

Extracción de datos contenidos en documentos aplicando Inteligencia Artificial

2 La Inteligencia Artificial aplicada en la gestión de los expedientes de ayudas de rehabilitación de viviendas del Gobierno Vasco.



Jornadas TEKgunea 3.0

6 Las Jornadas TEKgunea han celebrado este año en el Museo Guggenheim Bilbao su tercera edición.

ALBOAN



AENOR
Confía
Certificado de Conformidad
con el Esquema Nacional de Seguridad



La plataforma de eAdministración del Gobierno Vasco consigue la certificación ENS

10 El pasado mes de abril la Plataforma PLATEA consiguió la certificación ENS.



Revista «KomunIKT» de Osakidetza

KomunIKT tiene como objetivo dar a conocer los Sistemas de Información de las áreas que forman parte del Servicio Vasco de Salud.

12

Xia Peisu, la madre de la informática china

Conoceremos la vida y carrera profesional de Xia Peisu, considerada la «madre» de la informática china.





Extracción de datos contenidos en documentos aplicando Inteligencia Artificial

Incorporar la Inteligencia Artificial a la gestión de los expedientes administrativos de ayudas, subvenciones... está demostrando ser una gran ayuda.



El proceso de tramitación de expedientes administrativos implica, normalmente, el tratamiento y estudio de un gran número de documentos para certificar el cumplimiento de los requisitos por parte de la persona o entidad interesada, para recopilar la información necesaria para realizar el cálculo del importe de una subvención, para justificar la ayuda solicitada, etc.

Así, obtener la información que se precisa a partir de los documentos que la contienen, estando en formato de **datos no estructurados**, pasa por un proceso manual, que incluye la lectura del documento, la localización del dato y su traslado a la aplicación que posteriormente la procesa. En muchas ocasiones estos documentos tienen información adicional que no es necesaria, haciendo, por tanto, que su lectura sea pesada y no aporte ningún valor añadido.

En el caso de los expedientes de ayudas a la rehabilitación de viviendas y edificios promovidos por la Dirección de Planificación y Procesos Operativos del Departamento de Planificación Territorial, Vivienda y Transportes, en poco más de mes y medio se recibieron 37.674 documentos. Como nos podemos imaginar, la carga de trabajo que supone su gestión es enorme. Por ello, desde la Dirección de Planificación y Procesos Operativos, siendo conscientes de esta realidad, se propuso buscar una solución que mitigase esta problemática y, para ello, conociendo las posibilidades que ofrecen las técnicas de computación



actuales como la Inteligencia Artificial (IA) se planteó este reto.

ALCANCE DEL PROYECTO

El nuevo sistema permite tratar los documentos recibidos para **extraer automáticamente los datos** de interés que alberga, evitando de este modo que ninguna persona tramitadora deba leerlo por completo y detenidamente para recopilar la información necesaria.

Para el tratamiento de los documentos se aplican técnicas de Procesamiento del Lenguaje Natural¹ (NLP) de Inteligencia Artificial:

- Para los documentos en formato imagen, mediante un proceso de OCR («*Optical Character Recognition*») se obtiene el texto procesable que contiene.
- Se aplican algoritmos de filtrado y optimización del texto obtenido («*Tokenization, Stemming/Lemmatization, Stop Words*», «*Sentence Boundary Disambiguation*», etc.)
- Se determina el lenguaje (euskera, castellano o bilingüe) en el que está redactado.
- Se identifica la tipología a la que pertenece.
- En función del tipo de documento del que se trate, se extraen los datos de interés que contiene («*Named-entity recognition*», NER)

¹**Procesamiento del lenguaje natural:** los avances en el campo de la Inteligencia Artificial y, en concreto, en el ámbito del procesamiento del lenguaje natural («*Natural Language Processing*», NLP) basado en la aplicación de algoritmos de aprendizaje automático o aprendizaje de máquinas («*Machine Learning*», ML) han permitido obtener, entre otros logros, la identificación del **idioma** de un texto, la clasificación del **tema** principal sobre el que trata, y la localización y obtención de ciertas entidades de **datos** contenidas en el mismo.

El proyecto incluye la definición, normalización y estandarización de todo el proceso de tratamiento (desde el aprendizaje del sistema hasta la exposición de los datos obtenidos para las aplicaciones usuarias) lo que permitirá ampliar fácilmente, ya no solo el catálogo de tipologías de documentos que el sistema es capaz de tratar, sino también su uso por cualquier otro sistema informático.

LA SOLUCIÓN

La aplicación se compone de varios **módulos** que trabajan de manera coordinada, pero sin olvidar la fase previa de entrenamiento, en la que se construye el modelo que se aplica posteriormente en la fase de inferencia (ejecución), en la que, del documento que se esté procesando, se extraen los datos requeridos.

Evidentemente, esa información que se ha extraído debe integrarse con la aplicación de gestión que los necesita, en este caso la de ayudas a la rehabilitación de viviendas y edificios, de modo que el sistema ofrece también una interfaz amigable («*plugin*») que facilita el trabajo de la persona tramitadora en el uso de esa información.

Módulo OCR

El primer gran reto para que el sistema pudiese funcionar era obtener el texto «en bruto» que contiene cada uno de los documentos que se recogen. Además,

esto es necesario tanto en la fase de inferencia final (extracción de datos) como en la fase previa de entrenamiento en la que se construye el modelo.

«Para el tratamiento de los documentos se aplican técnicas de Procesamiento del Lenguaje Natural de Inteligencia Artificial»

El módulo de OCR se encarga de tratar los documentos, los cuales pueden contener texto, imágenes o una mezcla de ambos, extrayendo el texto que contiene, y, transformado, por tanto, esas imágenes en texto procesable. Para ello, y como primer paso, se realiza un preprocesamiento, que, utilizando técnicas de tratamiento de imagen, «gira» el documento hasta su posición vertical (en caso de que no se haya digitalizado con dicha orientación), y se «binariza»² para resaltar el texto en colores de grises. Este proceso previo permite mejorar el resultado final del segundo paso, que hace uso de la librería de software libre Tesseract OCR, y que consiste en realizar el análisis y extracción del texto del documento. Como resultado final, el texto extraído se convierte a un formato estructurado JSON.

Este módulo, además, es el encargado de detectar el idioma en el que está redactado el documento.

Módulo NLP

En tiempo de ejecución, cada vez que se aporta un documento al expediente, se obtiene el JSON³ con todo el texto que



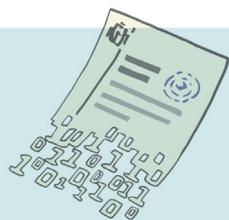
² **Binarización:** la mayor parte de algoritmos de OCR parten como base de una imagen binaria (dos colores). Por lo tanto, es conveniente convertir una imagen de escala de grises, o una de color, en una imagen en blanco y negro, de tal forma que se preserven las propiedades esenciales de la imagen. Una forma de hacerlo es mediante el histograma de la imagen, donde se muestra el número de píxeles para cada nivel de grises que aparece a la imagen. Para binarizarla tenemos que escoger un umbral adecuado, a partir del cual todos los píxeles que no lo superen se convertirán en negro y el resto en blanco.

Mediante este proceso obtenemos una imagen en blanco y negro donde quedan claramente marcados los contornos de los caracteres y símbolos que contiene la imagen. A partir de aquí podemos aislar las partes de la imagen que contienen texto.

[Fuente: Wikipedia]

Procesamiento OCR

El proceso de OCR puede parecer sencillo pero esta tecnología debe superar varios desafíos:



- Documentos mal escaneados (baja resolución, poco contraste, páginas con distintas orientaciones, etc.)
- Detección de párrafos y líneas
- Errores tipográficos por falta calidad imagen

- Líneas en tablas, caracteres especiales, etc.
- Documentos muy largos, con mucho texto (escrituras, presupuestos desglosados, etc.)

Es más, de todo el sistema construido, el módulo de OCR es el que podría obtener resultados menos óptimos, ya que depende en gran medida de la calidad de los documentos que se entreguen, hasta el extremo de que algunos de ellos son totalmente ilegibles, incluso para el ojo humano.



³ **JSON:** (acrónimo de «*JavaScript Object Notation*», en castellano «*Notación de Objeto de JavaScript*») es un formato de texto sencillo para el intercambio de datos. Se trata de un subconjunto de la notación literal de objetos de JavaScript, aunque, debido a su amplia adopción como alternativa a XML, se considera un formato independiente del lenguaje.

Una de las ventajas de JSON sobre XML como formato de intercambio de datos es que resulta mucho más sencillo escribir un analizador sintáctico («*parser*») para él.

[Fuente: Wikipedia]



contiene (más otros metadatos necesarios), por tanto, es el momento de aplicar la fase de inferencia.

«El nuevo sistema permite tratar los documentos recibidos para extraer automáticamente los datos de interés que alberga»

Con el fin de mejorar los resultados finales para cada tipología de documento, en la fase de entrenamiento, se ha construido un modelo de detección de entidades específico para cada uno de ellos. La secuencia de procesamiento para extraer los datos es la siguiente:

- Mediante el modelo de clasificación se obtiene la tipología de documento que se está tratando.
- Posteriormente, se le aplica el modelo de detección de entidades correspondiente para extraer la información que se requiere.

Los modelos se han construido con una arquitectura de redes neuronales («*Deep Learning*») diseñadas para modelar secuencias, llamada «*Transformers*» (<https://huggingface.co/models>) y se han usado los siguientes modelos base:

- LayoutLM. Que tiene en cuenta la posición y el contenido de los diferentes «*tokens*» de un documento.
- MBERT. Que tiene en cuenta el contexto de las palabras para aprender etiquetas.

Así, para cada tipología de documento, se ha seleccionado un modelo base concreto. Por ejemplo, para las facturas, en las que normalmente los datos del cliente y del emisor aparecen en una posición concreta (la parte superior), se ha aplicado el modelo LayoutLM, y para escrituras de compra-venta, el modelo MBERT.

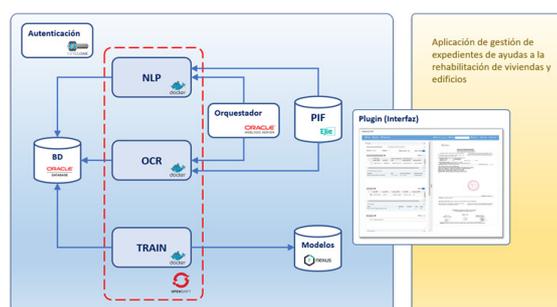
Módulo TRAIN

El objetivo principal de este módulo es dotar de capacidad de entrenamiento al sistema para:

- Reentrenar modelos existentes cuando haya documentos nuevos etiquetados
- Entrenar el modelo de clasificación para detectar nuevas tipologías
- Entrenar y producir los modelos de detección de entidades para nuevos tipos de documentos que se quieran incorporar

Puesto que el entrenamiento permite construir los modelos que posteriormente se usan para realizar la inferencia (extracción de información), para la puesta en servicio del sistema es necesario realizar al menos una vez este proceso de entrenamiento. Dado que estas tareas requieren mucho procesamiento y mucho tiempo (3-4 horas o incluso más), estas se han realizado en máquinas especializadas de EJIE que disponen de **GPUs** (procesadores gráficos con muchos más núcleos que la CPU y que facilitan el procesamiento paralelo).

En la construcción de los modelos se ha usado el aprendizaje supervisado, es decir, proporcionando al algoritmo documentos de ejemplo, en los que se han etiquetado los datos que se quieren obtener (p.ej. tipo de licencia, nombres...)

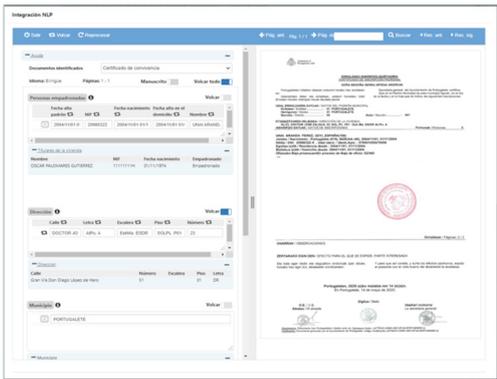


De esta manera, para cada tipología de documento, ha sido necesario definir qué datos en concreto se desean obtener, y después, etiquetar manualmente muchos documentos de ejemplo. Aunque, gracias al uso de «*transformers*» ese número se reduce frente a otros algoritmos y estrategias de «*Machine Learning*».

Plugin

Para facilitar a las personas tramitadoras el uso de la información extraída por el sistema, se ha construido un «*plugin*» que permite, entre otras funcionalidades:

- «*Localización*» de datos. Al pinchar sobre el campo en el que se visualiza el dato encontrado por el proceso de inferencia, en el documento que se presenta, se hace «*scroll*» automático hasta la página en la que se ha localizado dicho dato, resaltándolo para facilitar su localización.
- Rotado. Para el caso en que el documento no esté digitalizado en horizontal, se permite rotarlo.



- Carga manual de datos. Si el documento es manuscrito, se permite rellenar los datos, que no haya localizado el proceso.
- Reprocesado. Si el sistema no ha acertado con la tipología de documento al que pertenece, se permite marcar el tipo correcto, y volver a procesar.
- Correcciones. Se permite reutilizar el marcado de datos que pueda hacer la persona usuaria para regenerar el modelo de detección de entidades, permitiendo así mejorar la precisión del sistema.
- Volcado de datos en la aplicación usuaria.
- Búsqueda de texto en el documento.
- Zoom sobre el documento

De esta forma, la persona usuaria podrá comparar fácilmente si el dato que figura en su aplicación se corresponde con el que aparece en el documento recibido. Por ejemplo, que el certificado de empadronamiento se refiere al emplazamiento en el que va a realizarse la obra.

OpenShift y Docker

El desarrollo de los módulos de OCR, NLP y TRAIN se ha realizado en Python, y cada uno de los módulos está desplegado en contenedores *Docker*⁴, dentro de la infraestructura *OpenShift*⁵ de EJIE.

BENEFICIOS

Teniendo en cuenta que se han recibido 37.674 documentos, si descontamos los que son directamente procesables (solicitudes), las fotos de la obra (que no contienen texto), y un 10% de archivos con baja resolución, manuscritos, etc. para los que el proceso OCR podría tener problemas para extraer el texto, nos quedaríamos con 30.083 documentos procesables por el sistema.

Haciendo un cálculo aproximado, si el tiempo que una persona tramitadora necesita para leer y «trabajar» uno de esos documentos, con el nuevo sistema que se ha implantado, se reduce:

- 1 minuto, en ese mes y medio ahorramos 30.083 minutos = 501 horas.
- 4 minutos, en ese mes y medio ahorramos 90.248 minutos = 1.504 h.

Y si los traducimos en esfuerzo anual:

- Con una mejora de solo 1 minuto, en 1 año ahorramos 2.473 horas
- Con una mejora de 4 minutos, en 1 año ahorramos 7.419 horas

Incluso con estos números tan contenidos de, entre 1 y 4 minutos de ahorro de tiempo, nos permiten tener una idea de la reducción de esfuerzo al que se está llegando con el nuevo sistema. 



⁴ **Docker:** Docker es un proyecto de código abierto que automatiza el despliegue de aplicaciones dentro de «contenedores» de software, proporcionando una capa adicional de **abstracción y automatización** de virtualización de aplicaciones.

[Fuente: Wikipedia]



⁵ **OpenShift:** formalmente llamado «*OpenShift Container Platform*» (OCP), es un producto de computación **en la nube** de plataforma como servicio de Red Hat.

Red Hat OpenShift funciona con Kubernetes, lo que le permite ejecutar aplicaciones dentro de contenedores.

Los desarrolladores pueden usar Git para desplegar sus aplicaciones Web en los diferentes lenguajes de la plataforma.

[Fuente: Wikipedia]





⁶ **Ediciones anteriores:** para saber cómo se desarrollaron las dos primeras ediciones podéis consultar los siguientes artículos:

- «*Jornadas TEKgunea 1.0*» (boletín Aurrera nº 76, junio de 2021)
- «*Jornadas TEKgunea 2.0*» (boletín Aurrera nº 80, junio de 2022)

Jornadas TEKgunea 3.0

El pasado 26 de abril se celebró en Bilbao la tercera edición de las jornadas «TEKgunea» o, mejor dicho, «TEKgunea, Working Days-Egunak. Release 3.0».

La jornada TEKgunea 3.0, celebrada este año en el Museo Guggenheim de Bilbao, reunió a más de 200 personas relacionadas con los proyectos públicos de tecnología de Euskadi y Navarra. Dando ejemplo, la apertura fue realizada en colaboración, compartiendo escenario la Consejera de Gobernanza Pública y Autogobierno, Olatz Garamendi, y el Viceconsejero de Administración y Servicios Generales, Polentzi Urkijo.

Como en otras ocasiones, Xabier Arrieta, Director de la Dirección de Tecnologías de la Información y la Comunicación se encargó del papel de conducir y dinamizar toda la jornada.

LA JORNADA

La jornada tiene un formato ya definido en ediciones anteriores⁶: comienza con un repaso de los proyectos presentados en el año anterior para conocer su evolución, los objetivos logrados o comprometidos y ver si la reutilización y la colaboración sigue estando en su agenda.

Bloque 1. Evolución de proyectos de la edición anterior

Estos fueron los proyectos que «rendían cuentas» y que nos sorprendieron gratamente porque, como veréis todos «progresan adecuadamente».

Gonzalo R. Ordoñez desde Navarra (Nasertik) abrió la puerta del contenido de la

jornada hablando del proyecto de medicina personalizada de secuenciación genómica y supercomputación. Un proyecto con una hoja de ruta a muy largo plazo y en el que ya están colaborando Osakidetza y BioCruces.

Igor Goirizelaia Ziardegi, del Departamento de Salud del Gobierno Vasco, y Óscar Guadilla, de EJE, nos hablaron de cómo evoluciona «NIK Patrika Digitala», esa cartera electrónica que nos va a permitir quitarnos las tarjetas de plástico tradicionales del bolsillo. Un proyecto pionero en el Estado que nace y avanza en colaboración siguiendo las directrices europeas.

Ainhoa Anitua, de Izenpe, nos explicó que la opción de solicitar la BakQ desde casa evoluciona y este año se pondrá en la calle para toda la ciudadanía. Gracias a la identificación biométrica ya no será necesario hacer el trámite de forma presencial. En estos momentos, de un total de 15.000 emisiones mensuales, se hace ya de forma remota alrededor de un 25%.

Y cerró este bloque la Oficina de Experiencia de Usuario (UX) de la Diputación



El Museo Guggenheim de Bilbao fue la sede del TEKgunea 3.0.
[Foto: Museo Guggenheim Bilbao]



Olatz Garamendi, Consejera de Gobernanza Pública y Autogobierno del Gobierno Vasco. [Foto: Gobierno Vasco]

de Álava liderada por Andoni Martín, que aumenta equipo y volumen de proyectos poniendo Gustavo Echevarría como ejemplo práctico cómo convertir en un servicio usable una aplicación Cobol con visión del siglo XX.

Bloque 2. Micro Ponencias

Tras finalizar el bloque en modo *flash-back* volvimos al presente con los nuevos proyectos que se han puesto en marcha durante este año.

Como pequeña introducción de este bloque, contamos con la participación de Óscar Guadilla (EJIE) hablando de la importancia de la Innovación y Vigi-



Pero...

¿Qué es TEKgunea?

TEKgunea nace como una evolución natural de la Oficina de Software Libre que impulsó en su momento el Gobierno Vasco (conocida como «SALE», *Software Askea/Libre en Euskadi*). Busca **potenciar la colaboración** entre todas las administraciones públicas vascas y navarras a la hora de desarrollar proyectos tecnológicos, sobre todo los de tecnología emergente, aportando un punto de encuentro anual que busca un doble objetivo:

lancia Tecnológica como inversión de futuro, las cosas no son fruto de la casualidad, recorriendo ejemplos de distintas Pruebas de Concepto de las que 3 de cada 4 no llegan a Producción. Según Óscar, las personas no son innovadoras, sino que es la prospección de ideas que hacen las personas las que abren las puertas de la innovación.

Como segunda ponencia vivimos un gran ejemplo de colaboración entre Euskadi y Navarra. Ana López, del Departamento de Igualdad, Justicia y Políticas Sociales del Gobierno Vasco, y Javier Amezcua, de ITracasa, nos hablaron del proyecto Avantius⁷ que impulsa la visión de Justicia Digital innovadora en Euskadi. Un proyecto ya implantado en compromiso y plazo.

«TEKgunea 3.0 se celebró este año en el Museo Guggenheim de Bilbao y reunió a más de 200 personas relacionadas con los proyectos públicos de tecnología»

Y con visión de futuro, aún en fase -1, Alex Lara, de EJIE, nos habló de un proyecto que están analizando llamado «NIRE Card»,

- Dar visibilidad a proyectos tecnológicos con naturaleza de reutilización entre las citadas administraciones públicas.
- Fomentar el encuentro para el desarrollo de nuevos proyectos en colaboración entre distintas administraciones públicas, entre lo público y lo privado compartiendo: conocimiento, el talento de las personas y los recursos tecnológicos necesarios para disponer de más y mejores servicios digitales para la ciudadanía y nuestras empresas.



⁷ **Avantius:** para conocer los detalles del proyecto, podéis consultar el siguiente artículo:

- «El Expediente Judicial Electrónico en Euskadi» (boletín Aurrera nº 78, diciembre de 2021)



Memorias de Traducción:

para saber más sobre las memorias de traducción y sus características, podéis consultar los siguientes artículos:

- «Herramientas de asistencia a la traducción» (boletín Aurrera nº 12, diciembre de 2003)
- «IDABA: base de datos de traducciones» (boletín Aurrera nº 21, marzo de 2006)

con el que podremos tener un activo digital que nos permita disponer de dinero virtual inmediato y sin trámites obtenido de cualquier tipo de ayudas económicas de la Administración pública vasca.

A continuación, se recordó que en el año 2021 el Gobierno Vasco lanzó unas líneas de subvenciones para fomentar el desarrollo de Iniciativas relacionadas con el uso de Inteligencia Artificial (IA) y «Machine Learning» en la Administración Pública. Una de las empresas ganadoras, Legítimo Solutions, representada por Juan Bustillo, nos habló del caso práctico aplicado para la mejora en el flujo documental de la administración mediante IA. Visto el éxito de la convocatoria de 2021, este año 2023 se pondrá en marcha una nueva edición de esta subvención.

Iosu Uribe y Nekane Garitano, del IVAP, por su parte, hilando con otros proyectos de innovación, nos introdujeron en el problema de uso de datos para poder categorizar y sustituir en euskera y castellano. Una visión para compartir conocimiento sobre una problemática común a muchas administraciones y que han hecho uso de la IA para llevar a cabo la anonimización de datos en las Memorias de Traducción⁸.

Y como cierre, conocimos el proyecto de desarrollo regional «Data Spaces» europeos que van a aportar un modelo de dato como activo para la generación de actividad económica alrededor de la dinamización de servicios. La iniciativa BaiData

representa un modelo de colaboración pública-privada para toda la península con sede en Bilbao. Nerea Martiartu, Directora General de Lantik, y Eunáte Ramirez, de Versia, presentaron el proyecto como Vicepresidentas de BaiData.

Bloque 3. Ciberseguridad y colaboración interinstitucional para mejorar la resiliencia

El «postre» temático de la jornada venía de la mano de la ciberseguridad con una orientación de proyectos analizados desde distintas visiones y con una mesa redonda de cierre.

«TEKgunea busca potenciar la colaboración entre las administraciones públicas vascas y navarras para el desarrollo de proyectos tecnológicos»

Desde un punto de vista de la ciudadanía, tras 4 años de ideación, Manu Roibal desde BilbaoTIK nos presentó un proyecto único e innovador en el mundo que van a poner en Producción este año 2023. Un modelo de bienestar digital para la ciudadanía de Bilbao que se conecta a la red WiFi municipal. Algo tan sencillo y complejo como poder advertir, siempre con permiso explícito, a las personas que se conectan a la red de que su dispositivo



Ponencia de Josu Uribe y Nekane Garitano (IVAP).

[Foto: Gobierno Vasco]

está infectado o es inseguro. Una ciudadanía digital más segura permite mejorar nuestra ciberseguridad global en un proyecto con vocación de compartir conocimiento para su reutilización en otros municipios.



Xabier Arrieta (Dirección de Tecnologías de la Información y la Comunicación, Gobierno Vasco) y Oscar Guadilla (EJIE)

[Foto: Gobierno Vasco]

Con una visión también municipal, Ana Burgui, en representación de la entidad Animsa, nos abrió la puerta a la colaboración entre todos los municipios más pequeños de Navarra. Hablamos de municipios que no disponen de equipos o infraestructuras pero que sufren en igual medida incidentes o ciberataques. En colaboración y con un buen liderazgo se pueden mejorar sus capacidades de seguridad digital. Visión de presente y orientación a colaboración futura desde Navarra.

Y con una visión práctica orientada a la acción-reacción, desde Osakidetza nos hablaron de la gestión de contingencias. ¿Qué pasa cuando todo falla y debemos seguir dando servicios sanitarios a la ciudadanía? Un modelo que puede servir de ejemplo para la aplicación en otras problemáticas o sectores que no pueden parar un servicio público vital. En este caso, Maite Cuadrado nos aportó todos los detalles de su modelo con voluntad de seguir avanzando y de compartir conocimiento.

Y ya como cierre de la jornada, Javier Diéguez, Director del Basque CyberSecurity Center (BCSC) nos introdujo en un debate donde compartir experiencias, estrategias y opinión desde distintas visiones con la participación de: Óscar García

(EJIE), Imanol Sauto (Lantik) y Roberto Pérez de Arrilucea (CCASA). Echamos en falta a Miguel Ángel Hernández (IZFE) que finalmente no pudo asistir. En el ámbito de ciberseguridad tan importante es la **prevención** como la capacidad de **reacción** ante un incidente. Disponer de un Plan de Contingencia en estos casos siempre resulta vital pero la gran clave, según el Director del BCSC, y más en estos tiempos tan complejos que padecemos, es trabajar en equipo, colaborar,

compartir equipos y, por supuesto, conocimiento, y más teniendo en cuenta que cada día se incrementan los ciberataques y sus consecuencias.

TEKGUNEA 4.0

Según las encuestas de opinión recibidas, TEKgunea⁹ sigue despertando interés y se consolida como punto de encuentro anual para las personas con responsabilidad en los proyectos de tecnología de las administraciones públicas vascas y navarras. Desde la DTIC seguiremos empujando para que cada día surjan más proyectos en colaboración con visión de reutilización. Esto no sería posible sin vuestro apoyo y gracias a él ya estamos trabajando en la cuarta edición. Como dijo la Consejera, vamos a seguir manteniendo viva esa energía que surge de forma natural y que nos une denominada «*espíritu TEKgunea*». 



⁹ **TEKgunea**: si os perdisteis la jornada celebrada en Bilbao, tenéis publicados todos los vídeos y presentaciones en la siguientes página web:

<https://www.euskadi.eus/gobierno-vasco/tekgunea>



ALBOAN

La plataforma de eAdministración del Gobierno Vasco consigue la certificación ENS



El pasado mes de abril el Gobierno Vasco recibió la certificación ENS de su Plataforma de Administración Electrónica, en concreto, de los sistemas de información que soportan la Plataforma Tecnológica eAdministración (PLATEA): Web, Tramitación, Gestión Documental e Integración.

Ante estas premisas, el Gobierno Vasco decidió **certificarse en el ENS** para poder ofrecer un marco de **seguridad y confidencialidad** en los servicios que se ofrecen a la ciudadanía, y en las comunicaciones que se tienen con el resto de administraciones (locales, autonómicas o estatales), es decir, tener una política de seguridad que

¿QUÉ ES EL ENS?

Para la Administración Pública, el Esquema Nacional de Seguridad (ENS), según la definición publicada en su web, ofrece un marco común de principios básicos, requisitos y medidas de seguridad para una protección adecuada de la información tratada y los servicios prestados, con objeto de asegurar el acceso, la confidencialidad, la integridad, la trazabilidad, la autenticidad, la disponibilidad y la conservación de los datos, la información y los servicios utilizados por medios electrónicos que gestionen en el ejercicio de sus competencias.

Desde el pasado 5 de mayo de 2022, ha quedado derogado el Real Decreto 3/2010, en virtud de lo señalado por el RD 311/2022. No obstante, para facilitar una transición ordenada y la necesaria adecuación de las Entidades de Certificación al nuevo marco normativo, tal y como ha comunicado el CCN (Centro Tecnológico Nacional), los sistemas de información preexistentes a la entrada en vigor del RD 311/2022 podrán, hasta el 5 de mayo de 2024, seguir utilizando el procedimiento de certificación contra el RD 3/2010, sabiendo que la fecha máxima de validez de los certificados así emitidos no podrá superar el 5.5.2024. Por su parte, la certificación contra el RD 311/2022 podrá acometerse desde el 1 de diciembre de 2022 y las certificaciones tendrán la validez habitual de **dos años** naturales.



garantice la protección de la información tratada y de los servicios prestados a través de un planteamiento común de principios básicos, requisitos mínimos, medidas de protección y mecanismos de conformidad y monitorización para el sector público, así como los proveedores tecnológicos del sector privado que colaboran con la Administración.

Con ello se han conseguido los siguientes objetivos para todo el Catálogo de Servicios electrónicos que el Gobierno Vasco ofrece a la ciudadanía:

- Crear las condiciones necesarias de seguridad en el uso de los medios electrónicos
- Promover la gestión continuada de la seguridad
- Promover la prevención, detección y corrección
- Promover un tratamiento homogéneo de la seguridad
- Servir de modelo de buenas prácticas

NIVELES DE SEGURIDAD

El ENS prevé una serie de categorías para la seguridad de la información [ver Anexo I del Decreto 311/2022]:

- **Bajo:** Este nivel se utiliza cuando las consecuencias de un incidente de seguridad solamente afectan a alguna de las dimensiones de seguridad dentro de la empresa, pero no a todas. Es decir, este nivel supone un **perjuicio limitado** para las funciones de la organización, a sus activos o a las personas afectadas
- **Medio:** Este nivel se utiliza cuando las consecuencias de un incidente de seguridad afectan a alguna de las dimensiones de seguridad dentro de la empresa. Es decir, este nivel supone un **perjuicio grave** para las funciones de la organización, a sus activos o a las personas afectadas. [En nuestro caso, los sistemas de in-



formación certificados por el Gobierno Vasco son todos ellos de categoría MEDIA]

- **Alto:** Este nivel se utiliza cuando las consecuencias de un incidente de seguridad afectan a muchas de las dimensiones de seguridad dentro de la empresa. Es decir, este nivel supone un **perjuicio muy grave** para las funciones de la organización, a sus activos o a las personas afectadas.

PRÓXIMOS PASOS

Dado que el ámbito de la seguridad está en continua evolución, el Gobierno Vasco tiene previsto, entre otros, realizar los siguientes pasos:

- Progresar y alinearse con el marco legal y estratégico actual para facilitar la seguridad de la administración.
- Introducir la capacidad de ajustar los requisitos del Esquema Nacional de Seguridad a determinados colectivos o ámbitos tecnológicos.
- Revisar de forma detallada los principios básicos, requisitos mínimos y medidas de seguridad.
- Agregar un nuevo sistema de codificación de los requisitos de las medidas basado en refuerzos alineados con el nivel de ciberseguridad perseguido.

Una vez conseguido este certificado, a partir de aquí, el nuevo objetivo es conseguir la certificación en el denominado **ENS2** o **ENS2022**, tarea que el Gobierno Vasco tiene previsto abordar en el primer trimestre de 2024.



NORMATIVA

- **Real Decreto 3/2010**, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica.
- **Real Decreto 951/2015**, de 23 de octubre, de modificación del Real Decreto 3/2010.

Normas derogadas por la disposición derogatoria única del **Real Decreto 311/2022**, de 3 de mayo, por el que se regula el Esquema Nacional de Seguridad.



Real Decreto 311/2022:
<https://www.boe.es/buscar/act.php?id=BOE-A-2022-7191>

REVISTA «KOMUNIKT» DE OSAKIDETZA

La Subdirección de Informática y Sistemas de Información de Osakidetza elabora y publica la revista «*KomunIKT*».

Esta revista, que tiene como lema «*Osakidetza hacia la Salud Digital*», tiene como objetivo dar a conocer los Sistemas de Información de todas las áreas que forman parte del Servicio Vasco de Salud, las cuales permiten ofrecer un mejor servicio a los ciudadanos y ciudadanas de Euskadi.

La revista se elabora gracias a la colaboración del personal de las distintas OSIs (Organizaciones Sanitarias Integradas) que componen Osakidetza, con lo que forma parte de un proyecto colaborativo y de trabajo en red. Se edita un ejemplar al año.

El pasado mes de octubre se publicó el último ejemplar, el cual incluye artículos tanto en euskera como en castellano y consta de 68 páginas.

En esta ocasión algunos de los temas tratados son:

- Big Data y Gobierno del Dato en Osakidetza
- Web de Procesos Selectivos
- Área de ciberseguridad
- Mejorando las competencias digitales de los profesionales sanitarios para generar valor
- App donantes de sangre de Euskadi

La revista está disponible en Internet.



Más información en:

<https://komunik.osakidetza.eus>



XIA PEISU, LA MADRE DE LA INFORMÁTICA CHINA

Xia Peisu (Chongqing, 28/julio/1923 – Pekín, 27/agosto/2014) fue una pionera en el ámbito de la informática china.

En 1945 Xia obtuvo una licenciatura en ingeniería eléctrica. Dos años después, comenzó un doctorado en ingeniería eléctrica en la Universidad de Edimburgo, y en 1950 recibió su doctorado. Poco después, se casó con el científico Yang Liming y ambos regresaron a China.

En 1952, el matemático Hua Luogeng inició el desarrollo del primer computador electrónico de China, para lo cual reclutó a Xia y a otros dos científicos. Una vez que los otros dos científicos dejaron el proyecto, Xia se encargó de desarrollar el primer computador electrónico multipropósito elaborado en China, el Modelo 107, en 1958.

En marzo de 1956, Xia impartió el primer curso de China en teoría de la computación. Cuando la Universidad de Ciencia y Tecnología de China fue fundada en 1958, Xia se encargó de su Departamento de Informática, formando a más de 700 estudiantes entre 1956 y 1962, por lo que es conocida como la «madre china de la computación».

En 1978, fundó el «*Chinese Journal of Computers*», así como el «*Journal of Computer Science and Technology*», única revista de lengua inglesa en el campo computacional publicada en China, en 1986.

En 1991, Xia Peisu y su marido Yang fueron elegidos miembros de la Academia China de las Ciencias.



Foto de Xia Peisu

[Imagen: Wikimedia]



Más información en:

[Wikipedia y](#)

<https://mujeresconciencia.com>

