



Ibercaja Banco

INYCOM DEPURA LAS BASES DE DATOS DE DNI TRAS LA FUSIÓN DE IBERCAJA Y CAJA3, CON TECNOLOGÍA DE MULTITEC

PERFIL DEL CLIENTE

Ibercaja ha integrado Banco Grupo Caja3 (constituido por Caja Inmaculada, Caja Círculo y Caja Badajoz) formando un grupo sólido, solvente y competitivo, que da respuesta a las exigencias que impone el nuevo entorno económico y regulatorio.

Con la nueva dimensión, se configura un Banco más grande, más diversificado, con un claro liderazgo de negocio en Aragón, Burgos, Badajoz, Rioja y Guadalajara, y con una importante presencia en otros territorios clave, como Madrid, Cataluña y Levante.

Las fortalezas que definen al Grupo son la solvencia financiera, sus sólidos arraigos territoriales y su filosofía de gestión, basada en la honestidad, la excelencia y la cercanía. La entidad se denomina Grupo Ibercaja, marca que se incorpora visualmente en las oficinas de Caja Inmaculada, Caja Círculo y Caja Badajoz. Con esta estrategia se preserva el arraigo y valor de esas identidades centenarias en sus territorios, que suman ahora la fortaleza que confiere integrarse en el Grupo Ibercaja.

La sede social y los servicios centrales están en Zaragoza y posee direcciones territoriales en Burgos y Badajoz. Las cajas fundadoras de las entidades financieras originarias, accionistas ahora de Ibercaja Banco, no se fusionan y siguen existiendo de forma independiente convertidas ahora en fundaciones.



Durante el pasado mes de abril, Inycom llegó a un acuerdo con IberCaja y Multitec para prestar el servicio de depuración de las Bases de Datos procedentes de DNI de clientes de IberCaja y Caja3. El proyecto ha finalizado con éxito y ha permitido detectar duplicados, actualizar información e identificar otras casuísticas comerciales.



SITUACIÓN PRESENTADA

En 2013 se anunció el proceso de fusión mediante absorción de Caja3 por parte de Ibercaja. Mediante esta operación, Ibercaja Banco pasó a ser la novena entidad por volumen de negocio, aglutinando a 3,3 millones de clientes.

Con anterioridad a la fusión, cada una de las entidades ya había implantado el proceso de digitalización del DNI (o tarjeta de extranjería) de sus clientes y sus datos identificativos se informatizaban de manera independiente en las distintas aplicaciones informáticas de gestión puestas a disposición del personal de la entidad.

Muchos de estos clientes disponían de cuentas tanto en Ibercaja como en Caja3 lo que, al plantearse el proceso de fusión, suscitó la necesidad de depurar y actualizar las bases de datos de ambas entidades con el fin de evitar duplicidades, inexactitudes, y datos obsoletos.

Inycom llevó a cabo el proyecto de normalización de las bases de datos de la entidad financiera. Para ello, contó con la colaboración de la empresa tecnológica Multitec que se encargó de llevar a cabo el reconocimiento masivo de datos a partir de las imágenes de DNI procedentes de las citadas entidades bancarias.

EL PROYECTO

Dado que las imágenes digitales de los DNI habían sido generadas por multitud de empleados de las cajas con distintos modelos de escáner, su calidad (en cuanto a resolución y legibilidad) y formato variaban mucho entre ellas, imposibilitando en muchos casos la aplicación de la tecnología de reconocimiento automático de texto. Asimismo se detectó la existencia de disparidad de criterios en los datos preexistentes, de manera que nombres compuestos como «María José», podían aparecer como M. José o M^a José.

Para llevar a cabo el reconocimiento automático de textos, Inycom contó con la colaboración de Multitec que desarrolló una aplicación informática específica que combinaba un potente sistema de OCR especialmente adaptado para reconocimiento de textos de caracteres DNI y un ágil sistema de grabación manual de datos a partir de la imagen reproducida en la pantalla del ordenador al que se recurría cuando el nivel de fiabilidad del OCR no era el deseado.

Dicha aplicación incorpora un sistema de normalización de datos para garantizar la homogeneidad de criterios a partir de bases de datos preexistentes que contengan nombres, apellidos, sexo en función del nombre, nombres de localidades, etc..

Aunque la cadena de caracteres que figura en el reverso de los DNI actuales está especialmente diseñada para su reconocimiento por OCR, presenta limitaciones en cuanto a la longitud de los nombres de los titulares (cortándolos al superar el límite de la longitud de campo) y además no admite determinados caracteres y símbolos como «Ñ», diéresis o guiones. Por estos motivos se estableció como requisito necesario obtener la información a partir del texto impreso en el anverso del documento (más completo y exacto) cuyo reconocimiento automático mediante OCR resultaba más problemático por el tamaño y formato de la fuente.

La distribución de esta información en el DNI ha ido variando a lo largo del tiempo y, por ello, era preciso procesar tanto los formatos más actuales del documento como los más antiguos, lo que suponía otra traba al proceso de automatización del proceso.

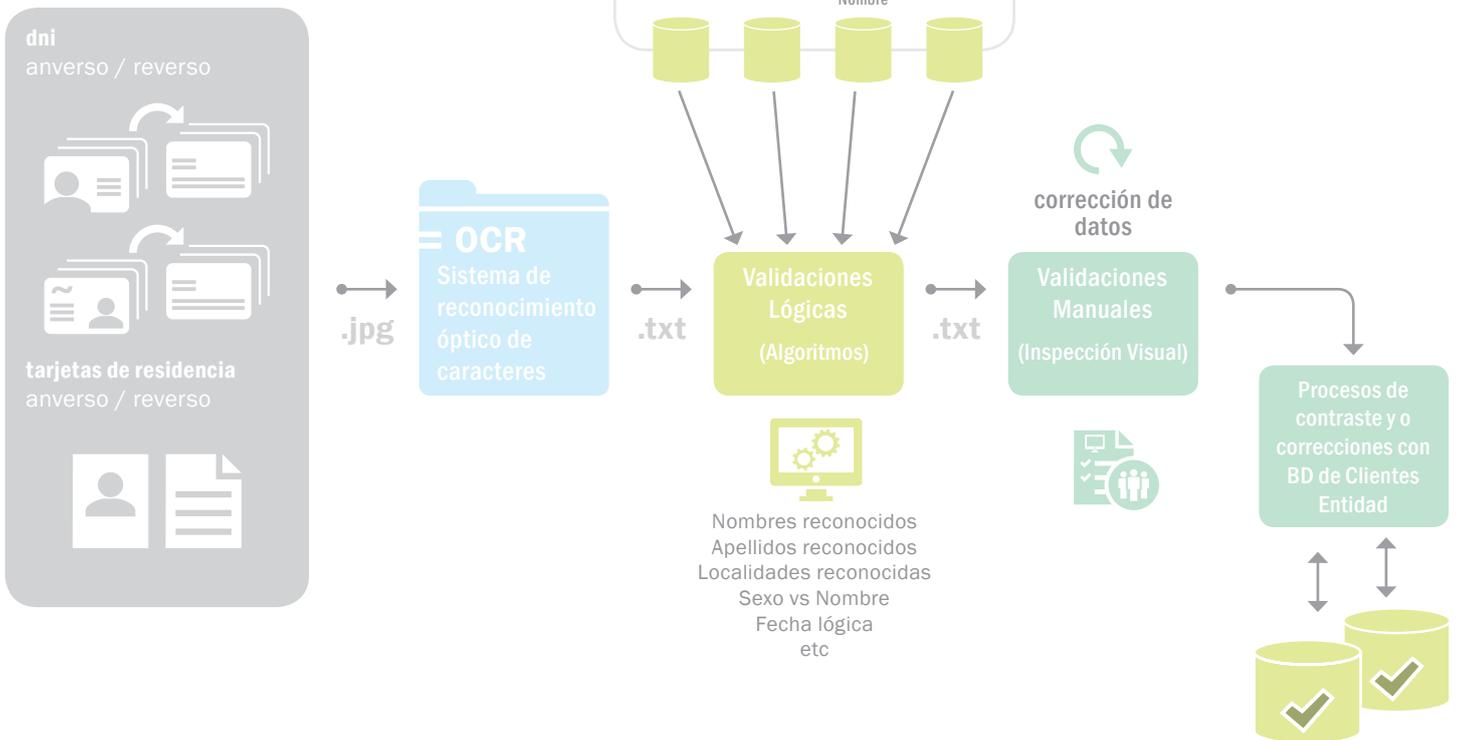
Finalmente, aunque el OCR arrojó excelentes resultados en un grupo muy reducido de documentos, fue necesario recurrir a un proceso de grabación manual del resto de las imágenes. El sistema de grabación manual incluía reglas de control que evitara que el operador pudiese introducir símbolos o números en campos de texto alfabéticos, que no se respetase la longitud del campo preestablecida o que quedasen campos en blanco.

Para garantizar la calidad de los datos entregados al cliente, se procedió a la verificación humana del 100% de los dígitos de cada DNI (tanto de los reconocidos por el OCR como de los capturados manualmente).

A partir de esta información depurada, actualizada y fiable, Inycom procedió a actualizar y normalizar las bases de datos de su cliente y a generar registros de los clientes de cuyos documentos no existía copia.

Imágenes escaneadas en oficina

Documentos identificativos de clientes



Rev. 001 septiembre 2014

