* en diferentes sectores.



02 Iniciativas singulares

Iniciativas y programas Relacionados



BAIC – BASQUE ARTIFICIAL INTELLIGENCE CENTER

El *Basque Artificial Intelligence Center*, BAIC, es una iniciativa promovida por el Gobierno Vasco en 2021 cuyo objetivo es el impulso de la Inteligencia Artificial en el País Vasco a través de la colaboración público-privada con un conjunto de empresas tractoras, constituyéndose como instrumento para la adopción rápida de la IA por parte de la industria.

Los principales retos a los que busca dar respuesta el BAIC son los siguientes:

- ▶ Centrarse en la aplicación y en los casos de uso próximos:
 - Especializar la IA en el apoyo a los proyectos empresariales en ámbitos RIS3 de Euskadi y a los servicios públicos como la Salud, el Transporte o la Educación.
 - Convertir a Euskadi en un territorio donde desplegar y testear soluciones integradas completas y a escala real.
- ▶ Potenciar el Talento:
 - Incluir tecnologías digitales, en especial la IA, en el currículo educativo desde etapas tempranas
- ▶ Involucrar a las empresas tractoras locales y a sus cadenas de valor.
 - Conectar la aplicación de la IA con la estrategia de estas empresas.
 - Apoyar a los clústeres principales para arrastrar a las cadenas de valor locales, globales y plataformas tecnológicas.
- ▶ Afianzar una Estrategia basada en el dato sobre la que apoyar los proyectos de IA:
 - Estrategia del dato, ciberseguro, soberano, privado...
 - En el BIDH: la IA y la Ciberseguridad van unidos.
 - Vincular proyectos de datos públicos y privados.



- ▶ Actuar en red y con visión global:
 - Estar conectado a los nodos y flujos internacionales donde se mueve el conocimiento.
 - Buscar alianzas con países y regiones avanzadas.
 - Alianza franco-alemana para crear un centro virtual para el desarrollo de la IA.
- ▶ Focalizar los recursos:
 - En el caso de las empresas orientarlos a potenciar la Estrategia RIS3.
 - En el caso de las Administraciones Públicas desarrollar servicios públicos avanzados.

El BAIC cumple su misión a través del desarrollo de diferentes servicios que, conjuntados, permiten construir soluciones completas a los retos complejos que se le plantean.



Sergio Gil-López es el coordinador por TECNALIA de la iniciativa BAIC – *Basque Artificial Intelligence Center*.

Tiene un doctorado en CC. Físicas por la Universidad de Granada y es licenciado en CC. Físicas por la Universidad de La Laguna, con estancias en centros de investigación en Alemania. Es experto en modelado y simulación numérica, en técnicas de Inteligencia Artificial, *Machine Learning* y Optimización; desde computación evolutiva, problemas LP, NLP, MINLP, métodos heurísticos, meta-heurísticos, entre otros; esquemas de *clustering/grouping*, técnicas de reducción dimensional, entrenamientos desbalanceados, esquemas regresivos, *linear, lasso, ridge, elastic-net*, SVMs, RBFs, ELMs, MLP, CART, RF, métodos de ensamblaje (*bagging, boosting, stacking*), xAI, entre otros.

Ha conceptualizado y diseñado aplicaciones para los sectores de Industria y Energía. Ganador del DataSciAwards 2019 (iniciativa europea) y de los *Advanced Factory Awards 2020* en la Categoría de "I+D en IA aplicada en plantas industriales". Coautor de 7 patentes, ha realizado más de 37 contribuciones científicas en revistas internacionales, más de 43 contribuciones en congresos nacionales e internacionales, y ha llevado la codirección de tesis doctorales y de trabajos fin de carrera y másteres de doctorado en la Universidad del País Vasco.



02 Iniciativas singulares

Iniciativas y programas Relacionados



PLATAFORMA TECNOLÓGICA PARA LA PROMOCIÓN Y DIFUSIÓN DE LAS TECNOLOGÍAS DIGITALES

La Plataforma Española para la adopción y la difusión de las tecnologías digitales PLANETIC tiene como misión impulsar la innovación en el sector tecnológico digital español a través de una visión global e integrada de estas tecnologías, aunando esfuerzos y complementariedades. Representa a una muestra significativa del ecosistema digital con casi 270 organizaciones.

El objetivo final que persigue es acelerar la incorporación y adopción de las tecnologías digitales para impulsar la transformación social y competitiva del país. Para este objetivo, PLANETIC atiende a las necesidades en materia de transformación digital de los diferentes sectores económicos españoles, impulsa el posicionamiento internacional del sector tecnológico digital y sirve como espacio de trabajo común para agentes I+D+i y empresas del sector.

PLANETIC apoya el desarrollo e implantación de un amplio abanico de tecnologías y campos de actuación, desde las micro y nanotecnologías hasta los sistemas inteligentes, sistemas embebidos, software, servicios, Big Data o Inteligencia Artificial, entre otros.

ROL DE TECNALIA

Forma parte del Comité Ejecutivo desde sus orígenes. Es responsable de la Agenda Estratégica y co-lidera los Grupos de Trabajo de Big Data e Inteligencia Artificial, que ha producido el primer documento en español de Posicionamiento sobre Espacios de Datos, y el de Cloud.



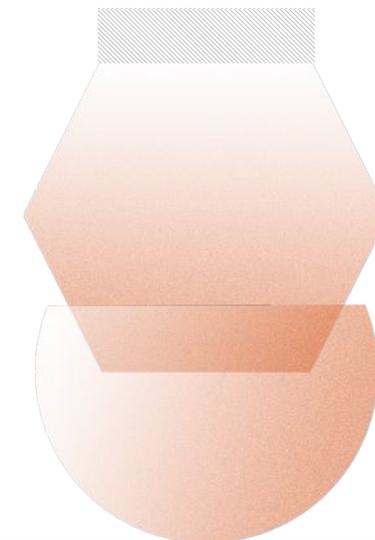
CLÚSTER BIG DATA DE MADRID

El Clúster Big Data Madrid está promovido por el Ayuntamiento de Madrid y reúne a entidades, tanto públicas como privadas, que desarrollan actividades en el territorio, y que se han asociado voluntariamente para impulsar el desarrollo del entorno Big Data y la Inteligencia Artificial. El Ayuntamiento contribuye a la Asociación con la aportación de recursos y servicios, y actúa como impulsor y dinamizador de una eficaz colaboración público-privada en esta iniciativa.



ROL DE TECNALIA

Forma parte de la Asociación desde sus orígenes y ha promovido la iniciativa dentro del clúster de "Espacio de datos federados para un Madrid + Inteligente". Durante el año 2021 se ha acometido el diseño de la arquitectura técnica para los servicios de este Espacio de Datos.





02 Iniciativas singulares

Iniciativas y programas Relacionados



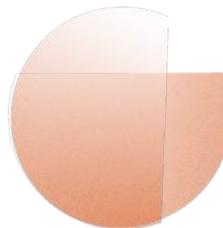
PROGRAMA HORIZONTE EUROPA

El Programa Horizonte Europa, con una dotación de 95.500 millones de euros en el periodo 2021-2027, es el programa de referencia en investigación y en innovación. El Clúster 4 Digital, Industria y Espacio, está orientado al apoyo del desarrollo y el dominio de las tecnologías habilitadoras digitales y clave del futuro con impacto, entre otros, en una economía de datos globalmente atractiva, segura y dinámica mediante el desarrollo de las tecnologías e infraestructuras de datos y de computación de próxima generación, con los espacios de datos correspondientes y permitiendo un ecosistema de inteligencia artificial.

ROL DE TECNALIA

TECNALIA ha sido la primera organización privada estatal en contratación, participación y liderazgo de proyectos en el Programa Horizon 2020 de la UE. Ocupa el puesto 13 entre 2.900 organizaciones europeas de investigación (con datos a marzo de 2020). Esta actividad le ha permitido generar una extensa red de socios y colaboradores en el contexto europeo. Actúa además con su capacidad de orquestación para involucrar en este tipo de programas a empresas, especialmente pymes, con limitadas capacidades para gestionar iniciativas en estos programas tan competitivos.

Con 524 empresas españolas | 451 proyectos contratados
178 millones de euros | 68 proyectos liderados
(datos a 31/12/2020)



Investing in the future:
Digital Europe Programme

PROGRAMA EUROPA DIGITAL

El nuevo Programa Europa Digital es un instrumento para facilitar la transformación digital de la sociedad europea y de su economía, potenciando las competencias digitales críticas y las infraestructuras que se precisan en la nueva sociedad digital. El Programa cuenta con un presupuesto previsto de 7.588 millones de euros para el periodo 2021-2027 y utilizará la contratación pública para desplegar estas tecnologías e infraestructuras digitales. Complementando a otros programas de apoyo a la transformación digital, financiará proyectos en cinco ámbitos interdependientes, cada uno de ellos dotado de su propio presupuesto indicativo:

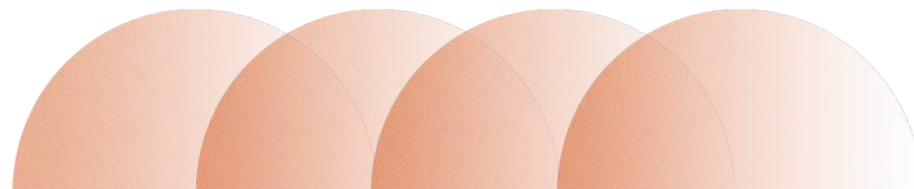
- ▶ Computación de alto rendimiento: 2.226.914.000 euros.
- ▶ Inteligencia Artificial: 2.061.956.000 euros.
- ▶ Ciberseguridad y confianza: 1.649.566.000 euros.
- ▶ Capacidades digitales avanzadas: 577.347.000 euros.
- ▶ Despliegue, mejor uso de la capacidad digital e interoperabilidad: 1.072.217.000 euros.

En el segundo de estos ámbitos se cita la "creación de espacios de datos comunes europeos que hagan que los datos sean accesibles en toda Europa, incluida la información recopilada a partir de la reutilización de la información del sector público, y se conviertan en una fuente de entrada de datos para las soluciones de IA. Los espacios deben estar abiertos al sector público y privado. Para facilitar su mayor uso, los datos dentro de un espacio deben ser interoperables, en particular a través de formatos de datos abiertos, legibles por máquina, estandarizados y documentados, tanto en las interacciones entre los sectores público y privado, dentro de los sectores y entre sectores (interoperabilidad semántica)". Aunque se encuentra todavía en versión borrador, se incluyen las siguientes convocatorias de interés:

- ▶ En la primera convocatoria, prevista en el segundo trimestre de 2021: 31 millones de euros para el soporte de los espacios de datos: salud – genómica, agricultura y pacto verde.
- ▶ En la segunda convocatoria, prevista en el último trimestre de 2021 y primero de 2022: 22 millones de euros para el soporte de los espacios de datos en fabricación.
- ▶ En la tercera convocatoria, prevista en el segundo trimestre de 2022: 68 millones de euros para el soporte de los espacios de datos: movilidad, salud - cáncer, seguridad y datos abiertos del sector público para IA.



03 Aportación tecnológica y proyectos



Páginas
6 – 9

01

Contexto y
motivación

Páginas
10 – 19

02

Iniciativas
singulares

Páginas
20 – 37

03

Aportación
tecnológica y
proyectos

Páginas
38 – 48

04

Casos de uso

Páginas
49 – 54

05

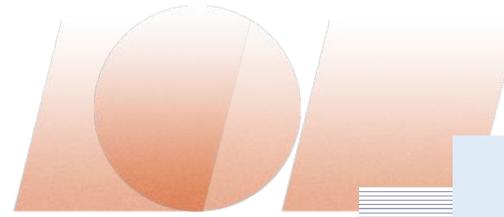
Equipo
TECNALIA

Páginas
55 – 56

06

Anexo

- Principales Retos para los Espacios de Datos
- Activos tecnológicos y proyectos TECNALIA relacionados



03 Aportación tecnológica y proyectos

Principales Retos para los Espacios de Datos

La construcción de los Espacios de Datos precisa de determinados servicios que se organizan dentro de arquitecturas técnicas concretas y que sirven para implementar y operar los requisitos de privacidad, seguridad, interoperabilidad, control de acceso y uso, valorización, etc. que caracterizan la visión más desarrollada de estos espacios. Su implementación posibilita ecosistemas de datos confiables, interoperables y soberanos para que los agentes participantes, tanto públicos como privados, desarrollen con la máxima eficacia y eficiencia servicios y aplicaciones inteligentes con impacto social, económico o medioambiental.

La implementación de estos servicios es una combinación de elementos legales, administrativos, organizativos y tecnológicos.



Confianza y seguridad

Gobernanza; Certificación de entidades; Acreditación; Trazabilidad; Certificación de componentes; Acceso identificado y seguro; Comunicaciones seguras.



Interoperabilidad

Catálogos y vocabularios; Conectores interoperables; Servicios de intermediación; Apps. facilitadoras



Soberanía y control

Gestión de políticas de acceso y uso; Compensación



Alberto Berreteaga, **investigador en TECNALIA:**
“las empresas cada vez generan más datos en sus procesos productivos. Esa información es valiosa y puede ser explotada como una nueva línea de negocio.”

EXPANSIÓN, 2/07/2021.

Alberto Berreteaga es Gerente del Laboratorio de I+D en ICT – Information and Communication Technologies y jefe de proyectos. Ingeniero de Software por la Universidad de Deusto, desarrolla su actividad de investigación en TECNALIA en el ámbito de los sistemas ciber-físicos, los espacios de datos y los entornos seguros DevOps para computación en la nube. Durante los últimos dos años ha participado en el desarrollo de componentes claves y Casos de Uso para la arquitectura IDS, una referencia europea en la implementación de espacios de datos. Como gerente del Laboratorio de I+D en ICT es responsable de sus infraestructuras de virtualización (VMWare vSphere, VirtualBox, AWS, Azure, Git, GitHub, Blockchain, etc.) para proveer soluciones a demostradores y componentes en el contexto de desarrollos de proyectos de I+D.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados

Desde el punto de vista tecnológico, existe y se está desarrollando una amplia colección de tecnologías que, combinadas, contribuyen a la realización de la visión completa de los Espacios de Datos.

TECNALIA está participando activamente en la I+D e implantación precomercial de algunas de estas tecnologías.





03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados

PRIVACY ENHANCING TECHNOLOGIES

En TECNALIA abordamos soluciones innovadoras para garantizar la privacidad y seguridad del dato desde la fase de diseño, aplicando estrategias que permiten:

- ▶ Minimizar la cantidad de información intercambiada o almacenada.
- ▶ Forzar las políticas de seguridad del dato mediante técnicas criptográficas y no confiar en acciones de terceros.
- ▶ Ofrecer control al propietario del dato para consultar, modificar o eliminar sus datos
- ▶ Ocultar la información tanto con técnicas de cifrado como mediante la pseudoanonimización o anonimización del dato en tránsito o almacenado.
- ▶ Separación de datos críticos del resto de información
- ▶ Agregación de datos en lugar de mantener el dato individual allí donde sea posible.

Algunos casos concretos son:

- ▶ **Identidad Digital.** A pesar de que en muchas ocasiones la identidad se asocia simplemente a un usuario/password, la identidad digital va más allá de las credenciales de acceso y se compone de todos los atributos que nos definen y componen nuestra identidad: edad, sexo, padrón, preferencias, etc. Todos esos atributos que nos definen en muchos casos pueden ser sensibles o, simplemente, un usuario puede desear minimizar los datos a compartir para acceder a un servicio o recurso. Técnicas como *Zero-Knowledge Proofs* permiten demostrar ciertas características de la identidad y son la base de los nuevos modelos de identidad auto-soberana (*Self-Sovereign Identity*) que devuelven el control al usuario sobre los atributos de su identidad. Igualmente, las técnicas de cifrado o firmas basadas en atributos (o roles) en lugar de identidades concretas de usuario, lo que aporta una mayor flexibilidad y privacidad a los sistemas.
- ▶ **Zero-Knowledge Proofs (ZKP).** Las pruebas de conocimiento cero permiten a una entidad convencer al verificador de la verdad de una declaración (generalmente matemática) sin revelar ninguna otra información que no sea la verdad de la afirmación. Si la declaración requiere que la entidad mantenga alguna información en secreto, sin que el verificador tenga acceso a la misma, matemáticamente es posible probar el conocimiento del secreto sin revelar el mismo. Las pruebas de conocimiento cero pueden ser interactivas o no interactivas, dependiendo de si las partes pueden comunicarse durante la prueba. En general, las pruebas no interactivas se consideran más complejas ya que no pueden utilizar protocolos reto-respuesta. Un uso típico de las pruebas de conocimiento cero es la autenticación de usuarios que poseen dispositivos o secretos criptográficos, como *smartcards*, sin revelar ninguna información sobre dichos usuarios salvo que poseen una tarjeta válida en el sistema.

- ▶ **Self-Sovereign Identity (SSI).** El modelo de identidad digital ha ido evolucionando. Inicialmente se optó por un modelo centralizado en el que cada proveedor de servicios recogía y gestionaba todos los atributos del usuario: esto requería que los usuarios recordaran una tupla usuario/contraseña por cada proveedor de servicios y que sus datos estuvieran repartidos por todos sus proveedores. Este modelo ha ido cediendo terreno al modelo de identidad federada en el que emerge una nueva figura, el proveedor de identidad, que gestiona la identidad/datos del usuario. El usuario puede centralizar así sus datos en uno de estos proveedores de identidad y logearse e intercambiar datos relativos a su identidad con terceros (servicios, aplicaciones, etc.). Este modelo, conocido como *Single Sign-On (SSO)*, se ha convertido en un estándar y los mecanismos de social *login* están presentes en la mayoría de servicios web. Con este modelo, los proveedores de identidad se han convertido en intermediarios en las relaciones de los usuarios con los proveedores de servicios digitales, y recopilan información y metadatos sobre la navegación y uso que los usuarios hacen más allá de sus servicios. Las grandes plataformas quieren convertirse en los proveedores de identidad únicos y procesar las identidades de usuarios y sus datos asociados. Estos proveedores no ayudan al usuario a gestionar su privacidad ni la compartición segura de su información personal, convirtiéndose además en un punto único de fallo del sistema, además de ser un gran atractivo para los cibercriminales.

Ante esta situación se está promoviendo un nuevo modelo descentralizado en el que se produce una relación directa entre el usuario y el servicio, sin que la identidad de una persona o entidad deba depender ni estar sometida a terceros: *Self-Sovereign Identity (SSI)*, identidad soberana o identidad autosoberana en castellano.

Estamos participando activamente en la normalización de la identidad soberana (ISO/TC 307), en la creación de la tecnología base desde Hyperledger (Aries, Indy) y en la difusión de la tecnología, participando en el primer libro sobre este nuevo modelo de identidad "*Self-Sovereign Identity: decentralized digital Identity and verifiable credentials*" de la editorial Manning. Ayudamos a los gobiernos y empresas de todo el mundo a adoptar los nuevos modelos de identidad digital para los usuarios, así como investigar la aplicabilidad de la tecnología SSI a otros ámbitos como el industrial, IoT o energético.

En TECNALIA hemos aplicado ZKP en sistemas de autenticación; actualmente trabajamos en la implantación de estas técnicas en diferentes sectores (administración pública, salud, etc.) para la privacidad e identidad soberana, minimizando la compartición de información de los usuarios de un sistema con sus proveedores.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados

PRIVACY ENHANCING TECHNOLOGIES (Cont.)

- ▶ **Attribute-Based Signatures & Encryption.** Los criptosistemas basados en atributos permiten generalizar los sistemas desarrollados gracias a sustituir las identidades de usuarios (firma) o contraseñas (cifrado) como credenciales únicas, sustituyéndolos por propiedades o atributos de los usuarios que serán utilizados como credenciales. En los esquemas de firma y cifrado basados en atributos, las claves secretas de los usuarios se generan en función de un árbol de acceso que define el alcance de los privilegios del usuario o en base a un conjunto de atributos directamente (KP-ABE vs CP-ABE). Tanto las firmas como el cifrado se produce en relación a la política de credenciales; por ejemplo, se podrán cifrar textos que únicamente puedan ser descifrados por los usuarios que cumplan la política en cuestión, como que el usuario tenga el rol de director general o sea responsable de negocio con una antigüedad mayor a 2 años en la empresa. Igualmente, a la hora de verificar una firma, podrá hacerse si el firmante cumple o no la política, con independencia de la persona que firma dicho texto.

Trabajamos en la aplicación de criptosistemas basados en atributos en el campo de la compartición de datos sensibles en entornos dinámicos y cambiantes que puedan beneficiarse de la utilización de políticas como, por ejemplo, el acceso a datos médicos, cifrado de datos o logs para su distribución a un conjunto grande de receptores o difusión en broadcast, etc.

Brokel.

En interés por la economía del dato y la Inteligencia Artificial está creciendo exponencialmente. Las empresas son conscientes del valor que generan los datos y del potencial de su compartición y explotación segura con terceros; pero el miedo y desconfianza a la hora de compartirlos con terceros se ha convertido en una gran barrera.

Brokel by TECNALIA contribuye a democratizar el acceso al dato, ofreciendo a las empresas y especialistas el acceso a más datos heterogéneos, permitiendo dar un salto de gigante en su explotación, poniendo en valor los datos aportados por cada empresa con el fin de poder obtener un mayor beneficio común y premiar a aquellos que aporten más datos y de mayor calidad al resto.

Brokel es el resultado de la aplicación de tecnologías de privacidad como la criptografía homomórfica o *secure multi-party computation*, entre otras, cuyo objetivo es la protección y ejecución segura de datos y servicios digitales. Se trata de garantizar la protección del dato a aquellas empresas que desean compartirlos con terceros, sin perder la soberanía y control sobre ellos. Al mismo tiempo, mediante criptografía avanzada se permite a los proveedores de servicios basados en el dato garantizar la confidencialidad de los algoritmos de Inteligencia Artificial con los que ofrecen servicios de optimización, predicción, etc. Brokel está en pruebas con el objetivo de que el tejido empresarial sea pionero en la compartición y explotación segura y soberana de sus datos para generar ventajas competitivas en escenarios como el de la Industria 4.0, la energía o la salud, entre otros.

Más información en: brokel.io



Óscar Lage es responsable de la línea de Ciberseguridad y Blockchain de TECNALIA.

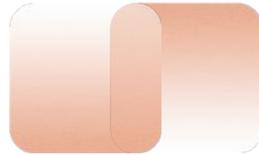
Gerente del primer laboratorio industrial blockchain de Europa, miembro del consejo asesor de varias compañías, conferenciante y coautor de varios libros, entre los que destaca el best seller "Blockchain: la revolución industrial de Internet". Es colaborador habitual en medios de comunicación generalistas (TV, radio, prensa).

Oscar es miembro de las dos principales alianzas internacionales sobre blockchain (*Enterprise Ethereum Alliance e Hyperledger*), coordinador regional de Alastria, vicepresidente de la comisión blockchain de AMETIC y miembro de los principales foros internacionales de ciberseguridad como la *European Cyber Security Organization (ECISO)* o el Centro de Ciberseguridad Industrial (CCI), entre otros. En los últimos años Óscar está centrando su actividad investigadora en las tecnologías denominadas *Privacy Preserving Technologies (PPT)*, con especial foco en las de Criptografía Homomórfica, *Multi-Party Computation (MPC)* o *Federated Machine Learning*.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados



COMPUTACIÓN EN CONTEXTO PRIVADO

El carácter sensible o privado de algunos de los datos que se desean compartir, como es el caso de los datos personales, es una de las barreras para el desarrollo de los espacios de datos y, en general, para el avance de la economía del dato. La GDPR en el caso de datos personales, o la sensibilidad de los datos de procesos industriales dificultan o impiden que otros actores puedan explotar u obtener nuevo conocimiento de ellos. Esto está dificultando que la Analítica Avanzada y la Inteligencia Artificial pueda conseguir unos mayores resultados y una mayor adopción.

La criptografía homomórfica o la computación multiparte permiten procesar datos de un tercero, o de un conjunto de actores de forma totalmente segura. En ambos casos es posible ejecutar diferentes algoritmos estadísticos o técnicas de *Machine Learning* o *Deep Learning* sobre datos cifrados, cuyo resultado será únicamente accesible por las partes acordadas, respetando siempre la confidencialidad de los datos de cada parte.

Las técnicas de aprendizaje federado (*Federated Learning*) permiten por su parte abordar casos de entrenamiento de modelos de Inteligencia Artificial en ecosistemas que quieren compartir el modelo resultante sin compartir sus datos o sin que estos salgan de sus organizaciones, siendo el modelo resultante un bien común para todos.

De hecho, la combinación de los dos primeros protocolos (criptografía homomórfica y computación multiparte) con las técnicas de *Federated Learning* permite obtener la seguridad de los primeros protocolos con la agilidad que proporciona *Federated Learning*.

TECNALIA está investigando actualmente en la utilización de criptosistemas homomórficos para la compartición y explotación segura de datos entre diferentes participantes. Gracias a la criptografía homomórfica es posible proteger tanto el algoritmo/modelo, en base al cual se ofrece un servicio basado en el dato, como los datos sobre los que se ejecuta dicho algoritmo. Se protege así tanto al proveedor como al consumidor de servicios basados en datos, en campos como la predicción de consumo energético, el mantenimiento predictivo, el análisis de ADN, etc.

- ▶ **Criptografía homomórfica.** Existen diferentes criptosistemas que presentan propiedades homomórficas, es decir, el resultado de la ejecución de una operación algebraica sobre un texto plano es equivalente a descifrar el resultado de la ejecución de dicha operación algebraica (no necesariamente la misma) sobre el texto cifrado. Así pues, el principal objetivo de los criptosistemas que presentan esta propiedad es realizar operaciones directamente sobre los datos cifrados, sin la necesidad de descifrarlos previamente para operar sobre los mismos, ni de contar con la clave con la que fueron cifrados para poder operar. Hay que tener en cuenta que el resultado siempre va a tener un mínimo nivel de ruido que irá incrementándose con cada operación que se realice sobre el texto cifrado. No obstante, gracias a las técnicas de bootstrapping, es posible controlar dicho ruido para que sea mínimo.

Los esquemas que presentan propiedades homomórficas para una única operación se conocen como *Partially Homomorphic Encryption* (PHE) y en su mayoría permiten un número de operaciones ilimitada sobre el texto como ElGamal o Paillier, entre otros. Finalmente, *Fully Homomorphic Encryption* (FHE) permite la evaluación de funciones arbitrarias y un número de operaciones ilimitado, eso sí, con una elevada carga computacional y una expansión de almacenamiento importante frente a los esquemas anteriores. Por esta razón, a veces se prefieren criptosistemas menos potentes conocidos como *Somewhat Homomorphic Encryption* (SWHE), que permiten varios tipos de operación con una menor carga computacional y requisitos de almacenamiento, pero con un número limitado de operaciones sobre el texto.



Cristina Regueiro es doctora en Tecnologías de la Información y Comunicaciones en Redes Móviles e Ingeniera de Telecomunicaciones por la Universidad del País Vasco.

Tiene una amplia experiencia en investigación en comunicaciones inalámbricas y ciberseguridad habiendo participado en numerosos proyectos de financiación pública y privada. Comenzó como Investigadora en la Universidad del País Vasco en temas relacionados con el procesado de señales digitales, comunicaciones inalámbricas y comunicaciones móviles. Continuó su labor investigadora en la Asociación Innovalia enfocándose más en las comunicaciones industriales, Industria 4.0, IoT, *fog* y *cloud computing*. Este enfoque lo completó posteriormente en IKERLAN con conocimientos más enfocados a ciberseguridad, formando parte del personal investigador de los equipos de IoT Ciberseguro y Ciberseguridad en Plataformas Digitales. En la actualidad desarrolla su actividad investigadora dentro del equipo de Ciberseguridad y Blockchain de TECNALIA, conceptualizando y desarrollando soluciones de ciberseguridad y sistemas basados en blockchain. Dirige en TECNALIA el proyecto *5G Euskadi*.



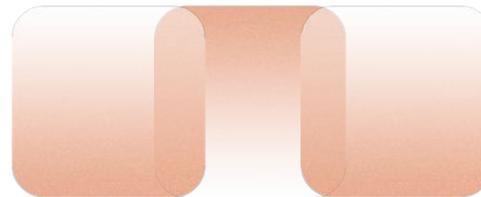
03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados

COMPUTACIÓN EN CONTEXTO PRIVADO (Cont.)

- ▶ **Multi-Party Computation (MPC).** Los protocolos MPC permiten a un conjunto de participantes calcular una función conjunta de sus entradas de forma segura sin la necesidad de un tercero de confianza. Durante la ejecución del protocolo, las partes no aprenden nada sobre las entradas de los demás, excepto lo que implica la propia salida parcial. A diferencia de los criptosistemas tradicionales, estos protocolos también consideran adversarios al resto de participantes que participan en el protocolo; de hecho, los protocolos utilizados en casos en los que la mayoría de participantes se considera honesta son diferentes a los entornos en los que no es posible garantizar un mínimo de participantes honestos. Los protocolos entre dos partes, *two-party computation* (2PC) son un subconjunto específico de los protocolos MPC en los que únicamente participan dos entidades. Existen protocolos específicos para este tipo de problemas y en muchos de ellos se crean canales privados, algo que no se suele utilizar con protocolos MPC en los que participan más participantes y que se basan en la compartición de secretos. Además, en este tipo de protocolos *multiparty* los participantes no suelen tener un rol específico. Uno de los principales retos de los protocolos MPC es reducir su elevado coste computacional y la falta de flexibilidad en las redes que la mayoría de implementaciones crean de forma específica para la ejecución del protocolo.

Investigamos y desarrollamos aplicaciones de los protocolos MPC para la explotación y compartición segura de datos, permitiendo la ejecución de algoritmos o entrenamientos totalmente seguros. Diseñamos nuevos esquemas que permitan garantizar la finalización de los procesos MPC de forma parcial. Combinamos la tecnología MPC con otras tecnologías como *Federated Learning*, o *blockchain* para mayor dinamismo y eficiencia.



- ▶ **Federated Learning.** El aprendizaje federado permite la colaboración entre diferentes participantes para el entrenamiento de un algoritmo de inteligencia artificial sin que las partes compartan sus datos de forma directa. Cada miembro del ecosistema realiza un entrenamiento parcial que comparte con el resto, de forma que se puedan juntar los entrenamientos parciales de cada parte y poner en común para que juntos puedan resolver mejor un problema sin que deban compartir los datos de las partes.

Existen dos tipologías principales de aprendizaje federado: horizontal y vertical. En el primero, federación horizontal, los conjuntos de datos comparten el mismo espacio de características, pero son diferentes en la muestra. En el aprendizaje vertical, o aprendizaje basado en características, se aplica en los casos en los que dos conjuntos de datos comparten el mismo espacio de IDs de muestra pero difieren en el espacio de características. Aunque sin todavía aplicaciones en producción conocidas, cabe destacar otro modelo conocido como *Federated Transfer Learning* que permitiría la transferencia de conocimientos entre dominios que no tienen muchas características y usuarios que se solapan.

En TECNALIA trabajamos en diferentes técnicas híbridas para integrar eficientemente en una misma arquitectura blockchain y las técnicas de FML, dando cabida a la trazabilidad continua y a la descentralización absoluta que añade mayor resiliencia al ecosistema. Además, estamos mejorando la seguridad del aprendizaje federado mediante la utilización de técnicas homomórficas y MPC en conjunto con el aprendizaje federado.



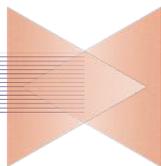
Esther Villar es doctora (Cum Laude) en Tecnologías de la Información y la Comunicación de la Universidad de Alcalá, una licenciatura en Ciencias de la Computación por la Universidad de Deusto y un máster en Lenguajes y sistemas informáticos por UNED (Universidad Nacional de Educación a Distancia).

Actualmente es científica de datos senior y coordina la Plataforma en Inteligencia Artificial en TECNALIA. Se especializó en modelos de aprendizaje superficial y aprendizaje automático profundo, estrategias de aprendizaje distribuidas con conciencia de privacidad (incluido el aprendizaje federado), aprendizaje reforzado y detección de valores atípicos, entre otros. Es autora de varias contribuciones en conferencias y artículos en revistas relacionadas con estas áreas de investigación, con enfoque a su aplicabilidad a problemas prácticos: predicción y optimización de procesos industriales y de gestión, planificación de recursos, programación, eficiencia energética y otras aplicaciones variadas por igual.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados



OTRAS TÉCNICAS DE SEGURIDAD Y PRIVACIDAD DEL DATO

Existen otras técnicas y protocolos de privacidad y seguridad del dato que también consideramos relevantes en nuestra investigación.

- ▶ **Differential Privacy.** Otro modelo de privacidad relevante es la privacidad diferencial. Se definió originalmente para bases de datos que reciben consultas sobre datos sensibles y se basa en la perturbación del resultado de la consulta original antes de compartirlo con terceros. Posteriormente también se ha aplicado esta misma técnica para perturbar una copia de la base de datos y ejecutar las consultas sobre los datos ya alterados. El objetivo principal de *Differential Privacy* es que el resultado de la consulta a un conjunto de datos no pueda utilizarse para inferir conocimiento sobre un individuo concreto, limitando así la inferencia sobre información privada de los individuos. Para ello se introduce ruido no determinista (aleatorio) que altera los resultados unitarios de la búsqueda pero que no supone una pérdida significativa de precisión en el procesamiento posterior de la muestra con fines estadísticos.

En TECNALIA trabajamos en la aplicación de la privacidad diferencial en dominios industriales como el energético, la automoción o el manufacturero, así como en aplicaciones innovadoras en el ámbito de la salud.

- ▶ **Blockchain y las tecnologías DLT (Distributed Ledger Technologies).** Son nuevas familias de registros que, a diferencia de los tradicionales en los que una única organización vela por la coherencia y confianza de la información albergada en la misma, la información es gobernada por una red o comunidad. Es por ello que cada vez que un participante de dicha comunidad desea registrar cualquier dato su propuesta será analizada por el resto de participantes de la red, y requerirá del consenso de la mayoría para poder formar parte de dicho registro. Precisamente, tras dicho consenso, esa información será inmutable y únicamente se podrán ir anotando nuevas "versiones" de dicha información. Por esta naturaleza singular, blockchain aporta un valor diferencial a muchos proyectos, especialmente aquellos en los que intervienen en un proceso diferentes organizaciones, es decir una comunidad, que tienen que establecer relaciones de confianza entre ellos. Algunas de las características diferenciales de blockchain son:

- Descentralización de procesos y modelos de negocio.
- Trazabilidad y transparencia del registro.
- Sincronización de procesos ofreciendo una visión única consensuada de la información relativa a los activos intercambiados.
- Automatización de procesos gracias a ofrecer un registro sincronizado, consensuado, inalterable y sobre el que podemos tener un no repudio de la información (cada participante firma sus transacciones).

Blockchain permite crear ecosistemas descentralizados sin un punto único de fallo en el que, por ejemplo, se puedan crear espacios de datos descentralizados. Además, su inalterabilidad está siendo utilizada como registro de confianza de la Inteligencia Artificial explicable (XAI), o garantizando la procedencia, calidad y seguridad del dato utilizado para entrenar algoritmos de IA. Por último, cabe destacar cómo blockchain ofrece una sincronización perfecta del dato compartido que posibilita tomar decisiones sobre activos compartidos.

En TECNALIA estamos investigando diferentes aplicaciones de la tecnología blockchain para la creación de ecosistemas abiertos de compartición y explotación del dato de forma segura, efectuando funciones de *clearing* y bróker en dichos ecosistemas, así como el uso de blockchain como registro de evidencias.



Blockchain Lab y Laboratorio de Ciberseguridad y Seguridad Industrial (ICSSY)

TECNALIA cuenta con dos laboratorios de ciberseguridad que pertenecen también al Basque Digital Innovation Hub para avanzar en el estado del arte y promover la adopción de soluciones de seguridad cibernética para usuarios industriales y científicos que aumenten la resistencia ante las amenazas, reduzcan los riesgos de seguridad y garanticen el cumplimiento de los estándares industriales.

- ▶ Laboratorio de Blockchain.
- ▶ Laboratorio de Ciberseguridad y Seguridad Industrial (ICSSY). Este lab forma parte de la Red Española de Laboratorios de Ciberseguridad Industrial (RNLI). Dispone de una gran base de estudios de casos y de:
 - Herramientas de software: gestión de garantía de seguridad / protección.
 - Soluciones de seguridad / protección: WhiteZone, Panzer.
 - Herramientas de prueba y evaluación: simulación, inyección de fallas, pruebas de penetración, herramientas SIEM.



03 Aportación tecnológica y proyectos

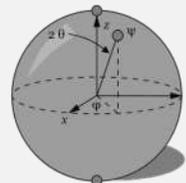
Activos tecnológicos y proyectos TECNALIA relacionados

OTRAS TÉCNICAS DE SEGURIDAD Y PRIVACIDAD DEL DATO (Cont.)

► **Criptografía Cuántica y postcuántica.** Hasta dentro de unos años no se espera la disponibilidad de ordenadores cuánticos generalizados. Sin embargo, las comunicaciones y datos sensibles que se intercambian actualmente podrían verse comprometidos más adelante cuando estos ordenadores cuánticos estén disponibles. Es posible que atacantes puedan estar recopilando información con vistas a esta disponibilidad.

Con el objetivo de evitar el impacto del ordenador cuántico sobre la seguridad tradicional surgen dos líneas de trabajo complementarias. La primera de ellas trata de buscar o identificar criptosistemas de base matemática que puedan ser utilizados por los ordenadores tradicionales, pero resistentes a ataques cuánticos, es decir la criptografía postcuántica. Precisamente, una de las iniciativas más reconocidas en la búsqueda de algoritmos postcuánticos es la iniciativa *Post-Quantum Cryptography* del NIST que actualmente se encuentra en su tercera ronda y que propone una serie de criptosistemas que no se verían afectados por dichos ataques. Por otro lado, la otra línea de trabajo se basa en la utilización de propiedades físicas para crear nuevos protocolos criptográficos. En este caso, estos protocolos deberán ser ejecutados por un ordenador cuántico y su seguridad se basará en las leyes físicas en lugar de matemáticas. En esta línea la generación de números aleatorios, así como las comunicaciones cuánticas, y en concreto *Quantum Key Distribution* (QKD) son las líneas de mayor relevancia.

Estamos analizando los criptosistemas candidatos de la iniciativa *Post-Quantum Cryptography* del NIST en diferentes entornos, en especial IoT, e introduciéndolos en los diferentes protocolos TI e industriales. Investigamos en nuevos protocolos QKD, así como la introducción de los existentes en los actuales protocolos TI e industriales.



QUANTEK

Las tecnologías cuánticas se están desarrollando con gran rapidez a partir de la disponibilidad de ordenadores cuánticos en la nube y con el avance de aspectos relacionados con las comunicaciones, como el intercambio de claves (*Quantum Key Distribution* – QKD).

La iniciativa colaborativa QUANTEK cuenta con el apoyo del Programa ELKARTEK del Gobierno Vasco para impulsar la investigación científica, el desarrollo tecnológico y la excelencia de los agentes de la red vasca de ciencia y tecnología más relevantes en torno a las tecnologías cuánticas.

Los socios de la red de excelencia están desarrollando tecnologías para la adopción de enfoques cuánticos que puedan facilitar la transición de las empresas, explorando la viabilidad de las tecnologías cuánticas para la resolución de algunos de los grandes retos de la industria y favoreciendo la generación de un ecosistema vasco en tecnologías cuánticas y sus aplicaciones industriales. Las áreas de trabajo son:

- Adaptación de la Ingeniería del Software al paradigma cuántico.
- Codificación de algoritmos clásicos para su ejecución sobre ordenadores cuánticos.
- Seguridad y comunicaciones cuánticas.



Xabier Larrucea se doctoró sobresaliente cum laude por la Universidad del País Vasco.

Experto en técnicas de privacidad del datos como *differential privacy* y *k-anonymity*.

Actualmente dirige el proyecto QUANTEK relacionado con tecnologías cuánticas. Es jefe de proyectos (PMP) e investigador senior (*IEEE Senior member*) en TECNALIA. Lidera proyectos e iniciativas relacionadas con la ciberseguridad como el proyecto SHIELD, *European Security in Health Data Exchange*. Pertenece al *Board* de la revista *IEEE Software*, y es Editor Asociado de revistas como *IEEE Software*, *IET Software* e *IEEE Access*. Ha realizado estancias doctorales en el *Software Engineering Institute*, Pittsburgh, USA. Esta actividad la compagina con su puesto de profesor asociado en la Escuela de Ingeniería de Vitoria-Gasteiz (EHU/UPV) donde imparte asignaturas del área Lenguajes y Sistemas informáticos.

Además dirige tesis doctorales, trabajos fin de master y fin de grado en la Escuela de Ingeniería de Bilbao de la UPV/EHU.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados



COMPUTACIÓN EN LA NUBE

Las tecnologías *cloud* son facilitadores clave en la economía del dato. Es preciso, sin embargo, seguir avanzando para superar algunos de los retos como la selección y composición optimizada de servicios *cloud* seguros, el despliegue automatizado y orquestación de servicios, así como la operación continua tanto de los servicios como de las aplicaciones desplegadas en dichos servicios:

- ▶ El reto sobre la selección de servicios *cloud* está relacionado con la comparabilidad entre temas tales como las capacidades funcionales (p.ej. disco), las no funcionales (p.ej. localización, precio), o el cumplimiento del marco regulatorio europeo, como la RGDP o el esquema europeo de certificación *cloud* (EUCS).
- ▶ Otro reto al que se enfrenta un proveedor de una aplicación como servicio tiene que ver con el despliegue automático de los diferentes componentes de la aplicación en los servicios *cloud* seleccionados o, si se decide por un despliegue *multi-cloud*, en la combinación optimizada de servicios *cloud* seleccionados. Este despliegue automático requiere de la capacidad de provisionar y configurar automáticamente los servicios y recursos *cloud*, así como su orquestación, tratando la infraestructura como software, en lo que se denomina actualmente como *Infraestructura como código* (IaC). La plétora de lenguajes existentes en este dominio requiere a los desarrolladores aprenderlos para cada una de las fases importantes en el despliegue.

- ▶ Asegurar la continuidad del servicio es uno de los elementos críticos para la adopción de *cloud computing*, en todas sus capacidades – infraestructura, plataforma y aplicaciones. Para ello es necesario disponer de mecanismos que monitoricen de manera continua los servicios e infraestructura, los comparen con los SLAs (*Service Level Agreement*) acordados y alerten de cualquier violación con respecto al valor negociado / acordado para lanzar un nuevo redespiegue automático u otra estrategia de (auto) recuperación que se considere apropiada. Sin embargo, en aplicaciones críticas, ser capaces de responder reactivamente a un problema no es suficiente y es necesario detectar anomalías que pueden llevar a fallas en el servicio o en la infraestructura auto-aprendiendo (*self-learning*) de las situaciones erróneas ocurridas en el pasado. El uso de técnicas de Inteligencia Artificial facilita estas actividades de predicción en lo que se denomina "*cognitive cloud*".



Dr. Iñaki Olabarrieta es licenciado y Doctor en Física Computacional por la Universidad de British Columbia (Canadá) en 2004, y Postdoctorado en el Instituto Hearne de Física Teórica de la Universidad de Louisiana (USA), centrado en el modelado de sistemas físicos complejos mediante métodos informáticos.

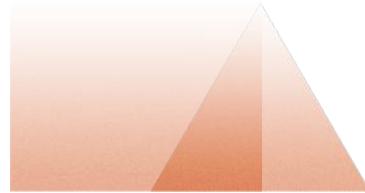
Investigador en TECNALIA desde 2006 y actual responsable científico del *Smart Mobility Lab* (sml.tecnalia.com), en donde ha sido también responsable del diseño de la arquitectura del sistema.

Es autor de varias patentes, múltiples artículos de investigación y tiene una gran experiencia en el desarrollo de proyectos para clientes cubriendo todas las etapas del procesamiento del dato desde la captación, el almacenamiento, procesamiento y visualización .



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados



COMPUTACIÓN EN LA NUBE

- La seguridad es otro de los puntos clave en la lenta adopción de *cloud computing*, de acuerdo con Eurostat y ENISA. Con el objetivo de generar confianza en servicios *cloud*, la Comisión Europea publicó el Cybersecurity Act que establece la creación de un marco de certificación de productos, procesos y servicios TIC. El esquema de certificación de servicios *cloud* (EUCS) es el segundo esquema en desarrollo, siendo el primero el de *Common Criteria*. EUCS define tres niveles de certificación, exigiendo en el nivel alto una monitorización y auditoría continua del cumplimiento de los requisitos de seguridad. Actualmente, la auditoría y certificación continua son temas poco maduros, ya que existen pocas herramientas independientes del proveedor *cloud* que puedan monitorizar los diferentes controles y requisitos de manera automática. Otro reto está relacionado con la posible manipulación de las evidencias recogidas, así como el ciclo de vida del certificado.



MEDINA proporciona un marco de certificación continua basada en auditorías para los proveedores de servicios de nube según el esquema de CSA de la UE, para la certificación de seguridad en la nube. Ofrece funcionalidades como la validación y las pruebas de seguridad, define un lenguaje de certificación legible por máquina, evalúa el desempeño de seguridad en la nube y gestiona las evidencias para la auditoría.

Para contribuir al primer reto, selección de servicios *cloud*, TECNALIA ha desarrollado un catálogo avanzado federado de servicios y recursos infraestructurales que permite descubrir qué servicio o combinación optimizada de servicios *cloud* es la más adecuada para las necesidades del desarrollador y operador de la aplicación en términos de disponibilidad, rendimiento, localización, cumplimiento del marco legal, coste, historial de violaciones del Acuerdo de Nivel de Servicio, valores medios monitorizados en determinadas métricas, etc. Además, TECNALIA participa en varios grupos de trabajo de GAIA-X relacionados con estos aspectos como por ejemplo "federated catalogue" o "compliance".

En cuanto al despliegue automático, el equipo de TECNALIA está desarrollando un lenguaje de modelado, DOML, que permite abstraer la complejidad de dichos lenguajes, así como de un gestor que permite ejecutar IaC de manera automática cubriendo todas las fases del despliegue, desde la provisión y configuración del servicio hasta el despliegue y orquestación.

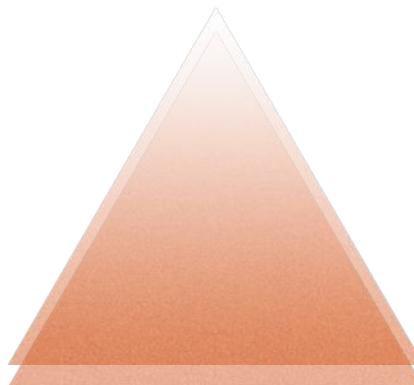
Para la continuidad del servicio, TECNALIA ha desarrollado componentes de monitorización de requisitos no funcionales de infraestructuras *cloud* y actualmente se encuentra extendiendo dicho componente con mecanismos de predicción de fallos y detección de anomalías.

En el ámbito de ciberseguridad *cloud* TECNALIA ha tenido un rol relevante, realizando primero un estudio para la Comisión Europea (SMART 2016/0029) y participando posteriormente en diferentes grupos de expertos (CSPCERT, ENISA *Ad-hoc Working Group*, *Stakeholder Cybersecurity Certification Group*) que han definido tal esquema de certificación (ENISA).



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados



COMPUTING CONTINUUM

En los últimos tiempos se está produciendo un avance paralelo y realimentado entre las tecnologías para el procesamiento masivo de los datos y las aplicaciones que hacen uso de estos. La tecnología habilita nuevos casos de uso que hasta ahora no habían sido posibles y, al mismo tiempo, se están explorando nuevas clases de aplicaciones que combinan el procesamiento de datos masivos en tiempo real con modelos complejos de Inteligencia Artificial, que tensionan de forma permanente el desarrollo de nueva tecnología.

Los datos deben ser pre procesados antes de ser transmitidos para maximizar la eficiencia de los sistemas y los requisitos de tiempo real de muchas aplicaciones hacen necesario desplazar los recursos de procesamiento a las proximidades en donde ocurren las cosas para garantizar los tiempos de latencia requeridos (*Edge Computing*). Al mismo tiempo, los recursos Cloud siguen siendo necesarios para cubrir eficientemente algunos aspectos de los sistemas, y el despliegue de redes 5G está poniendo a disposición de las arquitecturas IT nuevos recursos de procesamiento, no solo en el extremo (*Multiaccess Edge Computing* – MEC), también en puntos intermedios de la ruta de los datos (*Fog Computing*).

El resultado es la integración de los paradigmas Cloud, Edge y Fog de forma fluida para acomodar flujos de datos dinámicos en un concepto que se conoce como *Computing Continuum*, conjugando la conectividad avanzada y la computación con el razonamiento y la automatización basados en la Inteligencia Artificial.

La arquitectura IT de un sistema que utiliza la analítica de datos masivos en tiempo real puede resultar muy compleja por la necesidad de integrar todos estos niveles de procesamiento comunicándolos mediante canales de gran ancho de banda y baja latencia, o la conexión de cantidades masivas de dispositivos inteligentes.



Raúl Miñón es doctor (cum laude) internacional en informática por la Universidad del País Vasco, habiendo realizado una tesis enfocada en la generación automática de páginas web adaptadas al contexto y a las necesidades de personas con discapacidad.

Tiene un máster en sistemas de información avanzados y es ingeniero en informática por la Universidad del país Vasco, con estancias de investigación en el CNR, Pisa, Italia. Ha trabajado en la industria como *full stack developer* y ha promovido una startup en su campo de especialización. Desde 2016 trabaja en Investigación y Desarrollo tecnológico en las áreas de Big Data, DevOps, Cloud, *Edge Computing* y MLOps. Sus actuales intereses de investigación se centran en *edge computing* y en la gestión del ciclo de vida de los modelos matemáticos.

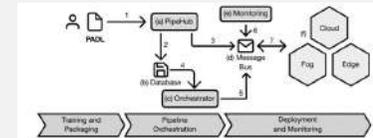


03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados

COMPUTING CONTINUUM

Además, el sistema tiene que ser capaz de orquestar recursos de proveedores de naturaleza muy diferente, combinando los recursos ofrecidos por los proveedores *cloud* con los recursos aportados por los operadores de redes 5G y los recursos *on premise*. Por último, los retos relacionados con la ciberseguridad que se desprenden de estas arquitecturas distribuidas representan también un desafío importante para la administración de estos sistemas, ya que a menudo se están gestionando procesos que son críticos para las empresas.



PANGEA by TECNALIA, Despliegue automático de procesos analíticos en el computing continuum

El activo PANGEA es una herramienta que, partiendo de una definición de un proceso analítico (*pipeline*) en formato PADL (*Analytical Pipeline Definition and Deployment Language*), una descripción de máquinas *edge*, *fog*, *cloud* y un *matching* entre ellas, permite desplegar cada paso del pipeline o de una carga de trabajo de manera automática en cada uno de los nodos establecidos en una infraestructura en producción. Además, si alguna máquina de la nube no existiese, la puede crear también automáticamente en la infraestructura proporcionando las claves requeridas. Finalmente, tiene un modo desarrollo, donde se genera un *docker compose* preparado para realizar pruebas en el entorno local.

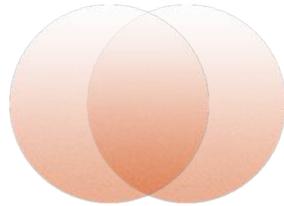
Desde TECNALIA aportamos innovación y tecnología en Arquitecturas TI avanzadas y ciberseguras para albergar modelos de computación en la nube en sistemas Big Data en tiempo real:

- ▶ Infraestructuras open source para Big Data, optimizadas para entornos híbridos y multinube, con capacidades de gestión de datos en nubes públicas, privadas y sin sistema operativo.
- ▶ Mecanismos de persistencia para Big Data: nuevos modelos de datos NoSQL; repositorios de alma-cenamiento centralizados con datos de varias fuentes en un formato granular y sin procesar, estructurados, semiestructurados o no estructurados (concepto Data Lake).
- ▶ Paradigmas Big Data (*real-time/batch*), arquitecturas Lambda, Kappa, ELK... y modelos de programación distribuidos.
- ▶ Gobierno del dato: Ciclo de vida, Seguridad Big Data, IDS...
- ▶ DevOps: Operacionalización de la IA (MLOps) e *Infrastructure as Code* (IaC).
- ▶ Sistemas distribuidos, IoT e interoperabilidad. Gestión de información. Interoperabilidad sintáctica, semántica y organizacional. Arquitecturas basadas en *middleware* (MQTT y DDS)
- ▶ Comunicaciones IoT en entornos industriales, entre máquinas, operarios, dispositivos IoT y vehículos autónomos heterogéneos, incluyendo las nuevas redes 5G.
- ▶ Gestión y visualización analítica de fuentes de datos. Procesamiento y agregación de datos heterogéneos. Herramientas de visualización de modelos analíticos.
- ▶ Ciberseguridad de arquitecturas avanzadas Computing Continuum. Seguridad de despliegues con comunicaciones 5G e integración de MEC y otros recursos de computación.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados



INTEROPERABILIDAD Y GOBIERNO DE DATOS

La implementación de los aspectos de interoperabilidad y gobernanza en el contexto de los Espacios de Datos se aborda con diferentes aproximaciones y tecnologías: IDS, FiWare, MyData, etc. En TECNALIA estamos trabajando específicamente en el desarrollo de tecnologías y componentes de interconexión de datos y de gestión de políticas de uso, un aspecto clave con relación a la soberanía:

- ▶ **Conectores IDS.** El conector es el componente principal de una arquitectura IDS y sobre el que se canalizan los diferentes servicios de las partes. Proporciona la conectividad estándar en el ecosistema habilitando las puertas entre las diversas entidades que intervienen. Es el elemento responsable del intercambio de datos entre los participantes del ecosistema, y del control de acceso y uso de dichos datos. Proporciona metadatos para su autodescripción (descripción técnica de la interfaz, mecanismo de autenticación, datos expuestos y políticas de uso) al intermediario principal, el Proveedor de Servicio de Catálogo o directorio (que en terminología IDS se denomina Broker). Cada conector debe tener un identificador único y un certificado válido. Además, cada conector debe ser capaz de verificar la identidad de los otros conectores.

Cada entidad Proveedora de Datos en un ecosistema con la arquitectura de referencia IDS debe ser capaz de asegurar que sus datos son manejados por el Conector de la entidad Consumidora de Datos de acuerdo con las políticas de acceso y uso especificadas; en caso contrario, no se dará acceso a los datos. Por tanto, una de las principales misiones es que se garanticen las políticas de acceso y uso de los activos de datos asociados

ESPACIOS DE DATOS CON ARQUITECTURA IDS

En TECNALIA se han diseñado y desarrollado arquitecturas que garantizan la confianza, interoperabilidad y seguridad para diferentes casos de uso:

- ▶ **Eólica OffShore:** en el proyecto DAEKIN se ha desarrollado una arquitectura de referencia basada en el ecosistema IDSA orientada a la creación de una plataforma de compartición de datos para el sector eólico *off shore*, promovida por el *Clúster Vasco de la Energía*. Se trata de uno de los sectores con mayor representatividad de empresas locales en la cadena de valor que puede experimentar un notable impulso integrando los datos en la cadena de valor.



- ▶ **Madrid + Inteligente:** Diseño de la arquitectura de un Espacio de Datos federados para el territorio de Madrid (Ayuntamiento de Madrid, Clúster Big Data del Ayuntamiento de Madrid).



Erik Maqueda es analista de datos y gestor de proyectos en el equipo de *Digital Energy and Environment* donde participa en distintos proyectos de I+D para la aplicación de tecnologías digitales emergentes en los sectores de la Energía y el Medio Ambiente.

En particular, actualmente es el coordinador técnico del proyecto H2020 PLATOON donde se están construyendo las bases para un espacio de datos de energía a nivel europeo. Además, participa activamente en distintos foros relacionados con los espacios de datos como IDSA, OPEN DEI, etc. Anteriormente, ha trabajado en ROLLS-ROYCE y LUFTHANSA TECHNIK en el desarrollo de herramientas inteligentes basadas en datos de mantenimiento predictivo para turbinas de gas de aeronaves.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados

INTEROPERABILIDAD Y GOBIERNO DE DATOS (Cont.)

- ▶ **Interoperabilidad semántica.** La interoperabilidad semántica se encarga de asegurar que el significado de la información intercambiada es comprensible de manera unívoca por cualquier otra aplicación que no fue inicialmente desarrollada para este propósito. Así, los sistemas son capaces de combinar la información recibida con otros recursos y procesarla explotando su semántica subyacente. Las ontologías son uno de los mecanismos con mayor adopción para este fin. La anotación semántica de datos crudos basándose en ontologías o vocabularios existentes permite una mejor representación de los datos, estructurándolos y estableciendo tipos, relaciones, propiedades y restricciones formales entre ellos. Además, los datos provenientes de múltiples fuentes se representan de una manera unificada, facilitando así su integración. Otro beneficio de la anotación semántica estriba en la incorporación de conocimiento de dominio al conjunto de datos, lo que contribuye a su enriquecimiento. La anotación semántica facilita la aplicación de técnicas de indexado, así como la realización de consultas más detalladas de los datos y la navegación a través de recursos relacionados.

En TECNALIA trabajamos en diferentes técnicas híbridas para gestionar el ciclo de vida completo del dato en diferentes dominios (Industria 4.0, salud, agro, *smart cities*...), desde su captación según los estándares y protocolos apropiados (MQTT, OPC-UA, DDS, ISOBUS...), pasando por los procesos de preparación, anotación y transformación automática (R2RML, JSON-LD, DCAT-AP) hasta su almacenamiento, visualización y explotación, garantizando la interoperabilidad de las soluciones en los tres niveles técnico, semántico y organizacional.



HODEI-X

HODEI-X es una red de investigación, financiada por el Gobierno Vasco a través del programa ELKARTEK, para fortalecer las capacidades del ecosistema de la CAPV con relación a la gestión de infraestructuras y explotación de datos, y para articular su participación en un eventual nodo de GAIA-X.

El equipo de trabajo HODEI-X está contribuyendo a la arquitectura general de GAIA-X con el desarrollo de pruebas de concepto de componentes y la validación en casos de uso en sectores estratégicos de la economía e industria vascas. Las contribuciones a la arquitectura y solución GAIA-X se focalizan en aquellos aspectos en los que la CAPV tiene más capacidades:

- Catálogo federado de servicios.
- Seguridad, interoperabilidad y gobernanza del dato.
- Monitorización continua de los nodos y servicios y certificación.
- Validación en dominios sectoriales identificados en RIS3 como energía e industria 4.0, y salud y otros alineados con la necesidad de modernizar la administración pública.

HODEI-X cuenta con un Comité de Contraste externo constituido por expertos seleccionados entre los agentes más relevantes cuyo objetivo es el contraste de los desarrollos planteados en el proyecto y la contribución a los diferentes grupos de trabajo.



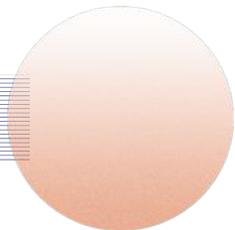
Ana Isabel Torre es doctora en Informática por la Universidad del País Vasco (EHU), Máster en Sistemas Informáticos Avanzados y Licenciada en Ingeniería Informática por la Universidad de Deusto.

Lidera el equipo de *High Performance Architectures*, focalizado en la conceptualización y desarrollo de Arquitecturas TI avanzadas y modelos de computación en la nube para el desarrollo de sistemas Big Data, desarrollando proyectos en clientes del sector energético, industrial y logístico mediante el uso de tecnologías Big Data y técnicas de *Machine Learning*. En investigación sus intereses se centran en la innovación mediante la aplicación de nuevas tecnologías semánticas (almacenamiento y consulta, integración semántica o mapeo de ontologías), de Big Data, analítica de datos y computación paralela, centrándose en el almacenamiento, transformación (ETL) y análisis/consulta en *batch* y tiempo real para diferentes sectores



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos
TECNALIA relacionados



INTEROPERABILIDAD Y GOBIERNO DE DATOS (Cont.)

- ▶ **Control de uso de datos compartidos.** Los mecanismos de control de acceso tradicionales únicamente proporcionan control en el lado del proveedor ante la solicitud de acceso a un recurso (p. ej. para lectura o escritura de un recurso de datos). Una vez los datos son consumidos, el control sobre estos se pierde. El control de uso en plataformas de almacenamiento y compartición distribuida extiende esta funcionalidad controlando cómo los datos están siendo utilizados por las partes consumidoras. Además, el control de uso de los datos no se limita únicamente al cumplimiento de un conjunto de condiciones para poder realizar su uso, sino que también, en ocasiones, se extiende al cumplimiento de una serie de obligaciones mientras y una vez los datos hayan sido utilizados. Estas obligaciones pueden consistir en un conjunto de acciones como, por ejemplo, eliminar los datos, notificar al proveedor del modo de uso de los datos, etc.
- ▶ **Tratamiento y compartición de datos personales.** Además de una infraestructura técnica de compartición de datos confiable y segura, otro factor que está obstaculizando el crecimiento de la economía de los datos es la falta de plataformas de datos personales confiables, seguras y basadas en la ética y métodos de análisis conscientes de la privacidad capaces de, por un lado, asegurar el intercambio de datos personales y de propiedad / comercial / industrial. datos y, por otro lado, de definir de manera estricta y justa cómo se puede capturar, producir, liberar y cobrar valor en beneficio de todas las partes interesadas involucradas.

DataVaults

DATAVAULTS

El proyecto DataVaults, financiado con fondos europeos del programa H2020 (ref. 871755), responde a la necesidad de disponer de plataformas de datos personales confiables, seguras y éticas, así como métodos analíticos que respeten la privacidad. La plataforma DataVaults gestiona los datos personales procedentes de diversas fuentes y define mecanismos seguros, fiables y de preservación de la privacidad, lo que permite a las personas asumir la propiedad y el control de sus datos, pudiéndolos compartir a voluntad mediante un intercambio flexible de datos y planes de compensación justos con otras entidades.

TECNALIA ha desarrollado los activos *"Secure Data Sharing and Access"* y *"Privacy and Trust Preservation"*. Sirven para implementar el concepto de Control de Acceso a los Datos mediante un editor de políticas de acceso y un motor de gestión de estas políticas. Con ellos, la parte propietaria de los datos personales puede definir las políticas de uso de sus datos personales usando el editor y cuando una empresa solicita acceso a dichos datos el "motor" se asegura de que cumple dichas políticas.

El objetivo de estas herramientas es facilitar el cumplimiento del RGPD - Reglamento Europeo de Protección de Datos, asegurando que los organismos que utilizan datos personales tengan la base legal necesaria y dotando al ciudadano de los mecanismos para ejercer sus derechos.



Josu Díaz de Arcaya **está actualmente desarrollando su tesis doctoral sobre el despliegue automático de modelos matemáticos, siendo codirigida por TECNALIA y la Universidad de Deusto.**

Es Ingeniero de Software (MSc) por la Universidad de Mondragón. Sus principales intereses son Big Data, DevOps y MLOps. Durante los últimos diez años, ha trabajado tanto en empresas de investigación como industriales. Al inicio de su carrera desempeñó su actividad como ingeniero de software haciendo streaming de datos en el Gran Colisionador de Hadrones del CERN. Después de eso, ocupó un puesto de desarrollador de software y administrador de sistemas en ICAA. En 2014, comenzó a trabajar para IBM UK, desarrollando sistemas distribuidos en tecnologías de almacenamiento de datos. En 2017, se unió a IKERLAN como ingeniero de datos diseñando e implementando ecosistemas intensivos en datos. Desde 2018 trabaja en TECNALIA como *Data Engineer* en el equipo de *High Performance Architectures*, donde se especializa en Tecnologías Big Data, microservicios y MLOps.



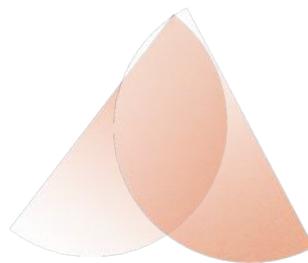
03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados

MODELADO Y OPTIMIZACIÓN BASADA EN DATOS

Trabajamos, en colaboración con instituciones nacionales e internacionales punteras, en el diseño, implementación y validación de nuevos métodos de aprendizaje automático (*Machine Learning*), aprendizaje por refuerzo (*Reinforcement Learning*) y meta-heurística evolutiva para abordar problemas prácticos de modelado de datos y optimización. En estas áreas se abordan retos específicos de vanguardia en la comunidad investigadora internacional:

- ▶ **Aprendizaje automático tradicional (*Machine Learning*)**, donde destacan las aportaciones sobre los modelos por ensembles, aplicaciones con redes neuronales basadas en randomización (*Reservoir Computing*), modelos que usan la computación bio-inspirada para mejorar su capacidad de generalización, o el uso de autómatas celulares para tareas de modelización sobre flujos de datos.



- ▶ **Aprendizaje profundo (*Deep Learning*)** para aprendizaje sobre datos de imagen, vídeo y series temporales, donde destaca la investigación llevada a cabo sobre métodos capaces de resolver simultáneamente varias tareas de modelado (*multitask learning*), los modelos generativos basados en *Deep Learning* (*Generative Adversarial Networks*) o el uso de algoritmos evolutivos para la optimización estructural de modelos de computación neuronal (*Neural Architecture Search*).
- ▶ **Metaheurística evolutiva e inteligencia de enjambre**, siendo especialmente reconocidos a nivel mundial nuestros trabajos metodológicos sobre el campo de la optimización bio-inspirada y la optimización multitarea.
- ▶ **Aprendizaje por refuerzo (*Reinforcement Learning*)**. Estudiamos en profundidad y aplicamos el aprendizaje colaborativo entre agentes heterogéneos (con diferentes capacidades), así como la no-estacionariedad del entorno sobre el que se formula la tarea de aprendizaje.



La "Cátedra TECNALIA en Inteligencia Artificial" tiene el objetivo de impulsar la aplicación multi sectorial de la Inteligencia Artificial desde el conocimiento y la excelencia científico tecnológica. La Cátedra, adscrita al Instituto Andaluz de Investigación en *Data Science and Computational Intelligence*, (DaSCI), está dirigida por el Prof. Francisco Herrera.



El equipo JRL está compuesto por investigadores de mentalidad abierta y totalmente colaborativos en ámbitos relacionados con el transporte, la ciencia de datos, la industria, la energía, la arquitectura, la agricultura o la ingeniería especial. Los temas de investigación del JRL son:

- Time series analysis
- Stream learning
- Optimization
- Lifelong ML
- Adversarial ML.



Javier Del Ser es doctor en Ingeniería de Telecomunicación (Cum Laude) de la Universidad de Navarra, 2006, y tiene un segundo Doctorado en Inteligencia Computacional (Summa Cum Laude, Premio Extraordinario) de la Universidad de Alcalá en 2013.

Licenciado en Ingeniería Eléctrica por la Facultad de Ingeniería de la Universidad del País Vasco. En TECNALIA lidera la investigación en Análisis y Optimización de Datos, colabora con el Centro Vasco de Matemática Aplicada (BCAM) y es profesor a tiempo parcial en la Universidad del País Vasco. Es asesor senior de Inteligencia Artificial en la startup Sherpa.ai. Su actividad se centra en la Inteligencia Artificial, la teoría de la información, análisis de datos, aprendizaje máquina, heurística y, en general, analítica descriptiva, prescriptiva y predictiva para la resolución de problemas relacionados con Energía, Industria 4.0, Telecomunicaciones, Economía, Salud y Movilidad Inteligente, entre muchos otros. En estos campos ha publicado hasta la fecha más de 360 publicaciones científico-técnicas (h index de 32), editado 6 libros, dirigido 30 tesis de máster y 10 tesis doctorales, así como participado/dirigido más de 50 proyectos de investigación. Ha sido el receptor, por dos ocasiones, del programa Torres Quevedo del Ministerio de Ciencia e Innovación (2007 & 2009). Es miembro Senior del IEEE, está acreditado en la figura de Catedrático de Universidad por la ANECA, y ha sido galardonado con el premio "Talento de Bizkaia" como reconocimiento a su experiencia y currículo científico.



03 Aportación tecnológica y proyectos

Activos tecnológicos y proyectos TECNALIA relacionados



SEGURIDAD Y CONFIABILIDAD DEL DATO, MODELO Y RESULTADO DEL PROCESO IA

El procesado analítico y los modelos de IA deben trabajar con datos y modelos de calidad. Trabajamos en:

- **Confiabilidad del dato.** El contexto donde los datos se recolectan en las aplicaciones habituales está sujeto frecuentemente a varias fuentes de incertidumbre, procedentes de factores exógenos no necesariamente reflejados en los propios datos en sí. Esta incertidumbre provoca que la salida de los modelos de IA aprendidos de los datos no sean confiables, lo que hace aún más complicado garantizar la confianza del usuario respecto a la certeza con la que el modelo dictamina su salida.

En TECNALIA estamos desarrollando nuevas técnicas para la estimación de la incertidumbre aleatoria (debida al dato) y epistémica (debida al modelo) en entornos con alto grado de incertidumbre, utilizando para ello varios enfoques agnósticos al modelo en sí (p.ej., predicción conformal o por cuantiles) y específicos para técnicas de aprendizaje automático tradicional (p.ej., quantile loss para ensembles) y computación neuronal profunda (Deep Ensembles, Evidential Deep Learning, Monte Carlo Test Dropout).

- **Explicabilidad.** En determinados casos la elevada complejidad necesaria para el modelo (como suele ser el caso con datos de imagen, vídeo o series temporales) tiene como contrapartida la incapacidad de entender el funcionamiento interno del modelo por parte del usuario. Se debe concebir por tanto la toma de decisiones a partir del dato como un proceso en el que el usuario que consume la salida de un modelo cuenta con información aumentada sobre el conocimiento atesorado por el modelo, que le hace generar una salida para una entrada dada.

Trabajamos en técnicas de explicabilidad para modelos de IA que proporcionan un grado adicional de confianza por parte del usuario, generando información útil, legible y comprensible ("relatos") sobre el procedimiento por el que un modelo, a priori opaco, genera su salida ante una entrada dada. Esta información es diversa en su naturaleza, destacando los hallazgos en la creación de contrafácticos multicriterio o la generación de explicaciones locales basadas en la atribución de importancia para modelos de computación profunda no convencionales.

- **Obsolescencia del dato/modelo.** Otro de los problemas habituales es la obsolescencia de los datos y/o de los modelos de IA, especialmente en contextos de aplicación donde los datos son generados a alta velocidad por procesos y fuentes con cierta variabilidad temporal (no estacionariedad). Detectar cuándo el dato o el modelo deja de ser prevalente es esencial para que el usuario pueda confiar plenamente en su salida y, así, tomar una decisión asistida por conocimiento extraído del dato.

Nuestros equipos están desarrollando nuevas formas de detectar dichos cambios, tanto en entornos de aprendizaje sobre flujos de datos sujetos a deriva de concepto (concept drift) como en aprendizaje offline (batch learning), donde el problema de la variabilidad de los datos se manifiesta en forma de outliers o los denominados "ejemplos fuera de muestra" (out of distribution samples).



3KIA: THE BASQUE NETWORK OF EXCELLENCE IN AI

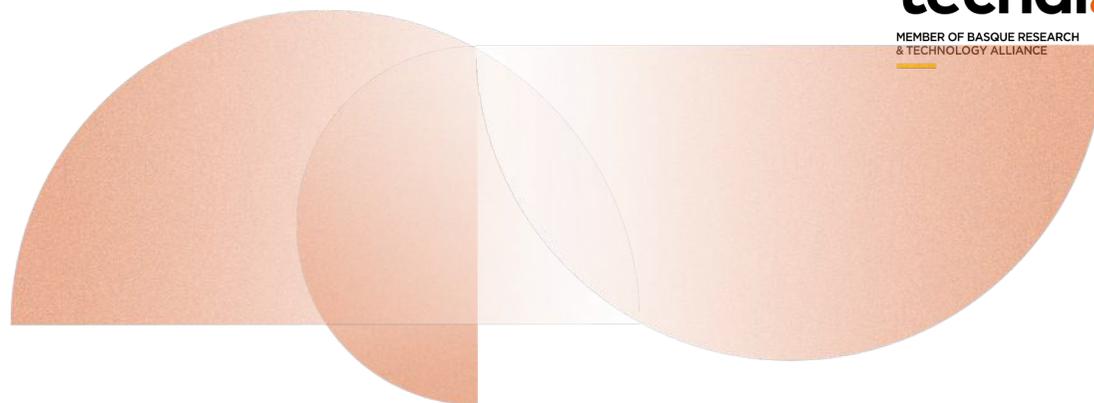
La red de excelencia está desarrollando técnicas y métodos para facilitar el diseño, desarrollo, validación y despliegue de sistemas prácticos basados en Inteligencia Artificial, con un enfoque multicriterio "3Q": calidad de los datos, calidad del modelo y calidad del resultado.

Contempla los criterios de:

- Privacidad – ¿Puede el modelo de IA utilizar cualquier fuente de datos?
- Equidad – ¿El modelo de IA se ve afectado por sesgos implícitos en la muestra de datos?
- Responsabilidad – Si falla, ¿se puede analizar por qué falla, depurar responsabilidades y evitar que vuelva a suceder?
- Robustez – Si el sistema sufre un ciberataque, ¿en qué medida podemos garantizar que sus resultados sean fiables?
- Confianza – ¿Se puede confiar en el resultado del modelo?
- Causalidad – ¿Se puede influir en la salida del modelo actuando sobre los datos de entrada?
- Explicabilidad – ¿Se comprende cómo funciona el modelo?
- Implementación y gobernanza de datos – ¿El modelo cumple con los requisitos del sistema? ¿Garantizan los modelos el uso efectivo, eficiente y lícito de la información?



04 Casos de uso



Páginas
6 – 9

01

Contexto y
motivación

Páginas
10 – 19

02

Iniciativas
singulares

Páginas
20 – 37

03

Aportación
tecnológica y
proyectos

Páginas
38 – 48

04

Casos de uso

Páginas
49 – 54

05

Equipo
TECNALIA

Páginas
55 – 56

06

Anexo

- Criterios de priorización
- Casos de Uso priorizados



04 Casos de uso

Criterios de priorización

La Estrategia Europea para Datos propone el desarrollo de espacios de datos en nueve ámbitos estratégicos y en ámbitos de interés público entre los que se citan los relacionados con la fabricación, con el Pacto Verde, con la movilidad, con la salud, con el ámbito financiero, con el de la energía, con el sector agrario, con el de las administraciones públicas y con las cualificaciones.

Lo cierto es que casi en cualquier ámbito se pueden identificar oportunidades concretas para el desarrollo de espacios de datos en los que la disponibilidad y el uso de datos faciliten la creación de servicios inteligentes con un impacto sistémico en todo el ecosistema y en el conjunto de la sociedad.

MAXIMIZAR EN IMPACTO

Desde la perspectiva TECNALIA, se puede maximizar el retorno de las inversiones en los espacios de datos priorizando los ámbitos de mayor impacto y en los que existen determinadas condiciones favorecedoras y menos barreras. El esquema representado debajo sintetiza los criterios de priorización¹.

Nota (1): este tipo de criterios, modificados o extendidos, se han utilizado también por el Grupo Impulsor del Hub ESP GAIA-X para su propuesta de Casos de Uso de Espacios de Datos.

En qué medida los agentes del espacio de datos están estructurados, cuentan con agentes tractores, agentes intermedios, tamaño, etc.	Articulación del ecosistema
En qué medida existen o resulta factible la disponibilidad de datos, de qué naturaleza y qué limitaciones.	Disponibilidad de datos
Existencia de iniciativas y experiencias previas en las que fundamentar el desarrollo de espacios de datos.	Existencia de iniciativas previas
En qué medida puede aprovechar el apalancamiento público – privado.	Colaboración Público – Privada
Contribución del desarrollo del espacio de datos a determinadas políticas y estrategias sectoriales u horizontales.	Alineamiento a políticas y estrategias
En qué medida el desarrollo en un dominio o ámbito concreto tiene externalidades positivas en otros ámbitos.	Generación de externalidades
Cuestiones que evidencien el dimensionamiento europeo y la alineación a otras iniciativas semejantes con las que puedan maximizar impacto.	Alineamiento y complementariedad internacional
Contribución del desarrollo del espacio de datos a la mejora de vida de las personas, al bienestar, a valores sociales.	Potencial de impacto social
Contribución al dinamismo económico, a la generación de empleo de calidad, a la cualificación y desarrollo profesional, a la innovación.	Potencial de impacto económico
Contribución a objetivos del Pacto Verde y medioambientales, incluidos el ecosistema natural, la flora y fauna.	Potencial de impacto medioambiental
Consideraciones relativas a los posibles retos o barreras de todo tipo que habría que superar para el éxito del Caso de Uso.	Retos y barreras para el éxito



04 Casos de uso

Casos de Uso priorizados

El análisis de priorización de TECNALIA considera determinados dominios en los que se dan condiciones más favorables para el desarrollo de I+D y el despliegue innovador de los espacios de datos, tanto desde la perspectiva de los factores habilitantes como del potencial de impacto. A su vez, dentro de estos amplios dominios pueden identificarse áreas concretas de oportunidad. Estos dominios, con algunos ejemplos de áreas de oportunidad, son:

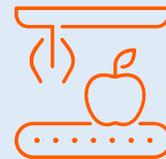
- ▶ Energía sostenible: Integración de energía renovable / Gestión activa de la demanda / Hidrógeno.
- ▶ Industria 4.0: Cadenas de suministro inteligentes y cooperativas / Gemelo digital y mantenimiento predictivo de plantas.
- ▶ Movilidad sostenible: Integración del vehículo eléctrico / Gestión de la multimodalidad.
- ▶ Comunidades inteligentes: *Smart Cities & Communities*.
- ▶ Turismo: Gemelo digital del territorio / Inteligencia turística.
- ▶ Comercio: Inteligencia en comercio.
- ▶ Salud personalizada: Salud personalizada / Medicina de precisión / Prevención de la cronicidad.
- ▶ Agro 4.0: optimización de la cadena de suministro desde la producción a la mesa.



Energía
sostenible



Turismo



Industria 4.0



Comercio



Movilidad
Sostenible



Salud
Personalizada



Comunidades
Inteligentes



Agro 4.0





04 Casos de uso

Casos de Uso priorizados



Energía sostenible

GEMELOS DIGITALES PARA LA OPTIMIZACIÓN Y EXTENSIÓN DE LA VIDA ÚTIL DE SISTEMAS DE GENERACIÓN / DISTRIBUCIÓN DE RENOVABLES

Área: energías renovables, eólica, fotovoltaica, termo solar; smart grid.

OBJETIVOS:

El Caso de Uso de Espacio de Datos en este ámbito está dirigido a posibilitar la generación de gemelos digitales en sistemas complejos multi-agente de generación o distribución de energía renovable (incluyendo almacenamiento y métodos de gestión activa de la demanda) e hidrógeno verde mediante la compartición voluntaria de datos en entornos interoperables para optimizar la operación, extender el ciclo de vida y contribuir a estabilizar los precios de la energía.

BENEFICIOS:

El aprovechamiento de los datos como herramienta de optimización de los sistemas de generación y distribución de energía eléctrica e hidrógeno verde tiene limitaciones por la falta de mecanismos interoperables de compartición, confianza y monetización.

IMPACTO:

Caso de uso con nítido impacto en los objetivos del Pacto Verde, invirtiendo en tecnologías limpias, contribuyendo a descarbonizar la energía e intensificar la generación renovable y su integración en el mix energético. La estabilización de los precios energéticos es también otro reto con impacto en la sociedad y en el contexto productivo.



PLATOON



Horizon 2020
European Union Funding
for Research & Innovation

PLATOON

PLATOON, con una financiación de 10 millones de euros del programa H2020 de la UE y ref. 872592, contribuye a la digitalización del sector energético, posibilitando niveles más altos de excelencia operativa. Despliega tecnologías de análisis de datos y procesamiento *edge* distribuido para una gestión optimizada del sistema de energía en tiempo real. Se utilizan conectores IDS (*International Data Space*) para garantizar la gobernanza de datos entre las diferentes partes interesadas en el intercambio de datos, así como para la cooperación entre múltiples actores en la cadena de valor. TECNALIA realiza la coordinación técnica de un proyecto que se está validando sobre 7 pilotos en 5 países que proporcionan casos reales de Big Data en el sector energético:

- Equilibrio eléctrico y mantenimiento predictivo.
- Estabilidad, conectividad y vida útil de la *smart grid*.
- Edificios: rendimiento operativo con modelos físicos y algoritmos IA.
- Sistema avanzado de gestión energética y modelos predictivos espaciales (multiescala) en la ciudad inteligente.
- Eficiencia energética y mantenimiento predictivo en el edificio terciario inteligente Hubgrade.
- Gestión energética de microrredes.



Eugenio Perea lidera en TECNALIA el equipo Digital Energy and Environment, orientado a la investigación, desarrollo e innovación de soluciones basadas en tecnologías digitales emergentes en los sectores de la Energía y el Medio Ambiente.

Es Doctor (Cum Laude) en Ingeniería de Telecomunicación por la Escuela de Ingeniería de Bilbao, UPV/EHU, y Máster en Gestión y Dirección de Empresas, MBA Executive de la UPV/EHU. Ha disfrutado de estancias de investigación en MATRA *Systemes et information* (actual EADS), INTESPACE y ARSENAL *Research* (actual *Austrian Institute of Technology*). Ha trabajado durante más de 20 años en decenas de proyectos de I+D+i relacionados con la digitalización y el sector de la energía, redes de distribución eléctrica, eficiencia energética en el uso final de la energía y Operación y Mantenimiento de infraestructuras de generación renovable. Es coautor de más de una decena de artículos de impacto y co-inventor de 5 patentes. Coautor del *Position Paper "Design Principles for Data Spaces, OPEN DEI Task Force"*, con contribuciones específicas para el sector energético. Es Profesor Asociado en la Escuela de Ingeniería de Bilbao, UPV/EHU.



04 Casos de uso

Casos de Uso priorizados



Industria 4.0

CADENAS DE SUMINISTRO INTELIGENTES Y COLABORATIVAS

Área: área industrial vinculada a cadenas de valor sectoriales: automoción, ferroviario, naval, aeroespacial, química, construcción, etc.

OBJETIVOS:

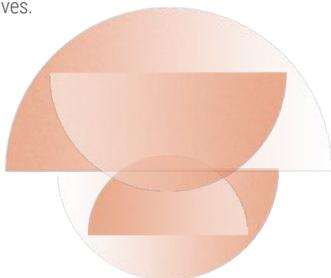
Contribución a la estandarización de plataformas de manufactura inteligente y datos IIoT – Industrial Internet of Things (p.ej., *Open Manufacturing*), superando la fragmentación actual en cada sector individual.

BENEFICIOS:

Cadenas de valor más eficientes para el consumidor; mejor diversidad de proveedores por facilitar la propuesta de valor de explotación de datos en el mercado para PYMEs. Mayor participación de todos los intervinientes en la cadena de suministro. Mayor personalización de los productos, hasta la "personalización en masa".

IMPACTO:

Facilitar la integración vertical de las cadenas de valor y posibilitar la incorporación de las PYMEs en las cadenas de valor internacionales de sectores claves.



5G EUSKADI

El proyecto 5G Euskadi es el piloto de 5G promovido por Red.es en la Comunidad Autónoma del País Vasco. El proyecto explora de la mano de empresas emblemáticas del País Vasco casos de uso avanzados que precisan de los recursos Edge (MEC) de la red 5G para alcanzar los requisitos de ancho de banda y latencia de aplicaciones que tienen que ver con el control de procesos industriales en tiempo real basados en la analítica de datos, o la seguridad de los operarios mediante el apoyo con tecnologías de realidad mixta. TECNALIA participa en el proyecto junto con el Basque Cybersecurity Center garantizando que tanto el despliegue de la red experimental como las arquitecturas de los casos de uso que hacen uso de los recursos Edge disponibles en la red (*Multi Access Edge Computing* - MEC) son ciberseguros.



Dr. Diego Galar lidera en TECNALIA el equipo de Mantenimiento Inteligente, actividad que compagina como profesor de Monitorización de Condición de la división de ingeniería de operación y mantenimiento en la Universidad Tecnológica de Luleå (LTU) en Suecia.

Así mismo, ha participado activamente en proyectos nacionales en la industria sueca y también financiados por agencias estatales suecas, como Vinnova. Además, es profesor de fiabilidad y mantenimiento de la Universidad de Skovde donde ostenta la cátedra VOLVO para Mantenimiento. En el ámbito internacional ha sido profesor visitante en la Politécnica de Braganza (Portugal), la Universidad de Valencia y NIU (EE.UU.), en la Universidad de Sunderland (Reino Unido) y la Universidad de Maryland (EE.UU.). En el mundo de la industria, ha sido director tecnológico y responsable de mantenimiento basado en condición (CBM, por sus siglas en inglés) de empresas internacionales y ha participado activamente en comités nacionales e internacionales para la estandarización e I+D en materias de fiabilidad y mantenimiento.



04 Casos de uso

Casos de Uso priorizados



Movilidad Sostenible

ORQUESTACIÓN DE LOS AGENTES PARA LA OPTIMIZACIÓN DE LA MOVILIDAD

Área: movilidad sostenible, e-movilidad, logística de la última milla.

OBJETIVOS:

Proporcionar mecanismos para la orquestación de las actividades de los agentes en el territorio de manera que se puedan optimizar los procesos asociados a la movilidad de personas o mercancías, y los servicios asociados (carga y descarga de mercancías, carga de vehículos, etc.).

BENEFICIOS:

El caso de uso se orienta a posibilitar la orquestación de las actividades de los agentes de movilidad mediante la integración de datos relacionados en un contexto cross-sectorial (movilidad urbana, particular, servicios públicos, logística, electrolinerías...), de una manera segura, interoperable, confiable y, eventualmente, monetizable.

IMPACTO:

Nítido impacto en los objetivos del Pacto Verde para el uso eficiente de los recursos y la mejora de los sistemas de transporte, reduciendo el consumo energético y las emisiones. La mejora de la movilidad tiene igualmente impacto social y económico, contribuyendo a la competitividad del sector productivo, con servicios logísticos más eficientes, por ejemplo.



SMART MOBILITY LAB by TECNALIA

El *Smart Mobility Lab* (SML) es un laboratorio para la validación de nuevos conceptos y soluciones en el ámbito de la movilidad inteligente basadas en el dato. El SML dispone de infraestructuras TI y diversas herramientas de SW que permiten a nuestros clientes realizar pruebas de concepto y avanzar en la implementación de herramientas Big Data. Contribuimos en la conceptualización de estas soluciones de movilidad inteligente. El equipo del SML también desarrolla algoritmos ad-hoc que puede ser necesaria para llevar a la práctica los conceptos desarrollados. Cabe destacar dos de las soluciones desarrolladas: (a) sistema para la predicción del flujo de tráfico urbano en el que se combinan múltiples fuentes de datos que son procesadas mediante tecnologías tanto de *deep learning* como de *big data*, y (b) para la optimización del reparto de última milla en la que se pueden optimizar cientos de repartos con restricciones espacio temporales y métricas de optimización personalizada. En la actualidad, este equipo trabaja en desarrollo de herramientas para dar soporte a la decisión en la implantación de políticas urbanas y en el proyecto INCOR en el que se recoge, almacena y procesa información V2X obtenidas con tecnologías ITS-G5 y 5G. Estos datos en conjunción con otros datos abiertos, son la base para que terceros puedan desarrollar servicios orientados a una gestión más eficiente de la movilidad.



Isidoro Cirió Director del Área de Negocio OPTIMA (Optimización, Modelización y Análisis de Datos) de Tecnalia, secretario del Clúster de Movilidad y Logística de Euskadi y representante de Tecnalia en AI Basque: el Grupo de Trabajo de Inteligencia Artificial del Clúster Gaia. Ingeniero de Telecomunicación.

Como director de proyectos primero y del área de negocio OPTIMA después, durante casi veinte años de vida profesional ha venido dinamizando iniciativas en los ámbitos de la movilidad, la logística y las infraestructuras de transporte, como el Smart Mobility Lab de Tecnalia (<https://sml.tecnalia.com/>), siempre desde la propuesta de valor de la Inteligencia Artificial y de las tecnologías Big Data. Desde esta aproximación ha impulsado con decenas de clientes de Tecnalia en el desarrollo de planificadores multimodales para mejorar la gestión del transporte público, de algoritmos para la predicción y caracterización de la demanda de movilidad a partir del análisis de datos de telefonía móvil, tanto individual (patrones personales), como agregada (matrices Origen / Destino), de algoritmos de predicción del tráfico y de la movilidad peatonal, etc. En cuanto al futuro, Tecnalia vislumbra un cambio total de paradigma con el advenimiento del vehículo autónomo, un vehículo que será eléctrico, conectado y, cuestión que suele pasar inadvertida, compartido.



04 Casos de uso

Casos de Uso priorizados



Comunidades Inteligentes

EL VALOR DEL DATO EN LOS EDIFICIOS, LA CIUDAD Y LAS INFRAESTRUCTURAS

Los edificios, las infraestructuras técnicas de servicios públicos y, en última instancia, las ciudades, emergen como agentes de innovación en la creación de valor basada en los datos, en los que se ha venido en denominar los *Gemelos Digitales para la Gestión Inteligentes de los Edificios o Infraestructuras y, también, Smart Cities*. En concreto, la Comisión Europea las considera como *"el lugar en el que las redes y servicios tradicionales se hacen más eficientes mediante el empleo de las TICs, para el beneficio de sus habitantes y negocios"*.

Como aspecto a considerar en la explotación del dato, en el entorno construido tienen especial relevancia la puesta en valor del dato y su contextualización mediante los *modelos inteligentes de la construcción (BIM) y los sistemas de información georreferenciada (GIS)*.

Por lo tanto, uno de los principales instrumentos para la materialización y concreción de la generación y extracción de valor del dato en los Gemelos Digitales está en la generación de un ecosistema híbrido, físico y digital.

En el contexto de edificación y entorno urbano, un gemelo digital es una réplica virtual de los principales elementos del edificio, la ciudad y/o sus infraestructuras críticas, conectados a bases de datos y sensores. Este modelo de la ciudad y sus procesos permite analizar, modelizar, simular y predecir escenarios o elaborar hipótesis para soportar mejores procesos de toma de decisión en el mantenimiento y operación del edificio, el planeamiento urbano y en la gestión de las infraestructuras y servicios públicos, en múltiples dominios y de una forma integrada.



KUBIK 4.0. RECIBE EL PREMIO DE DIGITALIZACIÓN EN LOS ADVANCED ARCHITECTURE AWARDS 2021

Los *Advanced Architecture Awards* reconocen el trabajo, liderazgo y esfuerzo de aquellas empresas y profesionales que apuestan por la transformación, disrupción e innovación.

Además reconocen a empresas y profesionales que generan nuevos materiales, que desarrollan productos alternativos, equipos más eficientes, sistemas de edificación acordes con criterios de sostenibilidad y procesos operativos que deben ayudar a una industrialización de su producción. También a aquellas que aplican la tecnología y la digitalización para mejorar los procesos en la edificación.

En el marco de estos Premios, KUBIK 4.0: GEMELO DIGITAL DE UN EDIFICIO, fue galardonado el pasado miércoles 22 con el primer premio a la Digitalización o Apuesta Tecnológica. **KUBIK** es una apuesta que nació hace 10 años, como infraestructura experimental singular para contribuir a la mejora de la eficiencia energética en la edificación y que en los últimos años *se ha reinventado iniciando una nueva etapa digital que ha sido liderada por Rosa San Mateos y Jose Manuel Olaizola*.



José Antonio Chica **Ingeniero Industrial Mecánico por la Escuela de Ingeniería de Bilbao, UPV-EHU, doctorado en Ingeniería Mecánica y Estructural por la Universidad de Burgos y MBA por ESEUNE y la Universidad Georgetown**. Además, orientado a los retos de la Transformación Digital, cuenta con un **Executive Certificate on Strategy and Innovation** en el **Massachusetts Institute of Technology - Sloan School of Management**.

Director del Área "Construcción Digital" en TECNALIA, miembro del Grupo de Seguimiento del "Fondo para la Investigación del Carbón y el Acero" de la Comisión Europea y del Grupo Temático 1 "Promoción de la Inversión en Renovación de la Edificación, Infraestructuras e Innovación" de la Dirección General para el Mercado Interior, Industria, Emprendizaje y PYMES de la Comisión Europea.

Miembro de las Plataformas Tecnológicas del Acero en Construcción, de la Construcción y grupos de trabajo del Eurocódigo y, actualmente, miembro de los comités de Digitalización del Entorno Construido de las Plataformas Tecnológicas española y europea de la construcción.



04 Casos de uso

Casos de Uso priorizados



Turismo

HABILITACIÓN DE LA INTELIGENCIA TURÍSTICA BASADA EN DATOS

Área: Territorios inteligentes, turismo inteligente.

OBJETIVOS:

Esta iniciativa se alinea con el componente 14 del Plan de Recuperación y Resiliencia, España Puede, contribuyendo a un entorno de compartición de datos relacionados con la operativa de turismo.

BENEFICIOS:

Oportunidad de afianzar un sector clave de la economía española como el turismo frente a la competencia de países emergentes con costes más bajos. Oportunidad de explotar otras vertientes de la marca turística del territorio apostando por la sostenibilidad, la diversidad y el turismo de calidad.

IMPACTO:

Garantizar la sostenibilidad social, económica y medioambiental, y fortalecer la convivencia entre residentes y visitantes en el territorio.



HODEIAN

H0de1an es una herramienta de inteligencia y gobernanza turística para el análisis de la actividad de turistas y visitantes en el territorio de Gipuzkoa. Sirve de ayuda a la toma de decisiones por parte de los gestores de la actividad. Se han desarrollado sistemas que aprenden automáticamente a partir de los datos para abordar predicciones inteligentes sobre la ocupación, la nacionalidad de los turistas que van a llegar o identificar las rutas habituales que estos realizan en un territorio. Se analiza el gasto generado a través de la información anonimizada de las transacciones realizadas con tarjetas y operaciones de TPVs de una entidad bancaria, y la movilidad y trazabilidad procedentes de una operadora de telefonía móvil. También se analizan de forma respetuosa con la privacidad las huellas digitales que dejan los visitantes con un sistema de conteo de personas y tracking en Puntos de Interés Turístico.



Jesus Herrero Arranz es responsable de Desarrollo de Negocio en I+D para el sector Turismo. Es doctor en Informática por la Universidad de Deusto, *Master Foundations of Advanced Information Technology* por el Imperial College de la Universidad de Londres y Licenciado en Informática por la Universidad de Deusto.

Ha dirigido numerosos proyectos de investigación en los ámbitos nacionales e internacionales relacionados con la innovación y la aplicación de las tecnologías habilitadoras en el sector turístico, incluido el proyecto HODEIAN por el que recibió el Premio AMETIC en la categoría Economía del Dato Turístico en 2020. Forma parte del comité AENOR que aborda la IoT y tecnologías relacionadas (CTN 071/SC41) y la normalización de las Smart Cities y los Destinos Turísticos Inteligentes (CTN 178/SC5). Se ha especializado en los Sistemas de Información y las herramientas de *Business Intelligence* de analítica y visualización de datos aplicadas a diferentes sectores.



04 Casos de uso

Casos de Uso priorizados



Comercio

DECISIONES INTELIGENTES BASADAS EN DATOS PARA EL COMERCIO

Área: Territorios inteligentes, comercio, logística de la última milla.

OBJETIVOS:

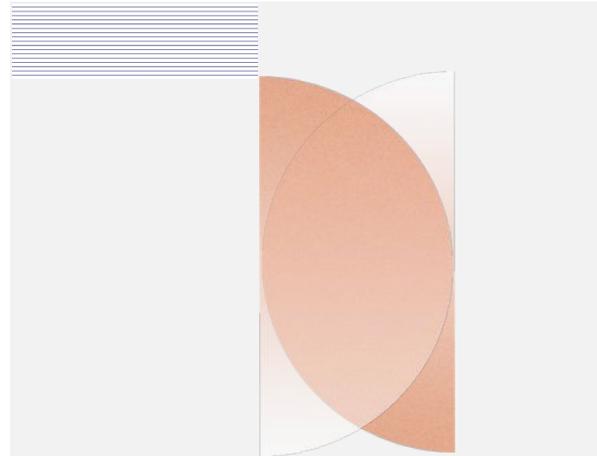
Las plataformas de inteligencia de cliente posibilitan la atención personalizada, la optimización de la cadena de suministro (predicción de la demanda, centrales de compra automatizada, reducción de stocks, la identificación de patrones de compra, la gestión de última milla); facilitan los escenarios de negocios colaborativos (co-creación, co-diseño de productos) y fomentan la colaboración entre retailers y marcas.

BENEFICIOS:

Aumento del nivel de fidelización del cliente local; mejora de la experiencia de compra y venta; mejora de los tiempos de entrega; gestión más eficiente de la trazabilidad.

IMPACTO:

Garantizar la sostenibilidad social, económica y medioambiental del comercio local, y fortalecer su competitividad respecto a las plataformas internacionales.



APORTACIÓN DE TECNALIA AL CASO

El sector del comercio tiene que abordar el gran desafío de su transformación estructural y la recuperación tras la crisis provocada por la covid-19, el crecimiento del comercio electrónico o el cambio en los hábitos de consumo. La inteligencia basada en datos contribuye a la modernización y mejora competitiva de este sector con impacto en múltiples procesos de negocio. TECNALIA aporta su experiencia en aplicaciones concretas como la gestión de agendas de los y las profesionales adaptadas a la predicción de actividad comercial, el ajuste del consumo de energía, la optimización de activos, como flotas y almacenes, la monitorización de residuos y emisiones, la identificación del momento de compra más ventajoso, la creación de paneles de ayuda a la visualización del negocio y a la toma de decisiones, el perfilado y segmentación de clientes, la predicción de la demanda a través de los modelos de ocupación y presencia peatonal en las vías, etc.



Sergio Campos es licenciado en informática (Universidad Politécnica de Madrid, 1999), Master en Ingeniería de Sistemas de la Decisión (Universidad Rey Juan Carlos de Madrid, 2014).

En TECNALIA es jefe de proyecto y responsable del foco *Smart Mobility*. Su actividad investigadora está centrada en ingeniería de sistemas críticos, plataformas para la gestión eficiente de movilidad (*Smart Mobility*), técnicas analíticas (fusión de datos, análisis de patrones y predicción) y su implementación óptima en arquitecturas avanzadas; con aplicación en problemas del ámbito urbano, movilidad/transporte inteligente, logística y fiabilidad en sistemas complejos. Previamente a su incorporación a TECNALIA, ha trabajado en THALES, TELVENT e IBERINCO. Como consultor ha realizado diferentes servicios de evaluación y soporte al despliegue de modelos AutomotiveSPICE y CMMI (+SAFE). Estudiante de doctorado en el ámbito de análisis de patrones de movilidad urbana.



04 Casos de uso

Casos de Uso priorizados



Salud Personalizada

SALUD PERSONALIZADA Y MEDICINA DE PRECISIÓN

Área: salud, médica, farmacia.

OBJETIVOS:

La medicina de precisión (también conocida como medicina 4P: personalizada, predictiva, preventiva y participativa), en la que la atención de la salud se adapta individualmente en función de los genes, el estilo de vida y el entorno de una persona, ha despertado un gran interés en el tratamiento individualizado de la práctica clínica. Los tratamientos médicos y preventivos deben considerar la variabilidad de los individuos para establecer el tratamiento que mejor se adapta a cada paciente. Hoy en día, la disponibilidad de datos clínicos es una realidad, pero debido a su dispersión, privacidad, volumen y complejidad, su gestión es compleja. En este Caso de Uso, los Espacios de Datos y las técnicas avanzadas de análisis de datos permiten procesar, analizar e interpretar una gran cantidad de datos heterogéneos de diferentes fuentes en períodos de tiempo relativamente cortos que sirvan para la toma de decisiones inteligente y personalizada de tratamientos mediante la integración de información clínica y patológica, de cronicidad y de las prescripciones farmacológicas.

BENEFICIOS:

Soporte a las decisiones de los profesionales que sirva para una mejor atención según las características de cada paciente.



IMPACTO:

La creciente disponibilidad de datos de salud y los avances en genética presentan una oportunidad para hacer de la atención al paciente personalizada y precisa una realidad clínica que sirva para mejorar la efectividad de los tratamientos y su coste, así como minimizar el riesgo para la salud o los efectos no deseados en los pacientes.



LABORATORIO DE GENÉTICA

El laboratorio de genética de TECNALIA está autorizado y registrado por el Gobierno Vasco para el diagnóstico genético. Trabaja en el diagnóstico de síndromes hereditarios (tumoraes y no tumoraes), en el diagnóstico de enfermedades neoplásicas y en el diagnóstico genético preimplantacional (DGP). Estos estudios proporcionan:

- Diagnóstico de enfermedades.
- Identificación de los factores genéticos asociados a la enfermedad.
- Pronóstico y predicción de la enfermedad.
- Selección del mejor medicamento y de la dosis correcta.
- Definición de los factores genéticos que aumentan el riesgo de enfermedad.
- Identificación de factores genéticos hereditarios.



Miguel Esteras Béjar es doctor en Investigación Clínica por el Imperial College London, Máster de especialización en Ciencia de los Datos en City, University of London.

Previamente a su incorporación a TECNALIA ha trabajado en NHS Digital (socio nacional de información y tecnología para el sistema de salud y atención del Reino Unido) en modelos de aprendizaje estadístico para la mejora de la eficiencia y calidad de los servicios, y de la seguridad de los pacientes. Su foco de especialización es el desarrollo de herramientas para la detección no-supervisada de anomalías en series temporales, desarrollo de algoritmos basados en aprendizaje estadístico para el análisis descriptivo y predictivo de las estrategias de operación de instalaciones, desarrollo de herramientas para la transferencia de conocimiento experto humano a los algoritmos de aprendizaje de estadístico, el despliegue de aplicaciones analíticas de servicios en la nube y el desarrollo de visualizaciones para la exploración e interpretación de datos y resultados.



04 Casos de uso

Casos de Uso priorizados



Agro 4.0

OPTIMIZACIÓN DE LAS CADENAS DE SUMINISTRO DESDE LA PRODUCCIÓN A LA MESA

Área: agroalimentario, pesca, acuicultura.

OBJETIVOS:

Proporcionar mecanismos para la orquestación de las actividades de los agentes de la cadena de valor alimentaria para aumentar la eficiencia operacional, la satisfacción del consumidor y la reducción de desperdicios y consumos (energía, riego, medicamentos, piensos, etc.).

BENEFICIOS:

Mejorar la eficiencia operacional de los procesos de explotación, distribución y comercialización de alimentos; reducción de consumos y desperdicios; optimización del ecosistema vivo.

IMPACTO:

Palanca impulsora para los planes nacionales en al lucha contra la despoblación, siendo el sector agrícola por ejemplo el de mayor implantación en las áreas rurales.



AFARCLOUD, *Aggregate Farming in the Cloud*

AFarCloud cuenta con una financiación de 8 millones de euros del programa H2020 de la UE y ref. 783221 para generar una plataforma distribuida para la agricultura autónoma que integra y soporta la cooperación de sistemas ciberfísicos agrícolas en tiempo real. El objetivo final es aumentar la eficiencia, la productividad, la salud animal, la calidad de los alimentos y reducir los costos de mano de obra agrícola. La detección confiable, la identificación precisa y la cuantificación adecuada de patógenos y otros factores que afectan tanto a la salud vegetal como animal son fundamentales para controlar los gastos económicos, las interrupciones del comercio e incluso los riesgos para la salud humana.

La plataforma se integra con software de gestión agrícola, apoyando el seguimiento y la toma de decisiones basadas en *Big Data* y técnicas de minería de datos en tiempo real en 3 demostradores (Finlandia, España e Italia). Estos demostradores incluyen escenarios de gestión de cultivos y ganado.

El consorcio representa toda la cadena de valor de las soluciones agrícolas basadas en las TIC, incluidos todos los actores clave necesarios para el desarrollo, la demostración y la futura asimilación por parte del mercado del marco de la agricultura de precisión que se apunta en el proyecto.



Sonia Bilbao es Licenciada en Ciencias Físicas (especialidad en Electrónica y Automática por la UPV/EHU) en 1998 e Ingeniera en Electrónica por la UPV/EHU en 2002.

Sonia lidera en TECNALIA proyectos e iniciativas ligadas al mundo de los datos en el sector primario. Cuenta con una experiencia de más de 15 años en Tecnologías de la Información, gestionando, planificando y desarrollando proyectos I+D+i tanto a nivel nacional como europeo. Sonia es una investigadora de prestigio en el ámbito de las tecnologías relacionadas con *Internet of Things*, con foco en Interoperabilidad, donde atesora una altísima especialización.



05 Equipo TECNALIA



Páginas
6 – 9

01

Contexto y
motivación

Páginas
10 – 19

02

Iniciativas
singulares

Páginas
20 – 37

03

Aportación
tecnológica y
proyectos

Páginas
38 – 48

04

Casos de uso

Páginas
49 – 54

05

Equipo
TECNALIA

Páginas
55 – 56

06

Anexo

- Quiénes somos
- Talento en Datos e Inteligencia Artificial
- Capilaridad en el tejido productivo
- Red de colaboradores y alianzas
- Aportación a los Espacios de Datos



05 Equipo TECNALIA

Quiénes somos



1.472

PERSONAS EN PLANTILLA



56%

HOMBRES



44%

MUJERES



31

NACIONALIDADES DIFERENTES



267

NÚMERO DE DOCTORES

SOMOS LA REFERENCIA EN SERVICIOS DE I+D+i

TECNALIA es un Centro Tecnológico de investigación y desarrollo multisectorial y multitecnología referente en Europa que pertenece a BRTA (*Basque Research and Technology Alliance*).

Nuestra misión nos une: transformamos la tecnología en PIB. Transformamos la tecnología en riqueza para obtener resultados visibles y beneficiosos para las empresas, la sociedad, nuestro entorno y, en definitiva, para las personas.

Y nuestra visión nos diferencia: *ser agentes de transformación de la industria.*

Apostamos por la colaboración estrecha con las empresas. Acompañamos a las empresas en un contexto internacional con nuestro potencial en I+D+i y servicios avanzados, siendo reconocidos por nuestra excelencia tecnológica y poniendo en valor nuestra capacidad de orquestación. Más de 7.800 empresas comparten su camino con nosotros en el ámbito de la I+D+i y los servicios avanzados.

Una referencia en Europa. TECNALIA es la 1ª organización privada estatal en contratación, participación y liderazgo de proyectos en el Programa Horizonte Europa 2020 de la UE. Ocupa el puesto 13 entre 2.900 organizaciones europeas de investigación (datos a mayo 2020).



TRANSFORMACIÓN DIGITAL



05 Equipo TECNALIA

Talento en Datos e Inteligencia Artificial

Somos personas y conocimiento, pasión y compromiso. Nuestros equipos y sus especialistas son un referente para las empresas.

En el área de conocimiento de Datos e Inteligencia Artificial disponemos de una combinación única de talento: especialistas apasionados con el desarrollo, experimentación y despliegue de sistemas de gestión de datos, ciberseguridad, modelos analíticos y sistemas de Inteligencia Artificial:

- ▶ 70+ investigadores/as y especialistas
- ▶ 22 grado doctor, 10 estudiantes doctorado.
- ▶ 200+ publicaciones
- ▶ 15+ familias de patentes y activos IPR
- ▶ 300+ proyectos industriales
- ▶ 30+ años de experiencia y referencias





05 Equipo TECNALIA

Capilaridad en el tejido productivo

Desarrollamos las tecnologías de gestión de datos e Inteligencia Artificial de vanguardia orientadas a crear productos y soluciones que hagan que los procesos de negocio sean más inteligentes y productivos en las organizaciones.

Trabajamos con arquitecturas avanzadas para el desarrollo de sistemas de datos: diseño, desarrollo y despliegue de plataformas de gestión de datos; plataformas de análisis de macrodatos (*Big Data*); migración, desarrollo e implantación de procesos operacionales y analíticos; soporte especializado para el despliegue, operación y mantenimiento por parte de terceros en infraestructuras tanto *on-premise* como *Cloud/Fog/Edge*.

Sobre los datos, trabajamos en la conceptualización, desarrollo e investigación de modelos analíticos y algoritmos de IA que den respuesta a los objetivos específicos de los Casos de Negocio según consideraciones específicas en términos de la naturaleza de los datos (series temporales, datos discretos/continuos, imágenes, etc.) y el rendimiento esperado (precisión, tiempo de cómputo, usabilidad, confianza).

- Smart Energy
- Smart Cities
- Industria 4.0
- Agro 4.0
- Smart Mobility & Logistics
- Salud personalizada
- Edificios e infraestructuras
- Ciberseguridad

Apreciamos sinceramente a nuestros valiosos clientes en este ámbito por la confianza que han depositado en nosotros y por hacernos mejorar cada día.





05 Equipo TECNALIA

Red de colaboradores y alianzas

TECNALIA pone en valor su capacidad de orquestación promoviendo la colaboración activa con diferentes entidades para impulsar el desarrollo y adopción temprana de innovaciones y tecnologías de gestión de datos y de IA.

Es una referencia en iniciativas de I+D cooperativas, tanto en el espacio nacional como en el internacional. Por ejemplo, en el Programa Marco H2020 de la UE, TECNALIA ha participado en más de 450 proyectos, liderando 68.

Una de las claves del éxito de TECNALIA en el Espacio Europeo de Investigación ha sido su posicionamiento estratégico en más de 60 iniciativas público-privadas y Asociaciones europeas relevantes, como Plataformas Tecnológicas, Partenariados Público Privados, Iniciativas Tecnológicas Conjuntas, Partenariados Europeos de Innovación, Asociaciones, comunidades de innovación del Instituto Europeo de Tecnología, etc. De manera similar, muchas de este tipo de iniciativas se trasladan al contexto nacional.

La colaboración estrecha con socios y clientes en estos foros sirve para visualizar, identificar y desarrollar soluciones tecnológicas integrales en cooperación con diversos agentes.

Algunos ejemplos de estos foros, plataformas y alianzas se indican en la figura.





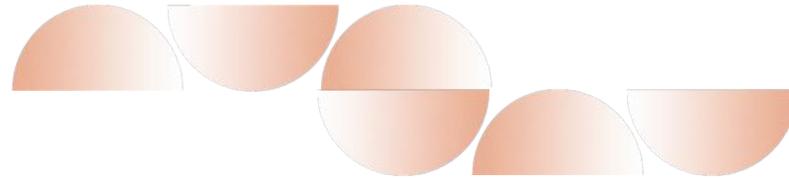
05 Equipo TECNALIA

Aportación a los Espacios de Datos

Creemos en la **Economía del Dato**, en su potencial de progreso, en su rol catalizador de innovación y en su capacidad para constituirse en motor de desarrollo económico. Este es uno de los grandes desafíos por los que TECNALIA apuesta en los próximos años. Nos apasiona idear y diseñar cómo serán los procesos, productos y servicios digitales en un mundo cada vez más interconectado, interactivo e inteligente.

Buscamos el objetivo final de impulsar la competitividad y diferenciación de las empresas a través de la incorporación, adopción y asimilación temprana e intensa de la Analítica Avanzada de Datos y la Inteligencia Artificial en sus procesos organizativos, productos y servicios.

Aportamos la capacidad y conocimiento de nuestros equipos, nuestra experiencia multisectorial y nuestra capacidad de liderazgo y orquestación de proyectos e iniciativas complejas basadas en la innovación y la tecnología.



COMBINACIÓN ÚNICA DE TALENTO

A través de nuestro equipo de especialistas y referencias, aportamos soluciones innovadoras y tecnológicas capaces de crear Espacios de Datos confiables, interoperables y soberanos que contribuyen a desarrollar todo el potencial de la Economía del Dato para la sostenibilidad social, económica y medioambiental.



EXTENSA CAPILARIDAD MULTISECTORIAL

TECNALIA ha colaborado con más de 7.800 empresas desde 2011, el 75% de ellas pymes, en ámbitos como la transición energética, la fabricación avanzada, la transformación digital, la movilidad sostenible, la salud o el ecosistema urbano.



CAPACIDAD DE LIDERAZGO Y ORQUESTACIÓN

TECNALIA aporta su extensa red de socios y aliados, así como su capacidad de liderazgo, para impulsar el desarrollo y adopción temprana de innovaciones y tecnologías de gestión de datos y de IA.



06 Anexo

Páginas
6 – 9

01

Contexto y
motivación

Páginas
10 – 19

02

Iniciativas
singulares

Páginas
20 – 37

03

Aportación
tecnológica y
proyectos

Páginas
38 – 48

04

Casos de uso

Páginas
49 – 54

05

Equipo
TECNALIA

Páginas
55 – 56

06

Anexo

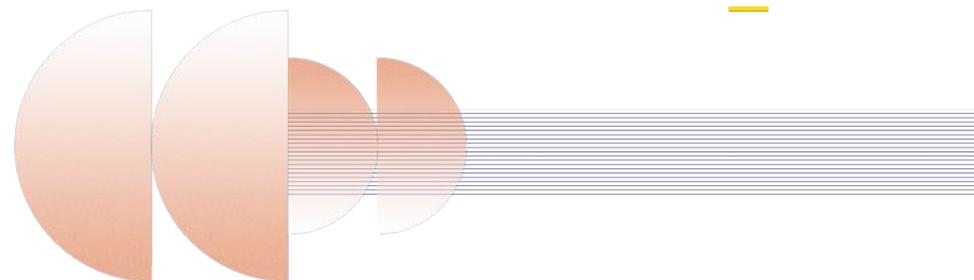
- Referencias



06 Anexo

Referencias

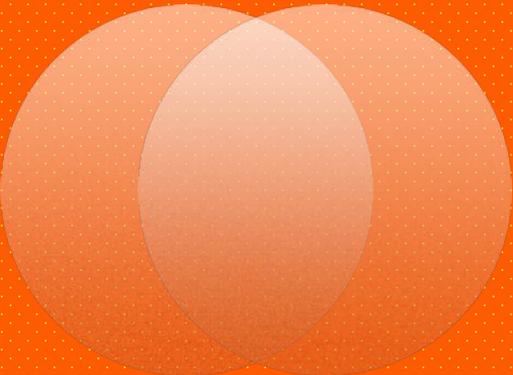
Para la elaboración de este documento se ha realizado una revisión de la documentación disponible, así como de referencias, casos y otros activos de conocimiento asociados al desarrollo los Espacios de Datos.



- ▶ Estrategia anual de crecimiento sostenible 2021, septiembre 2020; <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0575&from=EN>
- ▶ "Una Estrategia Europea de Datos", Comisión Europea, febrero 2020; <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0066&from=ES>
- ▶ SWD(2020) 205 final, Guidance to Member States Recovery And Resilience Plans, septiembre 2020; https://ec.europa.eu/info/sites/info/files/3_en_document_travail_service_part1_v3_en_0.pdf
- ▶ Digital Economy and Society Index; <https://ec.europa.eu/digital-single-market/en/digital-economy-and-society-index-desi>
- ▶ Estrategia Nacional de Inteligencia Artificial, diciembre 2020; https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2020/201202_np_eniav.pdf
- ▶ Plan España Digital 2025, julio 2020; https://www.lamoncloa.gob.es/presidente/actividades/Documents/2020/230720-Espa%C3%B1aDigital_2025.pdf
- ▶ Plan España Puede, Plan de Recuperación, Transformación y Resiliencia, octubre 2020; https://www.lamoncloa.gob.es/presidente/actividades/Documents/2020/07102020_PlanRecuperacion.pdf
- ▶ Página del Digital Europe Programme; <https://ec.europa.eu/digital-single-market/en/europe-investing-digital-digital-europe-programme>
- ▶ Towards a European strategy on business-to-government data sharing for the public interest; <https://www.euractiv.com/wp-content/uploads/sites/2/2020/02/B2GDataSharingExpertGroupReport-1.pdf>
- ▶ Towards a European-governed Data Sharing Space. Enabling data exchange and unlocking AI potential, BDVA, v2, septiembre 2020; https://www.bdva.eu/sites/default/files/BDVA%20DataSharingSpaces%20PositionPaper%20V2_2020_Final.pdf
- ▶ Data Governance Act site, noviembre 2020; <https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act>
- ▶ Data sharing in the EU – common European data spaces (new rules), 2020; <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12491-Legislative-framework-for-the-governance-of-common-European-data-spaces>
- ▶ Horizon Europe Cluster 4: Digital, Industry and Space; <https://ec.europa.eu/research/pdf/horizon-europe/annex-4.pdf>
- ▶ Posicionamiento sobre los Espacios de Datos, PLANETIC, enero 2021; https://planetec.es/sites/planetec.es/files/public/content-files/publications/2021/GL_BD_IA_Posicionamiento_Espacios_de_Datos_2021.pdf
- ▶ GAIA-X; <https://www.data-infrastructure.eu/GAIA/Navigation/EN/Home/home.html>
- ▶ IDSA; <https://internationaldataspaces.org/>
- ▶ Noticia INDESIA; <https://news.microsoft.com/es-es/2021/06/15/seis-grandes-empresas-crean-el-primer-consorcio-de-inteligencia-artificial-de-la-industria-en-espana/>
- ▶ European Alliance on Industrial Data, Edge and Cloud; <https://digital-strategy.ec.europa.eu/en/policies/cloud-alliance>



IMPULSO A LOS ESPACIOS DE DATOS



tecnal:a

MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE

blogs.tecnalia.com



www.tecnalia.com

